[Course Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live)

[Course Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live)

[Hi, my name is Vitthal Srinivasan, and I'd](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=0) [like to welcome you to this course on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=7.154444444444445) [architecting Production-ready ML Models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=9.1288) [Using the Google Cloud ML Engine. A little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=11.616) [bit about myself, I have master's degrees](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=15.142000000000001) [in Financial Math and Electrical](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=17.558) [Engineering from Stanford University and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=19.378) [have previously worked in companies such](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=21.682) [as Google in Singapore and Credit Suisse](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=23.493999999999996) [in New York. I am now co-founder at](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=25.96799999999999) [Loonycorn, a studio for high-quality video](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=29.161) [content based in Bangalore, India.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=32.025) [Building machine learning models using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=34.531) [Python and an ML framework is the first](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=36.162) [step towards building an enterprise-grade](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=39.369) [ML architecture. But two key challenges](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=41.744) [still remain, training the model with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=44.633) [enough compute firepower to get the best](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=47.097) [possible model and then making that model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=49.4) [available to users who are not data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=52.045) [scientists and may not even be Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=54.573) [users. You will gain the ability to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=57.26362499999999) [perform on- cloud distributed training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=59.184000000000005) [hyperparameter tuning, as well as to make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=61.68) [your ML models available for use in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=64.41) [prediction via simple HTTP requests.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=66.812) [First, you will learn to use the ML engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=69.83) [for models built in XGBoost. XGBoost is an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=72.532) [ML framework that utilizes a technique](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=75.81450000000001) [known as ensemble learning to construct a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=77.74433333333333) [simple, strong model by combining several](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=80.88066666666667) [weak learners as the unknown. XGBoost is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=83.37114285714286) [very popular in the Kaggle community, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=86.13528571428573) [it makes a lot of sense for use on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=88.0168888888889) [cloud because it features a fast and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=90.89666666666665) [reliable training algorithm that converges](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=93.3508) [quickly. On the cloud, time is money, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=95.6584) [so the speed of training is important.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=98.41288888888889) [Next, you will discover how easy it is to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=100.53742857142856) [port serialized models from on premise to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=103.34150000000001) [the GCP. You will build a simple model in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=106.04880000000001) [scikit-learn, which is the most popular](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=108.428) [classic ML framework, and then serialize](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=110.396) [that model and port it over for use on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=113.415) [GCP using ML Engine. Finally, you will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=115.769) [explore how to tap the full power of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=118.70262499999998) [distributed training, hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=121.33879999999999) [tuning, and prediction in TensorFlow,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=122.786) [which is the most popular library for deep](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=125.01033333333334) [learning applications right now.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=128.323) [TensorFlow was the first library that ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=129.8276666666667) [Engine supported, and both originated at](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=132.5812) [Google, so the integration between these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=135.34959999999998) [two technologies is very tight. You will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=137.1846666666667)[see how a JSON environment variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=140.1312857142857) [TF\_CONFIG is used to share state](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=142.34142857142857) [information and optimize the training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=144.66383333333334) [hyperparameter tuning processes. Once](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=147.60424999999998) [you're finished with this course, you will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=149.50400000000002) [have the skills and knowledge of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=151.00924999999998) [Google Cloud ML Engine needed to get the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=153.061) [full benefits of distributed training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=155.7805) [make both batch and online prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=158.32742857142856) [available to your client apps via simple HTTP requests.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=d59a2402-b242-4693-aed9-9e6780fe7eb7&clip=0&mode=live&start=160.484)

[Introducing Google Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live)

[Module Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live)

[Hello and welcome to this course on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=0) [Architecting Production-ready ML Models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=4.089) [Using the Google Cloud ML Engine. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=7.589) [module is all about introducing the Google](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=11.365857142857143) [Cloud ML Engine, which is a managed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=13.467500000000001) [service for enterprise-scale use of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=16.005) [machine learning on the Google Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=18.845000000000006) [platform. There are two services offered](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=21.79966666666667) [within the Cloud ML Engine. The first is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=23.73985714285714) [training, including both local and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=26.757571428571435) [distributed training, and the second is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=28.771000000000004)[hosted prediction on the cloud. ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=30.86966666666666) [was originally developed for supporting](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=34.07333333333333) [TensorFlow models. It has since been](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=36.336999999999996) [extended to support XGBoost and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=38.811571428571426) [scikit-learn as well. In addition, Keras](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=40.5834) [models can also be implemented, provided](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=43.12928571428571) [they use TensorFlow as the back end. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=45.417399999999994) [have a packed agenda with a large number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=48.280555555555566) [of demos coming up, so let's plunge right in.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=0&mode=live&start=50.353500000000004)

[Prerequisites and Course Outline](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live)

[Let's quickly spend some time talking](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=2.146) [about the prerequisites and the course](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=3.749714285714286) [outline. This course is very focused on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=6.4111666666666665) [the mechanics and the workings of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=8.774375) [Cloud ML Engine, and so it's definitely](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=10.618) [not a starter course on machine learning.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=13.297249999999998) [It's highly recommended that you have a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=16.384) [basic understanding of machine learning,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=18.02) [TensorFlow, and scikit-learn. It will also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=20.324) [help if you understand how cloud computing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=23.592624999999998) [broadly works. To give you a sense of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=26.495) [kind of background that you require, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=29.201111111111103) [take a topic like hyperparameter tuning.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=30.89628571428572) [This course will not spend a whole lot of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=34.103) [time explaining what hyperparameter tuning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=35.879) [is or why we should perform it; rather, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=38.127) [will just focus on how this is implemented](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=41.85) [on the Cloud ML Engine. If you feel that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=43.97016666666667) [your background on such topics is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=47.759499999999996) [inadequate, don't worry. There are plenty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=49.663799999999995) [of great courses here in the Pluralsight](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=51.94071428571429) [catalog. A couple worth discussing are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=54.56549999999999) [Understanding the Foundations of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=57.176500000000004) [TensorFlow and another called Building ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=59.13399999999999) [Models in Python with scikit-learn.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=62.15157142857142)[Likewise, this course is also not meant to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=64.58840000000001) [be a starter course on the Google Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=67.54422222222225) [platform, so it's important that you](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=70.00474999999999) [understand how Google Compute Engine and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=71.75828571428569) [Google Cloud Storage work. There are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=74.17266666666667) [courses on these topics on the Pluralsight](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=76.56885714285715) [catalog, Choosing and Implementing Google](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=79.4588) [Cloud Compute Engine solutions, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=81.32700000000001) [Architecting Google Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=84.01833333333333) [Configurations. Let's quickly outline the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=86.67166666666664) [flow of this course. We will begin by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=88.49500000000002) [discussing the theoretical concepts behind](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=91.04357142857143) [Cloud ML Engine in detail. We will spend a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=92.98928571428571) [lot of time talking about ML Engine's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=95.95533333333333) [pricing, which is notoriously opaque and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=97.83966666666667) [hard to understand. We will then discuss](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=100.90100000000001) [deploying models in each of the three](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=103.80157142857142) [frameworks, XGBoost, scikit-learn, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=107.323) [then TensorFlow. Please note, by the way,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=109.7935) [that Keras is supported on the Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=112.157) [Engine, but this is only with a TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=114.85966666666664) [back end, and that's why there's no need](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=116.72224999999996) [for us to separately discuss Keras.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=119.22700000000002) [Everything that we say about TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=122.071) [applies to Keras as well. All of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=123.964)[scenarios in this course are based in a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=127.184) [hypothetical online retailer called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=129.7105) [SpikeySales.com. As the name of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=132.64624999999995) [company would suggest, it specializes in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=134.59487499999994) [flash sales of trending products. As a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=136.69599999999997) [result, it is susceptible to spikes in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=139.69844444444442) [user traffic. This makes it an excellent](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=141.807) [candidate for moving to the cloud because,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=144.699) [of course, it does not need to provision](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=146.91) [excess compute capacity when the spikes are not taking place.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=1&mode=live&start=149.199)

[Introducing Google Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live)

[Let's start by introducing Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=1.967) [Engine. This is a managed service which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=4.560714285714286) [helps us to build and deploy ML models on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=6.556999999999999) [the Google Cloud platform. To understand](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=9.631750000000002) [clearly what this means, let's study how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=12.00266666666667) [Machine learning is actually implemented](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=14.586) [in enterprises. The first step is data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=17.105285714285714)[preparation. A whole bunch of data, some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=20.657166666666676) [structured, some unstructured, is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=23.190571428571428) [collected and cleaned and accessible in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=24.849999999999994) [something like a data warehouse. Then that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=27.447000000000003) [data is fed into a model. This is often](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=30.325) [something built by a data scientist. So at](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=33.848666666666674) [this point, we are still pretty firmly in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=36.8365) [the lab. The next step involves model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=38.97975000000001) [training and evaluation, and the reason](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=41.85575) [this box is colored a little differently](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=44.23966666666667) [is that this is at the cusp. It lies](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=45.962875)[somewhere between R&D use in a lab and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=49.502250000000004) [full-scale production use. This work is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=53.20419999999999) [still very likely being undertaken by a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=55.92828571428571) [data scientist, but the outcome of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=58.471799999999995) [step is going to be fed into production.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=60.471499999999985) [And the first step is, of course,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=63.597) [deployment. Once a model has been](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=66.18524999999998) [deployed, it will then be put to use for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=67.92333333333336) [actual production by other users within](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=70.1392222222222) [the organization or outside it, and over](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=72.22283333333334) [time, more versions of the model will be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=75.12471428571432)[developed, and therefore, this production](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=77.332) [system needs to have a way to manage those](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=79.7115) [versions. The first two steps of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=83.07299999999998) [workflow will rely heavily on external big](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=85.14049999999997) [data solutions, on Python libraries such](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=87.43183333333333) [as pandas and NumPy, and of course on ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=89.90283333333332) [frameworks such as TensorFlow, Keras, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=93.482) [so on. So in the context of this course,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=96.1) [we are focused on the Google Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=98.687125) [Engine. The frameworks supported are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=101.09516666666667) [TensorFlow, Keras, XGBoost, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=103.226) [scikit-learn. Let's spend a minute](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=106.57799999999997) [understanding their relative strengths and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=108.62) [weaknesses. TensorFlow is the most popular](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=110.83824999999999) [deep learning library out there right now,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=113.21116666666666) [and it is also most tightly coupled with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=115.923) [ML Engine. It's relatively complex to use,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=118.52779999999996) [it's relatively slow to train, but on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=121.991) [flipside, it does give an incredible](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=124.643) [amount of power and flexibility. Keras is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=127.03614285714283) [also a deep learning framework. The great](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=131.14857142857142) [bit about Keras is that it allows a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=133.53111111111113) [high-level way of interacting with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=135.65116666666665) [back-end frameworks, of which TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=137.657) [is one. So please note that when we use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=140.49271428571436) [the term Keras in the context of the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=143.27899999999994) [ML Engine, we explicitly mean Keras with a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=145.37325000000004) [TensorFlow back end. Support for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=148.407) [scikit-learn was added relatively recently](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=151.372) [to Cloud ML Engine. Scikit-learn is a very](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=153.44766666666666) [popular traditional ML package. Here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=157.12033333333338) [traditional is equivalent to non-deep](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=159.571) [learning. Scikit-learn is very easy to use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=162.52633333333335) [and integrates tightly with Python's other](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=164.68828571428574) [libraries. And the fourth framework](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=168.71183333333332) [supported on ML Engine is XGBoost. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=170.60542857142855) [by far the least popular, but it has some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=173.39312500000005) [important advantages for use on the cloud.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=176.82357142857143) [XGBoost is great for fast prototyping and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=179.968) [training, and this is important in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=182.76)[cloud because long-running training jobs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=185.044) [can run up huge bills. So you might](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=187.528) [seriously want to consider XGBoost for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=190.22300000000004) [your use cases. It's also particularly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=192.51260000000002) [well suited to small datasets and handles](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=195.42799999999997) [missing data well. Each one of these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=198.375) [frameworks has its own internal](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=201.10250000000002) [abstractions for a model. So when we use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=202.802) [the term model in the context of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=206.01533333333336) [TensorFlow or scikit-learn or Keras, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=207.756) [mean specific types of model objects which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=210.11599999999999) [exist in those frameworks. The next](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=213.587) [important step here is the training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=216.31442857142855) [evaluation step. Training and evaluation](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=219.44349999999997) [both have very specific meanings in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=222.02349999999998) [world of machine learning. Let's just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=224.194) [recap those. A trained model is one in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=226.79950000000002) [which the model parameters have been](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=229.98133333333325) [optimized to fit the training data. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=232.02499999999998) [training refers to that process of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=234.91728571428573) [optimizing those model parameters. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=237.928) [complex machine learning models, there can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=240.6094) [be thousands or even tens of thousands of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=242.32725) [parameters, so optimizing, or finding the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=244.66959999999997) [best values of their coefficients is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=246.8581428571429) [nontrivial. Let's say you've built a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=249.73899999999998) [model. If you decide to train it on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=251.97116666666668) [cloud, you might get dramatically better](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=254.31477777777778) [performance and production than if you](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=256.073) [train it locally. Another benefit of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=257.662)[performing these steps on the cloud is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=260.274) [that model evaluation is a lot easier as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=262.62687499999987) [well. This, again, has a specific meaning.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=266.02025000000015) [Evaluation refers to the process of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=268.585) [tweaking the hyperparameters, that is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=270.349) [model properties, rerunning a completely](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=272.901) [independent training process for each set](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=275.96) [of hyperparameters, and then selecting the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=278.622) [best-trained model after evaluating all of those candidates.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=2&mode=live&start=281.078)

[ML Engine and ML Frameworks](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live)

[So far, we were focused on the role of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=2.464) [framework and the data scientist. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=3.932) [now turn our attention to the ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=6.5365) [and how it interplays with the ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=9.097222222222223) [framework. As we shall see shortly, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=12.297777777777782) [Cloud ML Engine offers two services,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=14.561749999999996) [training and prediction. Training in turn](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=17.415428571428574) [can be performed either locally or in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=20.020285714285713) [distributed fashion on machines run by the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=22.567) [Cloud ML Engine. And prediction can also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=25.775428571428566) [be of two types, batch prediction, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=28.349) [seeks to maximize throughput, and online](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=30.450499999999998) [prediction, which seeks to minimize](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=33.13549999999999) [latency. Let's focus for now on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=36.006600000000006) [left-hand side of this tree, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=38.40224999999999) [understand the role of ML Engine in model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=40.164666666666655) [training and evaluation. This is a step](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=43.424499999999995) [which requires particularly close](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=46.413571428571416) [interaction between ML Engine and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=47.74675) [framework. The developer has a role to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=50.469249999999995) [play here as well in configuring the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=52.363666666666674) [training and the hyperparameter tuning,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=54.221142857142866)[but it's ML Engine which will have to take](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=56.086) [those configuration commands and go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=59.37142857142856) [and execute them on special machines on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=61.3495) [the cloud. This is relatively simple for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=64.2768) [scikit-learn and XGBoost, but is very](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=67.08849999999998) [complex for the deep learning frameworks](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=68.78450000000001) [like TensorFlow. The process of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=71.62460000000002) [optimization used in deep learning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=73.93028571428573) [frameworks is quite complex, and model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=75.6346) [parameters need to be held in a special](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=77.98811111111111) [kind of shared memory space called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=80.426)[parameter servers. All of the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=83.144) [instances need to be able to communicate](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=86.065) [with these parameter servers and update](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=88.668) [the values of those parameters at the end](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=90.49542857142853) [of each epoch. From this point on in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=92.99450000000003) [ML workflow, ML Engine takes over. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=95.948) [deployment, managing prediction and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=99.24380000000001)[managing versions, all of these are pretty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=100.979) [much ML Engine's forte. However, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=103.348) [crucial step here is the handshake between](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=107.09614285714287) [ML Engine and the ML framework. This needs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=109.58625) [to be orchestrated very precisely;](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=112.94733333333332) [otherwise, the distributed training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=114.8935) [deployment processes will fail. There are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=117.15060000000001) [four steps in this handshake. First, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=119.752) [model needs to be packaged correctly.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=122.27399999999999) [Second, ML Engine needs to be able to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=124.395) [stage the model; that is place it in some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=126.41099999999996) [GCS location from where it will be used](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=129.36344444444444) [during the training process. Third, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=131.73450000000005) [model needs to be exported so that it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=134.1145) [becomes available for prediction, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=136.84399999999997) [finally, the model can actually be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=138.3226666666667) [deployed. This is where ML Engine can take](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=141.0653333333333) [over using the saved or serialized model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=143.5857777777777) [Let's now understand precisely the roles](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=146.357) [of the developer and of ML Engine in this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=148.55728571428577) [handshake. Let's start with the role of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=152.80311111111112) [the developer. While packaging the model,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=154.94774999999998) [the developer simply has to use the right](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=156.938) [directory structure and then invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=159.138) [gcloud. Gcloud will go ahead and create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=161.96716666666666) [the Python package. This is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=164.25500000000002) [recommended way; the developer should not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=166.0634) [try to build the package herself. The next](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=167.86050000000006) [step is staging. Here, the developer must](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=170.7077142857143) [specify the staging location while](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=173.10766666666672) [invoking gcloud. This is where the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=175.56820000000002) [training package will be available for ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=177.613) [Engine to pick up. This will be a location](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=179.66728571428567) [in a Google Cloud Storage bucket. The next](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=182.36211111111118) [step is export, and here the developer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=184.75271428571432) [needs to include code to serialize the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=187.1358571428571) [trained model as well as the parameters](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=189.27999999999997) [inside the Python trainer package. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=191.828875) [developer also needs to specify the export](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=194.24599999999998) [location while invoking gcloud, and it's a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=196.3132) [great practice to save checkpoints along](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=199.29) [the way in case the VM instances running](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=201.0795) [the training process crash. And once all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=204.184) [of this is done, it's pretty easy to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=206.6824285714286) [deploy the model. All that we need to do](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=208.74871428571427) [is to tell ML Engine where the saved model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=211.1423636363636) [is, create a model name, and assign a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=213.51099999999997) [version number. Most of the heavy lifting](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=216.533) [is done for us by ML Engine, so let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=218.5505) [understand the platform's role. In the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=221.04133333333328) [packaging step, gcloud will actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=223.27900000000002) [create the Python package for us. Staging](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=225.04300000000003) [of the model is done for us as well. ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=227.42949999999996) [Engine is going to copy the trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=230.367625) [package to all the training instances.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=232.094)[Once the code runs, ML Engine will export](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=234.161) [the model to the location that we had](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=236.921) [specified. And finally, when it comes to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=239.32633333333337) [deployment, ML Engine takes over](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=241.33357142857147) [completely. It hosts the model and makes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=243.12679999999995) [it available to accept HTTP REST API](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=245.23875000000007) [requests. Hopefully by this point, we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=248.7795) [a clear sense of the exact handshake](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=250.76549999999997) [between the ML framework, the developer,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=253.30187500000002) [and the ML Engine. It's important for us](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=255.71885714285713) [to be very clear on what exactly we need](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=258.3124444444444) [to do and what information we need to supply the framework.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=3&mode=live&start=260.59850000000006)

[Python Package Structure](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live)

[This is probably also a good place for us](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=2.456) [to talk about the standard directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=4.494) [structure into which we have to organize](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=6.544) [our projects. It's important for us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=8.904) [understand the structure and follow it;](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=11.328) [otherwise, gcloud will have a hard time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=13.496) [creating our trainer package. Remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=15.433) [that all of this is happening inside our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=17.780142857142856) [project folder. There we should have a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=20.302) [directory, which is our trainer directory,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=23.185000000000002) [and inside this, we should keep all of our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=25.454)[model code. This example structure here is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=27.361) [for TensorFlow, where we need to have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=30.621) [separate task.py and model.py files. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=32.756) [addition, each package folder needs to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=37.36671428571429) [have a file called init.py. This is going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=39.903) [to indicate to the Python package builder](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=43.84300000000001) [what folders should be included in this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=45.495999999999995) [package. Finally, one level up, that is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=48.44) [outside our trainer package, we need to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=51.279) [include a file called setup.py. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=53.7775) [used by Python's setup tools package to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=56.97571428571428)[locate all of the dependencies and place](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=59.80350000000001) [them inside the trainer package. So this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=61.948499999999996) [is the complex directory structure which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=64.80528571428572) [we need to follow if we would like to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=67.1722857142857) [build our own package. But the good news](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=70.3033333333333) [is we don't really need to do much of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=72.7242222222222) [ourselves. All we need to do is to use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=75.85700000000003) [gcloud to build our packages. As long as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=78.02136363636367) [we have organized our code, that is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=80.62612500000003) [task and the model.py in this directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=83.502) [structure, gcloud will do the rest. It](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=86.1197142857143)[will create the setup.py and package it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=88.96987500000002) [all up into a trainer package. In the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=91.66250000000001) [demos up ahead, we will demonstrate the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=94.40125) [creation of a setup.py, just in case. But](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=96.56949999999999) [again, gcloud is the recommended way of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=99.899125) [doing this. We should not try and do this manually ourselves.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=4&mode=live&start=102.259)

[Training and Deploying Models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live)

[Let's talk a little bit about the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=2.334) [mechanics of training and deploying models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=3.840750000000001) [on the Cloud ML Engine. Let's pick up from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=5.3988571428571435) [where we left off. We were discussing the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=8.280000000000001) [use of gcloud to build packages, and this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=10.3905) [happens right when we use gcloud to invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=13.29985714285714) [the Cloud ML Engine training service. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=16.362444444444442)[service is going to take our training code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=19.394875) [and go ahead and execute it, either](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=21.701999999999998) [locally or on Google Compute Engine VM](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=24.50842857142857) [instances. In either case, this training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=27.5902) [code is going to access the training data,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=30.256) [which is going to be in a Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=32.254333333333335) [bucket, and it's going to go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=34.253750000000004) [save the model as well as checkpoints and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=36.372749999999996) [other artifacts such as the model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=38.68485714285715) [parameters into an output directory, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=40.75633333333333) [is also going to be in a Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=43.58844444444444) [bucket. So by the end of the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=45.12025) [process, a saved model will have come into](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=47.47) [existence on some GCS bucket. The saved](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=50.48912499999998) [model is a serialized, trained model. When](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=52.893) [we say it's trained, we mean that the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=56.56933333333333) [model parameters have been optimized to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=58.228) [best fit the training data. This saved](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=60.792) [model can now be used for prediction using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=63.4375) [the Cloud ML Engine prediction service.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=65.72399999999999) [When serializing our model from either](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=68.516) [XGBoost or scikit-learn, we have a choice](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=70.935) [of three serialization libraries, pickle,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=73.507) [joblib, and xgbBoost.Booster. The last is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=77.016) [available only in XGBoost. A saved model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=80.82014285714286) [file has to be named either model.pkl or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=83.33833333333334) [model.joblib or model.bst. So we should](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=87.1306) [not try to get too creative with our saved](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=90.62866666666667) [model names. Once the training service](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=93.4872) [goes ahead and writes out the saved model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=95.5502857142857) [to some location, this same saved model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=97.5675714285714) [can then be deployed. These deployments](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=100.45411111111115) [are once again going to be hosted on GCE](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=103.07614285714287) [Compute Engine instances, and this is what](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=107.00785714285713) [brings into play the Cloud ML prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=109.74044444444442)[service. Now it's important for our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=112.47866666666667) [prediction code to follow the same set of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=115.13999999999999) [data processing steps as our training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=117.70525) [code. If we do not do this, we are likely](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=120.48460000000003) [to end up with poor performance in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=123.09118181818177) [production, and this is something known as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=125.75649999999996) [training- serving skew. In order to avoid](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=127.71685714285717) [this, it's really important that training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=130.53628571428564) [and prediction follow the same data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=132.84649999999996) [transformations. But other than this,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=136.59300000000005) [there is one important difference between](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=138.20300000000003)[prediction and training. The prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=140.2002) [capabilities of our model are going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=142.714) [accessed by clients anywhere, and that's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=145.49685714285715) [why the ML Engine service provides](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=147.4301428571428) [endpoints. These are HTTP endpoints which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=150.35099999999997) [clients can access, even if they are not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=152.85699999999994) [in Python. So our clients are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=156.11700000000008) [send an HTTP request with the data, these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=158.3256666666667) [are the features of the X variables, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=160.61171428571427) [our endpoint is going to send back the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=163.5557777777778) [prediction response. That is the Y value.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=165.314) [All of this is done for us by the ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=168.319) [Engine. These endpoints are maintained,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=170.55172727272728) [and they're associated with a specific](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=172.934) [model. So we create a named model and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=176.32571428571433) [associate it with an underlying saved](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=179.111) [model, a version, and a description. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=181.073) [ability to manage versions and associate](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=184.65966666666668) [them with saved models is an important](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=186.82585714285713) [feature of Cloud ML Engine. Let's also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=189.35583333333335) [talk about the two types of prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=192.30399999999997) [supported by the ML Engine. The first of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=194.12233333333333) [these is online prediction where the aim](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=197.74928571428572) [is to minimize latency. The second is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=200.15100000000004) [batch prediction, where the aim is to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=202.69800000000004) [maximum throughput. Currently, as of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=205.0685714285715) [November 2018, XGBoost and scikit-learn are only available in the online mode.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=5&mode=live&start=207.03339999999997)

[ML Engine Pricing: Scale Tiers and Training Units](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live)

[Let's spend some time understanding the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=1.999) [pricing of Cloud ML Engine. At the very](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=3.559) [outset, let me start by saying that ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=6.6256666666666675) [Engine pricing is fairly complex, and at](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=8.672) [least some of that complexity comes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=11.255857142857144) [because of terms in Cloud ML Engine being](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=13.225285714285715) [slightly different or duplicates of terms](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=16.2495) [in the Google Compute Engine. I'm](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=18.559875) [referring here to scale tiers and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=20.902) [underlying machine types. In any case,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=23.1) [let's go ahead and get started. It starts](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=25.972999999999995)[simply enough. Local training, which is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=28.792500000000004) [always something we can do using Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=30.899333333333338) [Engine, is free. No job is submitted, no](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=34.47200000000001) [special hardware is permission to run the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=38.1455) [training job, and there is no extra cost.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=40.849333333333334) [Next, let's talk about distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=44.826) [training. This is where the training is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=47.11516666666667) [performed by Cloud ML Engine via a job,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=49.14199999999999) [which is running on special hardware. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=51.702333333333335) [pricing for this can be a little complex,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=54.49175) [but if we break it down, hopefully we'll](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=56.094)[understand it clearly enough. Training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=58.80466666666666) [jobs are charged per hour, per training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=61.49966666666666) [unit. Now the obvious question that you're](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=64.77333333333333) [going to ask is, well, what's a training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=67.43200000000002) [unit? Well, a training unit is an internal](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=70.06474999999999) [GCP abstraction. The number of training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=73.7833333333333) [units that you consume is determined by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=75.84457142857146) [the machine configuration that you opt](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=78.284) [for. Broadly speaking, there are two types](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=82.00320000000002) [of configurations. There are the scale](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=84.92350000000002) [tiers where you go with predefined](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=86.48555555555555) [configurations, and then there are custom](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=88.039) [configurations. Notice here that we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=91.311) [not only talking about the machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=93.07087499999997) [hardware properties; these configurations](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=95.5458) [also govern the number of worker nodes,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=97.45766666666667) [master nodes, and parameter servers. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=100.23260000000002) [cycle through both of these, starting with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=103.77833333333334) [the scale tiers. There are a number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=105.61120000000001) [predefined scale tiers, BASIC, STANDARD\_1,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=107.9705) [PREMIUM\_1, BASIC\_GPU, BASIC\_TPU, and then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=111.21319999999997)[the special value of CUSTOM. Now a scale](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=114.36933333333333) [tier is a combination of a machine spec,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=117.73933333333333) [along with the number of parameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=120.699) [servers, workers, and masters. For](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=123.03385714285714) [instance, the simplest scale tier is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=125.415) [BASIC. Here we have a single worker](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=127.539) [instance. This works best for learning how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=129.85585714285716) [ML Engine works or trying out new models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=131.90874999999994) [with small datasets. That one single](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=134.59399999999997) [worker instance is of type n1-standard-4.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=136.49575000000002) [One step up is scale tier STANDARD\_1. Here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=138.555375)[there's one master instance, four workers,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=143.4432) [and three parameter servers. The machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=145.67133333333334) [configurations defer as well. The master](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=148.3536666666667) [and the workers are n1-highcpu-8; the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=150.75033333333332) [parameter server is n1-standard-4. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=153.11033333333333) [other scale tiers are variations on this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=156.67885714285714) [team. So in PREMIUM\_1 we have 19 workers](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=158.715375) [and 11 parameter servers, and then there's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=161.7936) [BASIC\_GPU where we have a single worker](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=164.68687500000001) [instance with an NVIDIA GPU. There's also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=167.41157142857142) [BASIC\_TPU, which runs Google's proprietary](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=171.195) [hardware, which is optimized for training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=173.3538) [jobs using deep neural network models. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=176.68175) [addition to all of these, there was also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=180.5738888888889)[the scale tier CUSTOM. This is a special](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=182.221) [scale tier in which we as the end users](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=184.93600000000004) [have much greater control over the number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=188.4481428571429) [and type of the master, worker, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=191.37787500000002) [parameterServer instances. When we opt for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=194.903) [the CUSTOM scale tier, we can also specify](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=197.656) [the exact machine type. There are several](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=200.42749999999995)[available choices, including standard,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=203.11120000000003) [large\_model, cloud\_tpu, and so on. Coming](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=205.431) [now to the pricing, here is what prices](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=208.89800000000002) [look like as of November 2018 for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=211.58474999999999) [United States geographic region. Each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=214.8266) [scale tier is equivalent to a certain](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=217.52100000000002) [number of training units and is then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=219.453) [charged at a deferring dollar rate per](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=221.725) [hour. So for instance, the BASIC scale](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=224.99662500000002) [tier offers 0.566 training units and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=227.39942857142853) [currently incurs a charge of around 27](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=231.22699999999998) [cents an hour. In contrast, the PREMIUM](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=234.48800000000006) [scale tier offers 49.32 training units and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=237.6102857142857) [will set you back by $ 24.16 per hour. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=241.068) [there's a pretty wide difference between](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=245.35199999999998)[the pricing of these different scale](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=247.38400000000001) [tiers. It's also interesting to see how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=250.05199999999996) [much more capacity we get when we go with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=252.47924999999998) [a BASIC\_GPU or a BASIC\_TPU. Remember,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=255.22455555555555) [however, that not all machine learning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=258.0210000000001) [applications will actually benefit from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=260.48049999999995) [the use of GPUs or TPUs. And finally, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=262.122)[pricing for CUSTOM scale tier](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=266.19314285714296) [configurations depends on the machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=268.93125000000003) [type. Here is a quick rundown of those](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=270.9834) [custom prices as of November 2018. Without](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=273.50924999999984) [going into a whole lot of details, you can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=276.47075) [see that the training units offered for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=278.8365) [the custom machine types also vary very](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=280.9598888888889) [widely. At the low end, a standard machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=283.59111111111105) [type offers around 0.566 training units. A](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=285.8001250000002) [complex\_model\_1\_v100 offers 63.64 training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=290.81100000000004) [units, and the prices are correspondingly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=295.4501666666667) [inflated as well. Because of how different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=297.63225) [these numbers are, it's actually fairly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=300.31771428571426) [important for us as end users of ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=302.00928571428574) [to have a handle on exactly what kind of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=304.83849999999984) [scale tier we are going with. And there's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=307.67750000000007) [also some more fine print we should be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=309.87125000000003) [aware of. These training costs are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=312.1935714285715) [calculated in 1-minute increments, so the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=314.45980000000003) [hourly rate is going to be multiplied the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=317.043) [number of minutes divided by 60. But there](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=319.22925000000004) [is a minimum of 10 minutes per training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=322.79577777777774) [job. And the timer on these costs kicks](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=324.93787499999996) [off when the resources are provisioned,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=328.8115) [and the timer ends when the job finishes.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=330.4395714285714)[Later on in the course when we actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=334.02425) [run these in code, we will see, for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=337.76866666666666) [instance with that XGBoost job, that our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=340.16766666666666) [training job takes about 5 minutes and 50](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=342.93174999999997) [seconds to run, but it's only 2 or 3](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=345.41327272727267) [minutes into that process that our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=348.74) [training instance is even provisioned. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=350.83057142857126) [so that is when we start actually paying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=353.499625) [charges. We are not actually being charged](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=355.60871428571426) [for the entire 5 minutes and 50 seconds](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=357.79925000000003) [that we see in the ML Engine UI. Now if](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=360.44)[you were paying really close attention,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=364.4071428571429) [you would have noticed that that UI](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=365.73283333333336) [mentioned consumed ML units, not training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=368.2096666666667) [units, and that's because the consumed ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=371.76728571428566) [units are simply the training units](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=374.67785714285714) [multiplied by the job duration in minutes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=376.89257142857144) [divided by 60. So overall, it's pretty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=380.3185) [clear that the pricing for distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=383.526) [training on the Cloud ML Engine is quite](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=385.33950000000004) [complex. You definitely would want to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=388.6401666666668) [Google the latest docs. These are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=390.6168571428572) [available at the links onscreen now. Right](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=392.6697142857143) [now, the pricing for scikit-learn,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=396.4676666666667) [XGBoost, and TensorFlow, are identical.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=398.561) [But because there are different doc pages](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=401.489) [for these, it is entirely possible that they will diverge in the future.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=6&mode=live&start=403.457)

[ML Engine Pricing: Online and Batch Prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live)

[We just saw how pricing for training is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=2.328) [quite complex. Let's now turn our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=4.6254285714285714) [attention to pricing for prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=6.554) [Thankfully, this is a lot simpler to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=9.194)[understand. Remember that there are two](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=12.246875000000005) [types of prediction models available,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=13.492285714285712) [batch and online. As of the recording of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=16.033) [this course in November 2018, batch is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=19.260624999999994) [only available for TensorFlow. Online is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=21.864) [available for all of the supported](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=24.331857142857142) [frameworks, namely TensorFlow and Keras,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=26.0626) [scikit-learn, and XGBoost. Both these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=29.121000000000002) [modes are priced in the same way. As of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=32.21522222222222) [November 2018 for geographic region United](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=35.083000000000006) [States, the pricing is as follows: for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=37.84628571428571)[batch prediction, we are charged at the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=41.204875) [rate of 9.26 cents per node hour. For](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=43.157) [online prediction, we are charged about](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=47.88771428571428) [5.6 cents per node hour. Once again, these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=50.082) [rates are very likely to change over time,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=53.81450000000001) [so please be sure to refer to the docs for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=56.05324999999999) [the latest versions. Now this as it stands](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=59.648875000000004)[seems simple enough. There's just a little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=62.52249999999999) [bit of additional fine print, which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=64.03833333333333) [definitely should be aware of. A node hour](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=66.1851666666667) [is basically the time spent by a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=69.30712500000003) [prediction job running on a VM instance.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=71.45885714285713) [This definition implies that time spent in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=75.032) [a ready state waiting for requests to come](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=77.784) [in is also charged. Remember that the aim](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=80.317) [of batch prediction is to minimize the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=83.9874285714286) [total job time, so if that's the mode that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=86.1032857142857) [we are going with, scaling is not really](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=88.21866666666669) [going to impact how much money we pay. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=90.47375) [way the ML Engine pricing has been worked](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=93.95625) [out, if we choose a high compute capacity](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=95.949125) [for a short period of time, that ought to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=98.85614285714286) [cost us roughly the same as a small](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=101.9908) [compute capacity for a long period of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=104.39677777777779) [time. So for batch prediction, it probably](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=107.88122222222228) [just makes sense to set the maximum number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=110.29471428571428) [of nodes and beyond that, let the service](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=112.51712499999996) [do its thing. Online prediction is more](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=115.26787499999998) [subtle. Here, the aim is to minimize the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=118.92028571428573) [latency of individual requests, and this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=121.63611111111106) [means that the service might often keep](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=123.73566666666665) [the model ready for a few minutes after](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=126.04171428571432) [each request, assuming that there is some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=128.69599999999997) [locality in the frequency of requests.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=130.80012500000007) [Because we are going to be billed for each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=134.506) [of those few minutes for each of the nodes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=136.36) [that we have reserved, scaling is directly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=138.607) [going to affect our costs. The more](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=141.18) [numerous and the more frequent our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=144.26825000000002)[requests, the greater the costs are going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=146.012) [to be. While using the online prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=148.332) [service, we have two choices. We can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=151.4881666666667) [either go with automatic scaling, or we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=154.11599999999999) [can fix the number of nodes up front. If](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=155.9554) [you are looking to minimize your Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=158.692625) [bills, you certainly ought to consider](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=161.58471428571428) [automatic scaling. That's because this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=163.833) [will scale down to zero in zero-traffic](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=165.806) [periods. On the other hand, if you go with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=168.944625) [a fixed number of nodes, you are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=170.94180000000003) [be billed for all of those nodes for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=173.12955555555558) [entire running time, whether or not any](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=175.44737499999994) [client requests are actually coming in.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=177.95585714285718) [The downside of automatic scaling is that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=180.094) [the first client request after a long](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=182.372) [period of inactivity is going to suffer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=184.78) [poor latency because it's going to take](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=186.798) [the service some time to provision the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=189.15175000000005) [nodes and fire the model back up. With](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=191.37885714285716) [that, we've reached the end of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=193.707) [module, in which we introduced Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=194.981) [Engine and covered all of the theoretical](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=197.333) [concepts required for the demos up ahead.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=199.795) [Cloud ML Engine is a production-ready ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=202.747) [system. It provides training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=206.30912499999994) [prediction services on the cloud and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=208.78120000000004) [supports TensorFlow, Keras, XGBoost, and scikit-learn.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=f2e2c4e2-8c32-4429-a174-a325dc6c61d8&clip=7&mode=live&start=211.4628333333333)

[Deploying XGBoost Models to Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live)

[Module Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live)

[Hello and welcome to this module on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=0) [Deploying XGBoost Models Using Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=4.515) [Engine. XGBoost is a Python library built](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=8.264142857142856) [on a technique called ensemble learning.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=11.189999999999998) [This is a fairly popular library with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=13.838) [Kaggle users, and it is also gaining](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=16.798) [popularity for use on the cloud because](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=19.299) [training converges quickly when we use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=21.594) [ensemble learning. This means that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=23.998) [relative to deep learning techniques using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=26.843) [TensorFlow or Keras, XGBoost can often be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=29.138) [quite cost effective. You will remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=32.11) [that XGBoost is only supported in online](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=35.078) [prediction mode on the Cloud ML Engine. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=37.217999999999996) [will explore ML Engine support for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=40.407) [building, training, and deploying models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=43.165) [in XGBoost. We will also code up a simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=45.269600000000004) [Python client to use our published model for prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=0&mode=live&start=48.36677777777778)

[Implementing Models in XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live)

[Let's start with a very quick introduction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=2.376) [to building models using XGBoost. XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=4.625) [is an ML library which focuses on a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=7.912666666666667) [specific ML technique called ensemble](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=10.566) [learning. The word ensemble means](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=13.808500000000002) [together. Ensemble learning techniques are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=16.4394) [of two types, bagging and boosting.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=18.18433333333334) [Boosting techniques sequentially build a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=21.312) [number of learning models. Each of these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=23.498) [models might individually be weak, but](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=26.47671428571429) [when combined, they give a strong learning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=28.3435) [model. That's basically the idea behind](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=30.820999999999998) [XGBoost. XGBoost is far less widely used](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=33.514857142857146) [than scikit-learn or TensorFlow/Keras, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=37.898199999999996) [other options available on ML Engine. So a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=40.688428571428574) [quick word of introduction. It is popular](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=44.21666666666667) [with the Kaggle community, and it is also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=46.554571428571435) [the basis of a pretty popular R package.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=48.72333333333333) [These in turn have been driving the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=50.913875000000004) [increase in popularity of the XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=53.204142857142855) [Python library in recent times. Google's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=55.309666666666665) [decision to add XGBoost support in ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=58.54257142857143) [Engine was well founded. XGBoost is well](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=61.4758) [known for quick and efficient training.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=64.03385714285714) [This is important while running on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=66.3278)[cloud, because after all, long-running](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=69.2598) [training jobs which do not converge](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=71.20560000000002) [quickly can end up costing your](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=73.66199999999998) [organization a significant amount of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=75.876)[money. XGBoost, like other R packages,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=77.8262857142857) [also deals with missing data well, and it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=80.60733333333334) [is pretty quick and efficient while](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=83.60175000000001) [working with small datasets. For features](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=86.0632) [which are well structured and datasets](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=88.82342857142856) [which are small XGBoost is a really good](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=90.94014285714286) [option, and you should definitely consider](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=93.714) [using it. Before we plunge into the demos](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=96.0195) [on XGBoost, let's take a quick segue to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=99.24774999999997) [understand how decision trees work. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=101.40650000000001) [start with a simple classification](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=104.58242857142857)[problem. Is an individual a jockey or a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=106.76957142857141) [basketball player? Jockeys tend to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=110.00622222222226) [light, and basketball players tend to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=112.42840000000002) [tall, strong, and heavy. So the question](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=114.8062857142857) [is, given physical attributes of an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=117.385) [individual such as the height and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=119.94583333333335) [weight, can we correctly classify this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=121.73214285714288) [individual as a jockey or a basketball](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=123.78366666666669) [player? Given what we know about the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=126.39385714285714) [general characteristics of these two types](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=128.479) [of athletes, we could go ahead and fit](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=130.363) [that knowledge into rules, and each rule](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=132.70966666666658) [would involve a threshold. So for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=135.04642857142852) [instance, we might start by looking at the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=137.96699999999998) [weight of an individual and impose a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=140.32877777777776) [threshold of 150 pounds. We would classify](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=143.135) [anyone heavier than that as a basketball](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=146.7315714285714) [player, and anyone lighter than that,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=148.97942857142857) [well, we would then need to examine their](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=150.90224999999998) [height. Here once again, we would have to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=153.55566666666664) [use a rule and a threshold. Let's say that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=155.771) [we went with a threshold of 6 feet. Any](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=158.73433333333332)[individual taller than 6 feet was a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=160.965) [basketball player. Any individual shorter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=163.88899999999998) [than that was a jockey. And in this way,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=166.21025) [we have constructed a fairly crude](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=169.27149999999997) [decision tree. We knew something about the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=171.2425) [decision variables height and weight, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=173.871) [we knew roughly where some reasonable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=175.93542857142853) [thresholds might lie, so using those, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=178.3372) [went ahead and constructed rules. Now of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=180.621) [course, to formalize this, we would need](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=184.264) [to determine the order of the decision](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=185.911) [variables in some scientific way, and we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=187.72083333333333) [would also need to use some kind of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=190.47220000000002) [mathematical technique in order to find](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=192.201) [the correct thresholds. And this is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=193.9967142857142) [exactly what training data is good for, so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=197.27016666666663) [we could find the rules and the order](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=199.57922222222217) [using training data and get a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=201.81125000000006) [well-constructed decision tree using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=204.66400000000004) [machine learning. One common technique](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=206.9925) [used to build such trees is called CART.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=209.0185714285714) [That's an acronym for classification and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=211.783) [regression tree. The way this would work](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=213.767) [is we would build an ML-based classifier.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=217.166) [That would then take in the weight and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=219.897)[height of an individual, and it would](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=221.647) [output a label, telling us whether that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=223.615) [individual was a jockey or a basketball](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=225.829) [player. This is the idea of an individual](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=229.81899999999996) [decision tree. Ensemble learning takes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=232.64222222222216) [that a little further. Here we build many](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=234.58628571428574) [different trees and then combine their](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=237.81900000000002) [outputs. That combination could be done](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=240.67866666666666) [either in series or in parallel. Based on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=243.00971428571424) [this, we have bagging and boosting. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=247.0145) [bagging, we build the trees in parallel.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=249.5115) [In boosting we build the trees](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=251.671) [sequentially and feed the output of one](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=253.92271428571433) [tree in as the input of another. The term](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=255.96420000000006) [gradient boosting is used to refer to one](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=259.58149999999995) [specific form of boosting. It refers to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=262.495) [the value of the shrinkage factor, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=265.025) [is used while combining the individual](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=266.8741428571429) [learners. That's not really relevant here.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=268.93899999999996) [In general, the idea behind gradient](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=271.928) [boosting is that we combine many different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=273.609) [trees, many different learners, and this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=276.089) [diversity gives us strength. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=280.29166666666663) [individual learners might be very weak,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=282.5074285714286) [but when we combine them, we get a fairly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=284.73) [strong, robust learner. The term ensemble](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=286.804) [means together in French, and it is a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=290.6397142857144) [fairly standard way to aggregate models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=293.93250000000006) [together. Because we are combining the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=296.19233333333335) [output of many different learners, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=298.22437499999995) [final learner tends to be fairly robust.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=300.4375) [This is called a regularization technique.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=304.193) [This reduces overfitting and variance](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=306.745) [error. Overfitting is a problem which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=309.2974999999999) [occurs when a model performs really well](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=311.17442857142845) [in training but really badly with new data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=313.7167142857143) [that is in live prediction. This was just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=316.21500000000015) [meant to be a quick indicative and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=318.9131249999999) [intuitive introduction to gradient](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=321.6971999999999) [boosting, which is a form of ensemble](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=323.96775) [learning, and of course is what drives the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=326.0412857142856) [XGBoost package. We'll now break into a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=329.51411111111116) [series of demos that show how XGBoost can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=331.77977777777784) [be used to build and deploy models on the Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=1&mode=live&start=334.21)

[Enabling ML Engine APIs, Creating Service Account Keys and Storage Buckets](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live)

[Let's get started with a simple demo. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=2.038) [the course of this demo, we will first](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=4.465999999999999) [enable the ML Engine API. We'll then go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=6.074) [ahead and create a service account. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=9.418666666666665) [is required because we'll be writing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=11.396285714285714) [Python code directly from Cloud Shell. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=12.958) [finally, we will create a cloud storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=16.38175) [bucket which we will be using again and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=18.186) [again in the remainder of this course.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=20.58) [Make sure that the project you're using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=22.917) [for these demos has billing enabled. You](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=24.993374999999993)[can find the project name up top. In this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=28.004375) [case, it's spikey-ml. And you can also see](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=30.9678) [the billing status in the billing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=33.69500000000001) [dashboard over on the right. Next, please](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=35.751000000000005) [click on the three horizontal lines over](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=38.800428571428576) [in the top left. These three horizontal](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=41.514333333333326) [lines are sometimes called the hamburger](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=44.187) [menu or the navigation menu or the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=46.39071428571428) [products and services pane. They're a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=49.428142857142866) [great way to access different services on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=51.73957142857142) [the GCP. So we click on the navigation](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=54.473)[menu, navigate into APIs and Services, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=57.72014285714287) [Library. There we type in the word ml.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=60.887428571428565) [This brings up three results. The one that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=63.894) [we care about here is the ML Engine API](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=66.43239999999999) [over on the extreme right. Let's click on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=69.35457142857143) [that and go ahead and hit ENABLE. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=73.3517) [takes us to a page with some details about](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=76.30144444444444)[the API and then helpfully, there is a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=78.25885714285714) [button over in the top right called CREATE](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=81.028) [CREDENTIALS. In this course, we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=84.21066666666665) [basically be using Python directly from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=86.4875) [Cloud Shell to write our code. And for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=88.52857142857144) [that, we are going to need to create some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=91.29511111111111) [credentials. Please note that this step is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=94.96640000000001) [not required if you're going to be writing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=97.19699999999997) [Python from a Cloud Datalab instance, but](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=99.24866666666664) [for our purposes here, the code editor in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=101.7363333333333) [Cloud Shell is very handy. So when we're](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=104.47957142857139) [prompted for what kind of credentials we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=107.89025000000001) [need, we click on service account. Then we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=109.92525) [click on the CREATE SERVICE ACCOUNT](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=113.06525) [button, type in the name, which here is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=114.464) [spikey-project-admin. We hit CREATE, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=116.513) [then when we are prompted for a role, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=120.46199999999999) [give this the Project, Owner role. We're](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=123.57387499999999) [fine for this service account to have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=126.97325) [access to all resources in our project. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=128.69) [hit CONTINUE and then go ahead and create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=132.81574999999998) [a key. This is a key which we are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=134.9502) [download to our local machine and then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=138.18750000000003) [upload to the Google Cloud Shell and use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=140.459625) [it from there. There are two types of keys](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=143.24277777777777) [available, JSON and P12. We go with the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=146.993) [JSON key. This saves the private key to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=150.77766666666662) [our computer. In a moment, we will switch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=154.15974999999992) [back to Cloud Shell and use this to set](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=156.28633333333335) [our credentials, but for now, let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=159.41825000000003) [ahead and create the bucket which we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=161.2684444444445) [be using in the remainder of this course.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=163.00549999999996) [For that, we navigate into the products](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=166.331) [and services menu into the Storage,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=168.37849999999992) [Browser. There we click on the Create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=171.72100000000006) [button and go ahead and use a regional](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=174.48299999999998) [bucket. This is an important point. Make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=178.11150000000004) [sure to use regional buckets while working](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=180.64385714285714) [with Cloud ML Engine because the docs warn](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=182.58700000000002) [us that there can sometimes be issues with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=185.1345) [eventual consistency during the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=188.16500000000002) [process. So we create a regional bucket.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=190.867375) [This one is called spikey\_ml\_engine. Once](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=194.216) [that's done, let's go ahead and create a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=197.82537499999998) [folder, one for the datasets which we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=200.1038888888889) [going to be using in this course. Within](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=202.631) [this datasets folder, we can go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=205.31271428571426) [upload some files from our local machine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=208.01588888888887) [Here we have an automobile price dataset.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=211.351) [This is a fairly common and simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=214.667) [dataset, which is available at the URL](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=217.14237499999996) [onscreen now. This dataset is also linked](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=220.64899999999997) [off of Kaggle. This is also a good time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=223.58562500000002) [for us to very quickly take a look at the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=225.84499999999994) [columns and rows in the data. One point](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=228.2723333333333) [worth paying attention to, certain fields](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=230.8335714285714) [in this dataset may contain missing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=232.87633333333335) [values; those missing values are indicated](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=235.56100000000004) [by a question mark. This is something we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=237.50442857142852) [will need to handle in our ML models.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=239.93585714285715) [Please also note that several columns in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=244.468) [this dataset will have categorical or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=246.731) [discrete values. All of the frameworks we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=248.778) [plan to use will be able to handle such](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=252.338) [issues. In any case, the upload is now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=255.27289999999996) [complete. The file appears in our Google](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=257.80837500000007) [Cloud Storage bucket. Let's open up Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=260.152857142857)[Shell. Cloud Shell can be activated using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=263.7412857142858) [that little button over on the top right.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=266.0367142857142) [Remember that Cloud Shell represents a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=269.243) [terminal session on an ephemeral VM](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=271.553) [running on the Google Cloud Platform. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=275.04585714285713) [a really useful feature of the Cloud Shell](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=277.8295) [is the code editor. Launch the code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=280.1214) [editor, click on that little pencil icon](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=282.67742857142866) [in the Cloud Shell window. This will bring](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=285.29375) [up a nice UI in which we can navigate the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=288.38910000000004) [file system of our little home directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=291.24775) [on the Google Cloud Platform. Using this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=294.04042857142855) [UI, we can create files, write code, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=297.196) [even upload files into our home directory.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=300.3766250000001) [The very first thing we are going to do is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=303.03857142857146) [to upload the JSON key file associated](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=306.03700000000003) [with our service account. Let's go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=308.2273333333334) [and upload this. Once the upload is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=310.99749999999995)[complete, we can then go ahead and point](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=312.99233333333325) [our application credentials to it. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=315.9765000000001) [switch back by closing the code editor by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=319.09425000000005) [clicking on the pencil icon. Now note that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=322.8076250000001) [our Cloud Platform project is set to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=325.69125) [spikey-ml, so all of the commands that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=328.1385) [run in this window are going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=331.1086) [executed in the context of that project.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=332.807) [Before we move on, let's just change our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=335.878) [prompt to reduce its size. This will give](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=337.811) [more screen real estate for the commands](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=341.232875) [that we wish to run. Of course, this step](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=343.1705714285714) [is strictly optional. Next, we want to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=345.9422857142858) [make sure that any commands which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=347.8474999999999) [execute in this Cloud Shell window use the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=349.6391428571429) [credentials we just uploaded. To ensure](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=353.0135) [this, we point the environment variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=355.97585714285714)[GOOGLE\_APPLICATION\_CREDENTIALS to the JSON](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=358.84433333333334) [key file containing the private key of our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=361.5315714285714) [service account. We now can go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=364.20480000000003) [run our first ML Engine command. We make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=366.8667999999999) [use of the gcloud command line utility. At](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=369.8329999999999) [this point, the command we run is gcloud.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=373.5811111111111) [The second parameter we specify is always](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=376.582) [going to be ml-engine because that is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=378.68) [service which we are going to hit. We want](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=381.242) [to list out all of the models. At this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=383.73033333333336)[point, we haven't created any, and so we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=386.2112857142857) [get a response saying that there are 0 models at this point in our project.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=2&mode=live&start=388.42699999999996)

[Implementing a Simple XGBoost Model in Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live)

[We will now move into a second demo in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=2.882) [which we will write a simple XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=4.6) [model, train this using our automobile](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=7.512875) [dataset, and we will perform this training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=10.561500000000002) [process twice, once locally and once on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=12.840857142857137) [the cloud. Just to be clear, whenever you](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=16.033999999999995) [see the term local in the context of ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=18.31833333333333) [Engine, it refers to the machine on which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=21.116428571428568) [the command is being run. So here the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=23.033999999999992) [local training is going to be executed on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=25.83933333333333) [the VM where our Cloud Shell session is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=27.92575) [being hosted. Let's pick up the action](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=30.929) [from our code editor window. We click on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=33.58899999999999) [New and open up a new folder. Let's call](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=36.109444444444456) [this folder spikey-ml. Within this folder,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=40.79550000000002) [let's go ahead and create a new subfolder.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=43.61785714285714) [This is going to have only the data for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=46.929) [the XGBoost model. Let's call it auto\_xgb.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=48.8496) [Within that, we create a further subfolder](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=54.387) [called xgb\_trainer. This xgb\_trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=56.317) [directory is going to contain our Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=59.843) [package, and so it must have a pretty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=62.345) [specific structure. Within this, we go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=64.509) [ahead and create a new file called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=67.74455555555556) [task.py. This is going to house our simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=71.05733333333333) [XGBoost model. We start off with a bunch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=75.13211111111113) [of import statements. These include the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=78.71833333333336) [pickle package which we require to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=81.21714285714285) [serialize our model. In additional to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=83.5976)[these system-level imports, let's also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=86.42899999999999) [import the data processing packages we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=88.70149999999998) [need. These include pandas, XGBoost,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=91.48725000000002) [sklearn, and finally, let's also import](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=95.02059999999997) [Google Cloud Storage. We make use of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=98.67287499999999) [client library to get access to the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=101.76544444444446) [Storage bucket which contains our data. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=104.1627142857143) [we instantiate the client and then get](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=106.7737142857143) [hold of the bucket. Remember that bucket](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=109.58300000000003) [names have to be globally unique. Let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=112.14928571428572) [ahead and load the data from within that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=115.34185714285715) [bucket, so we access the file](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=118.78333333333336) [Automobile\_data.csv from the datasets](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=122.298) [folder and access this using the blob](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=124.08314285714285) [handle. We then go ahead and download this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=127.326) [file to the local machine, that is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=129.5264444444445) [local Cloud Shell VM instance. And we load](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=132.35349999999997) [it into a pandas DataFrame using the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=135.85599999999994)[pd.read\_csv method. Let's perform a few](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=140.51000000000005) [simple preprocessing steps. Let's restrict](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=144.27571428571432) [the number of columns which we care about,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=146.36457142857145) [so we have a list of the required columns.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=148.39000000000004) [These include the make, body style, number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=151.18233333333333) [of cylinders, engine size, bore,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=154.58360000000002) [horsepower, and the price. In this simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=157.29057142857144)[model, we are going to try and predict the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=160.2896666666667) [price using all of the other columns in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=162.40024999999997) [here. So we restrict our DataFrame to only](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=165.69375000000002) [those required columns then we go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=168.73987500000004) [and drop all of the missing fields and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=171.80350000000007) [drop all of the NAs. Here we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=174.98166666666674) [explicitly searching for the question mark](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=178.47271428571432) [which is used to signify a missing field.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=180.34650000000002) [Now because we want to predict the price,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=183.408) [we obviously should not be including it in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=185.572) [our list of features, so let's create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=187.618) [another DataFrame with only the features,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=190.147) [that is the x variables, by dropping the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=192.96) [price column. And we explicitly store the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=195.23) [price column in a variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=198.351) [auto\_target. This is the target variable.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=200.236) [We are also careful to convert this to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=203.567) [numeric values using pd.to\_numeric. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=205.243) [also had a column with the number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=210.65444444444444) [cylinders, which was expressed in the form](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=212.986) [of a string. Let's convert that from a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=214.911) [string to a number using a LabelEncoder](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=217.314) [object. This is from the preprocessing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=221.38362500000005) [module in the sklearn package. Next, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=223.5107142857143) [convert some of our categorical variables](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=227.65142857142857) [so that they are in one-hot encoding form.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=229.81314285714285) [This can easily be done using the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=232.70600000000002) [pd.get\_dummies method. We invoke this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=236.896) [method on the make and the body-style](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=239.709) [columns, which had categorical data and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=241.92620000000002) [which requires the one-hot encoding to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=244.567) [performed. By this point, we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=247.69825000000003) [successfully preprocessed all of our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=249.6438333333333) [columns, so we can go ahead and load these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=251.56049999999993) [columns from our pandas DataFrame, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=254.77137499999992) [is called auto\_features, into an XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=256.8368) [DMatrix. DMatrix, or data matrix, is an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=261.08919999999995) [XGBoost internal data structure which is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=264.8868571428573) [optimized for memory efficiency and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=268.3431666666668) [training speed. DMatrices can be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=271.26233333333323) [constructed from NumPy arrays, pandas](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=273.8495) [DataFrames, and from csv files. Notice](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=276.70399999999995) [when we perform this invocation, we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=280.7365) [to specify both the features and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=283.58425) [label, so we specify that the label is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=285.6187777777778) [equal to the auto\_target, which, of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=288.449) [course, is the price of our automobiles.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=290.48985714285703)[This is clearly a good use case for linear](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=293.11) [regression, and so we go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=295.386) [instantiate a model. Note that if we do](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=297.369) [not specify parameters or details about](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=301.4618750000001) [the model type, XGBoost is going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=303.775) [perform linear regression, so that's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=306.4573333333333) [what's happening here. The first argument](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=308.75000000000006) [we pass into xgb.train is an empty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=311.68700000000007) [dictionary. This means that the parameters](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=314.53625) [are empty. If we were building a different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=317.05328571428566) [model, we might have specified parameters](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=319.58642857142866) [such as the maximum depth of our tree and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=321.8267142857143) [other hyperparameters. The second](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=326.13355555555563) [parameter here is the training data, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=327.85685714285717) [the third is the number of rounds, or the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=330.25249999999994) [number of epochs. Here we set this to 20,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=333.30300000000005) [which means that our training process will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=336.2003333333332) [go through every point in the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=338.461) [data 20 times. By the end of this process,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=341.56071428571425) [we have a trained model stored in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=344.2333333333332) [variable called bst. Now we are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=347.77437499999996) [go ahead and save this model using the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=350.495) [bst.save model method. There are three](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=353.85249999999996) [possible libraries that we could use to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=357.0467142857142) [serialize our data, pickle, joblib, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=359.22114285714287)[xgboost.Booster. The file name must be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=363.1646666666667) [model.bst. This is required by ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=366.64066666666673) [At this point, we've saved our model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=370.303) [locally. Let's upload it to Google Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=372.23587499999985) [Storage. Let's construct the Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=374.7431428571429) [part and then go ahead and invoke gsutil](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=377.15928571428566) [followed by the cp command. Notice how we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=380.9725000000001)[are making use of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=383.9585000000001) [subprocess.check\_call method to execute the system command gsutil.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=3&mode=live&start=386.7014999999999)

[XGBoost Model: Train Locally](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live)

[That does it for our task.py. Let's also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=0) [go ahead and create a new empty init.py.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=5.7665999999999995) [This file is required by setuptools and is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=9.529) [a part of the process of packaging up our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=12.559) [trainer into a standard Python package. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=14.969) [let's save this empty file, get out of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=19.024444444444445) [code editor, and return to Cloud Shell.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=21.729) [Let's quickly run a few pip installs to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=24.172) [make sure we have all of the required](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=26.666) [packages. So we go ahead and run pip](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=29.801999999999996) [install scikit-learn, pip install pandas,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=31.955375000000004)[pip install xgboost, and finally pip](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=35.071) [install, the google-cloud-storage client](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=38.35) [library. Once all of these are done, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=42.1428) [cd into the exact directory which contains](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=44.87755555555556) [our code. That is spikey-ml/auto\_xgb.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=47.6187142857143) [Let's run the ls command to orient](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=52.639) [ourselves. The only directory contained](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=56.988625000000006) [here is the xgb\_trainer. This is what we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=58.90542857142857) [need for our training package, so let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=61.869499999999995) [initialize an environment variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=64.83433333333333) [TRAINING\_PACKAGE\_PATH. This is going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=67.217) [point to the xgb\_trainer directory right](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=70.43100000000003) [here. We are going to use this environment](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=73.26666666666668) [variable when we invoke gcloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=75.463) [in a moment. Let's also go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=78.26300000000003) [create the other environment variables for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=80.644) [that very delicate handshake. Let's list](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=82.96533333333336) [the contents of the current directory and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=86.28585714285714) [one level down. Within our trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=88.8136) [package, the module that we care about is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=91.51300000000002) [the task.py, so let's create an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=95.03033333333332) [environment variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=97.26633333333331)[MAIN\_TRAINER\_MODULE and point this to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=98.4628) [task.py file within our xgb\_trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=102.27775) [directory. ML Engine will use this module](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=105.90712500000001) [as the entry point during the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=108.75233333333337) [process. Finally, we are ready, so let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=111.99920000000002) [actually invoke gcloud ml-engine. For now,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=114.44028571428568) [we are just going to run this training job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=118.27222222222224)[locally; that is on the VM hosting our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=120.353) [Cloud Shell session. That's why the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=123.4794) [command is gcloud ml-engine local train.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=125.8902857142857) [This then has the package path, which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=128.24314285714286) [point to the environment variable we just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=131.644) [created. Notice how we did not actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=134.95675000000003) [create our own Python package. That's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=137.44649999999993) [because gcloud will automatically do this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=139.873) [for us. All that we need to do is to make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=142.5906666666666) [sure that our directory structure is set](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=145.06736363636364) [up correctly, and gcloud will do the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=147.03200000000004) [actual package creation. And this is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=149.5938571428571) [recommended way of doing it. You are not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=153.078) [recommended to manually create your own](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=155.58557142857143) [package. Finally, we also specify the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=157.70585714285713) [module name. This is the task.py entry](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=160.624875) [point into our training code. Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=163.543) [Engine take over and runs the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=166.84375)[process. We can see that the tree pruning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=168.688625) [process which XGBoost relies on is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=171.58557142857143) [happening under the hood, and at the end](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=174.199) [of the process, our trained model is saved locally to the file model.bst.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=4&mode=live&start=176.01855555555557)

[XGBoost Model: Train on the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live)

[So far, nothing that we did actually made](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=2.694) [use of Cloud ML Engine's true capabilities](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=5.315) [because the training process we ran a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=8.383) [moment ago was a local one. Let's change](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=10.47) [that. Let's actually submit a training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=13.901571428571428) [job. Let's start by initializing a set of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=16.587333333333333) [environment variables. The first of these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=20.807800000000004) [is the BUCKET\_ID. This is going to refer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=23.433285714285713) [to the bucket which will hold our trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=25.72522222222223) [package and which will be used by the ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=27.78928571428571) [Engine for staging. Next, let's construct](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=31.009999999999994) [a job name. Notice here how we append the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=33.4825) [date and time so that we can uniquely](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=37.08233333333334) [identify this particular job. The JOB](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=39.82714285714286) [directory, this is a subfolder within the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=43.19439999999999) [bucket that we created just above. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=45.790749999999996) [follow this with the region. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=50.0595) [us-central1. This is the same region that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=52.82933333333333) [we used for our bucket. Next, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=54.477) [specify the runtime version of the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=56.81666666666667) [ML Engine. This is 1.10. This is a version](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=59.79014285714284) [which supports XGBoost and scikit-learn.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=63.81824999999998)[Previous versions used to support only](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=66.892) [TensorFlow and Keras. After this, we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=69.054) [to specify the version of Python. Here,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=73.2425) [let's just pick version 2.7. This is for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=75.92262499999998) [completeness. There is nothing in our code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=79.87824999999998) [which was specific to version 2.7. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=81.71249999999998) [could have chosen a version like 3.5 or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=84.78933333333333) [3.7 instead. We will use versions of 3.x](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=87.144) [for scikit-learn and TensorFlow. Remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=92) [that ML Engine devotes compute capacity](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=95.62516666666666) [based on the scale tier, so here let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=98.14116666666668) [over the scale tier BASIC. This is going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=101.88233333333332) [to cause a single n1-standard-4 worker](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=104.50814285714287) [instance to be devoted to our training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=106.926) [process. We now have set up all of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=111.519) [required environment variables. Let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=114.06577777777773) [ahead and invoke the Cloud ML Engine. Once](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=116.14744444444443) [again, this is done using gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=119.17) [ml-engine, but now there's something a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=120.832) [little different. Previously, we had used](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=123.777) [the word local. That's because we were not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=126.735) [actually submitting any job. This time, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=129.476) [go with a different syntax, gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=132.87987499999997) [ml-engine jobs submit training. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=135.3844) [because we are using the training service.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=139.18133333333333) [Later we will use jobs submit prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=141.321) [when we want to invoke the models for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=143.82) [actual predictions. Next up is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=145.923) [job-dir. This is the part on the Google](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=150.17271428571433) [Cloud Storage where training outputs and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=152.60037499999993) [other data needed for training are going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=154.84266666666673) [to be stored. This, by the way, is also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=157.71133333333339) [going to be used as the staging bucket if](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=160.19759999999994)[we do not specify a separate staging](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=162.847) [bucket argument. ML Engine will also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=166.09299999999996) [validate this location and make sure that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=168.85699999999997) [it exists. Next up is the package-path.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=171.74366666666663) [This is simply the path to the Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=175.026) [package to build. This is on our local](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=176.971) [machine. Remember again that gcloud will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=181.0723333333333) [take care of building that package for us](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=183.1635714285714) [using setup rules. This is the recommended](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=186.3423333333334) [way to go. It is not recommended that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=188.54742857142858) [manually build up our own package. Next](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=191.27142857142863) [comes the module name. This the task.py](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=194.32) [contained within our trainer package. Next](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=197.495) [is the region in which the compute engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=200.6762857142857)[instances are going to be brought up.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=204.41333333333336) [Ideally, we want this to be the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=206.919) [region as our regional bucket, but that's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=208.9447777777779) [not always necessarily the case. There's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=211.71971428571428) [nothing preventing us from instantiating](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=214.29059999999998) [our GCE compute engine VMs in a different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=215.91757142857142) [region than the bucket. Then comes the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=219.97733333333332) [runtime-version, which we had set to 1.10,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=223.758) [the Python version, which we had set to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=226.947) [2.7, which, by the way, is the default.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=229.277) [Python 3.5 will only work with the runtime](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=232.633) [version set to 1.4 and above. Python 2.7](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=235.361) [will work with all runtimes. Then comes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=239.52428571428572) [the scale-tier, which we had set to BASIC.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=242.63000000000002) [Let's submit this job, and the result](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=245.90766666666667) [tells us that our job is still active. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=249.13700000000003) [can now monitor this job using the web](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=251.578375) [console. In addition, this output gives us](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=255.40850000000006) [two helpful tips. It tells us that we can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=257.8537142857142) [view the status of our job with a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=260.50066666666663) [particular command, and we can also stream](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=262.72875000000005) [the logs being written by this job out](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=265.56475) [using gcloud. So let's go ahead and try](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=269.0602500000001) [out both of those commands. Let's first](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=271.23754545454534) [try and describe our job, pass in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=273.1451428571429) [correct job name, and we get a bunch of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=276.0323) [information, the createTime, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=278.8032) [trainingOutput, and the Cloud Console URL](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=280.5243333333333) [where we can view our job and its logs.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=283.0954) [Let's try out the other command, which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=286.78) [were just told about. That is the gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=289.04966666666667) [ml-engine jobs stream- logs command. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=291.8710000000001) [in this way, we can pipe out into our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=293.9196666666666) [command shell all of the log entries being written out.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=5&mode=live&start=297.16166666666663)

[XGBoost Model: Examine Results](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live)

[Now if we switch into the web console and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=2.326) [navigate to ML Engine and to Jobs, we can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=4.915) [see that all of the details for our job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=8.1554) [are available there. This is a training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=10.331499999999998) [job, and it's already been running for 1](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=13.658000000000001) [minute and 46 seconds. Remember that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=16.534499999999998) [are charged for a minimum of 10 minutes,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=18.9615) [and it typically takes at least 5 minutes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=21.399874999999998) [for even very simple training jobs because](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=24.50971428571428) [of all the overhead associated with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=26.912000000000006) [starting up the compute engine nodes. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=28.899) [training on the cloud in this manner](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=31.861749999999997) [really makes sense for complex machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=34.35528571428571) [learning tasks where you require a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=36.52657142857143) [significant amount of compute capacity.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=38.87) [Our job is a really lightweight one. It](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=42.318) [still only consumed 0.01 ML units. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=45.42742857142857) [a measure of the training units multiplied](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=49.93375) [by the job duration so far. We can also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=52.43271428571429) [view the job logs from right here. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=56.677800000000005) [click on that link, and we get a list of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=58.7116) [the same job logs that we were streaming](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=61.260444444444445) [out to gcloud a moment ago. It takes a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=64.04) [while, but soon we start seeing the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=68.058) [kinds of log statements that we would have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=69.8255) [when we ran XGBoost locally. It talks](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=72.35828571428571) [about tree pruning and so on. And in just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=75.18325) [a short while after that, the training is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=77.88550000000001) [complete. The model.bst file is now going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=81.20724999999999) [to be copied into the Google Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=83.6425) [bucket that we had specified. Let's switch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=86.74557142857141) [back to the Cloud Shell window where we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=89.87374999999999) [held the logs still being streamed, and we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=91.98657142857142) [can see that this has run through to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=94.01499999999999) [completion as well. The job now shows up](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=96.8426) [as having succeeded in the Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=100.07633333333335) [web console, and it took quite some time.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=102.78366666666666) [It took 5 minutes and 50 seconds. We can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=104.923) [view a graph of the CPU utilization on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=109.19655555555555) [master if we scroll down just a little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=111.71888888888888) [bit. Notice how most of these 5 minutes 50](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=113.647125) [seconds were spent just waiting. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=117.42914285714286) [vertical dashed line shows when our VM](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=119.8565) [instance was provisioned. After that, it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=122.128) [took a while for the process to start, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=124.70199999999998) [then it quickly finished. At this point,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=127.78314285714286) [we can switch back to the storage browser](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=130.85633333333334) [and check on the contents of our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=132.87014285714287) [spikey\_ml\_engine bucket. Here we will see](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=135.91699999999997) [that our job directory, the xgb\_job\_dir](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=139.00400000000002) [has a packages subfolder, inside which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=141.98566666666667) [there is a long system- generated string.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=145.27242857142858) [If we click on that, we can find the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=147.71200000000002) [Python package. This is the trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=151.1853333333333)[package which was built for us by Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=152.76500000000004) [Engine. Remember that this is merely where](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=156.1096666666666) [the trainer package was uploaded. At the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=158.34514285714278) [end of the training process a saved model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=161.52811111111112) [was saved down to the same bucket that is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=164.092875) [in a different folder called xgb\_model. If](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=167.256) [we click on that subfolder, we find the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=170.38666666666666) [saved model with the name model.bst.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=172.248) [Remember that the naming conventions of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=175.775) [Cloud ML Engine are such that for our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=177.905) [model to be used for online prediction,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=180.61877777777775) [it's got to have the first name model and the extensions bst, pickle, or joblib.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=6&mode=live&start=183.345)

[XGBoost Model: Deploy Using the Web Console](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live)

[Let's move into another demo in which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=2.814) [will go ahead and deploy the XGBoost model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=4.555) [using the web console. This model will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=7.63) [then be available for use in prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=10.702499999999997) [Let's pick up the action in the Models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=13.45) [section of ML Engine. In order to get](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=16.082) [here, we use the navigation menu, find ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=19.12788888888889) [Engine, and then click on the Models tab](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=21.255666666666666) [over on the left. We don't have any models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=24.373333333333335) [at this point. Let's go ahead and create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=27.80575) [our first model. We give it a name. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=30.393125000000005) [call it xgb\_auto-model. Once we've created](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=34.147666666666666) [a model, we can select it and then go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=36.153000000000006) [ahead and create a version. This is going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=38.902111111111104) [to be v1 of our model. It requires us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=41.85433333333333) [select a version of Python. Here, we go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=45.07277777777778) [with 2.7. The other choice is 3.5. Next,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=48.288599999999995) [we've got to explicitly select the ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=52.459500000000006)[framework which was used in developing the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=54.745) [model. The choices are scikit-learn,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=57.470625000000005) [TensorFlow, and XGBoost. We will go with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=59.44879999999999) [XGBoost, of course. We will then be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=62.84866666666668) [prompted for the version of the framework](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=65.33933333333331) [XGBoost, which, in this instance, is 0.72](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=67.6805) [.1. Then we have to pick the ML runtime](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=72.78333333333333)[version, and here is something a little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=75.564) [odd. The recommended version is 1.10, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=78.58974999999997) [the other choice available right now is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=82.86788888888888) [1.9. So the numbering actually went from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=85.8841111111111) [1.4 through 1.5, all the way up to 1.9,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=88.96957142857147) [and then it reset to 1.10. So again, 1.10](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=93.074) [is greater than 1.9 in this context.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=98.2412) [Support for scikit-learn and XGBoost was](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=102.542) [introduced relatively recently so the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=104.935) [version we choose has to be greater than](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=107.23) [1.4. We are prompted for the machine type](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=110.86911111111111)[that we want our prediction to be hosted](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=112.79575000000001) [on. We just go with a single core CPU. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=115.09987500000003) [now for maybe the most important input](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=118.462625) [here, which is the model URI. So we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=120.56) [to browse. This takes us into a browser,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=124.23159999999999) [where we can navigate into our bucket](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=126.58133333333329) [spikey\_ml\_engine. We find the directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=130.1732)[containing the saved model with the name](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=132.5288571428571) [model.bst. We hit Select, and this is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=136.30259999999996) [model which is going to be copied onto the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=139.05388888888893) [replicas for the prediction service. Next,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=142.3124444444444) [we are prompted for the kind of scaling we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=145.86075) [would like. Remember that there are two](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=147.753) [choices, auto scaling and a fixed number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=150.01250000000002) [of nodes. Both of these have their own](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=152.71399999999994) [strengths and weaknesses. Let's just go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=155.02333333333328) [with auto scaling with a minimum number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=156.50812500000004) [nodes equal to 2. Now let's carefully](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=159.2015714285714) [parse the fine print. This is telling us](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=162.42328571428573) [that if we specify a minimum number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=164.9654444444445) [nodes which we have done in this instance,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=166.86587500000002)[those nodes will be kept running even if](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=169.17212500000002) [there are no incoming requests. And if](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=172.034875) [those nodes are running, you can be sure](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=174.53988888888887) [we are going to be billed for them as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=176.34863636363636) [well. So while you actually deploy your](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=178.8653636363637) [models into production, be very careful](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=181.15671428571423) [with the level of auto scaling and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=183.43985714285714) [minimum number of nodes. Once we've read](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=185.91871428571426) [the fine print, let's go ahead and hit](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=188.21425) [SAVE. And now our version 1 of this model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=191.34928571428566) [comes into existence. Next, we will code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=194.05600000000004)[up a simple client which will hit this model and use it for prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=7&mode=live&start=196.47525000000005)

[XGBoost Model: Access Using Python Libraries](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live)

[Now that we have published our model,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=2.335) [let's go ahead and access it for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=3.442) [prediction. We will code up a simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=5.430874999999999) [Python client. Let's switch back to Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=7.96725) [Shell. Remember that we are in the code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=11.005499999999998) [editor, which allows us to navigate the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=13.245375000000005) [file system of our home directory on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=15.597749999999998) [GCP. Let's pick up the action from the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=18.734714285714293) [same directory, that's the xgb\_trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=20.43512500000001) [directory in which we had our training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=23.392800000000005) [data as well as the locally saved version](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=25.52466666666667) [of the model.bst. Let's go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=28.956500000000002) [create a new file there. Let's call this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=31.69111111111111) [get\_predictions.py. We are going to import](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=35.2755) [a library here. This is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=37.64712499999999)[googleapiclient.discovery library. Using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=40.933400000000006) [this, we are going to make use of a handy](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=42.642) [utility service. This is the API's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=44.876666666666665) [discovery service provided on the GCP.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=47.72333333333333) [This is a great way of exploring what APIs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=50.772) [are available, finding the associated](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=53.152) [resources and method details. Within this,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=55.748) [let's create a simple function called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=59.058428571428564) [predict\_json. The inputs into this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=61.8955) [function include the project name, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=64.32442857142857) [is spikey-ml, the model name, which is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=66.66039999999998) [xgb\_auto\_model, the version which we care](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=70.27333333333333) [about, as well as the instances, which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=72.16777777777779) [want to pass in for prediction. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=75.04199999999999) [instances input argument is going to be a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=77.95575)[list of lists. Each list is going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=80.10655555555554) [contain all of the x coordinates for one](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=82.991) [point for which we want to perform the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=85.369) [prediction. Let's go back and understand](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=88.32633333333328) [what exactly the predict\_json method does.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=90.20957142857144) [The first thing that we want to do here is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=93.015) [to going to a handle to the ML Engine API.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=95.08) [The way to do this is by using the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=98.823) [googleapiclient.discovery .build method.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=100.598) [This is going to return a resource or a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=105.495) [handle for interacting with the ML API.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=107.677) [This is a resource object. The input](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=110.755) [arguments into this build method include](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=113.49528571428571) [the service name, which here is ml, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=115.358) [version, as well as other parameters which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=118.62214285714286) [we have not used here such as the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=121.15455555555556) [discovery service URL, the developer key,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=123.167) [a model, credentials, and so on. Next, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=126.993) [want to construct the URI for the model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=130.11819999999997) [This has a specific format. It is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=133.595) [projects/, followed by the project name,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=136.68585714285712) [/models, followed by the model name. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=139.7067142857143)[next level in the hierarchy is for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=144.06266666666667) [version, so we have /versions, followed by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=145.852) [the version that we care about. We are now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=149.6415) [ready to actually invoke our model for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=153.07312499999995) [prediction, so for this, we make use of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=155.152) [the handle which we had up above to the ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=157.66240000000002) [Engine service. Then we invoke the predict](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=160.88133333333334)[method. The arguments to this method](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=164.17183333333332) [include the model URI, which is of course](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=166.02571428571426) [something we just constructed in the name](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=168.24475000000007) [variable, as well as the body of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=170.93150000000003) [request. That has to include the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=173.91912499999992) [instances. Now that we have constructed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=176.6316666666667) [the request, we have to execute it. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=178.42814285714286) [throw in a little bit of error handling,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=181.04955555555557) [and if there is no error, we go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=183.246) [return that response, which is going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=185.062) [contain the output of our model. Let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=187.346)[ahead and give this method a spin, so we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=191.45544444444445) [invoke the predict\_json method, passing in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=194.10466666666667) [the one instance. When we execute this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=196.834) [code from the command line using Python,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=199.54449999999997) [it nicely returns a price prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=201.86457142857142) [Because we had a single instance, there is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=204.38828571428573) [only a single prediction, and our model is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=207.14100000000002) [telling us that for the values of the x](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=209.988) [variables that we passed in, the forecast](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=212.77777777777771) [price of this car is $7, 107. Our client](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=214.87875) [was successful in invoking the published](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=219.80850000000007) [model. That gets us to the end of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=223.10479999999995) [module, so let's go ahead and summarize](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=225.13633333333334) [the territory that we just covered. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=226.76642857142855) [module was all about XGBoost, which is an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=229.96116666666666) [ensemble learning-based library. It's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=233.84300000000002) [pretty popular with the Kaggle community.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=236.19285714285715) [It works particularly well for relatively](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=238.376)[small datasets. It converges quickly,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=241.294) [which is an important attraction on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=244.555) [cloud, where, of course, time is equal to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=246.714) [money. XGBoost is a pretty cost-effective](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=250.25266666666658) [way to build models on the Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=252.4314285714285) [Engine. It's worth keeping in mind here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=255.76342857142848) [that support for XGBoost and scikit-learn](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=257.7261428571429) [was added relatively recently to Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=259.8275000000001) [Engine, and as of now, only online](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=262.361) [predictions are supported. Currently as of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=266.02360000000004) [November 2018, batch predictions are only supported with TensorFlow and Keras.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=8542c6c3-dd63-4d77-9b8e-0350dbef42e7&clip=8&mode=live&start=269.0152)

[Deploying Scikit-learn Models to Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live)

[Module Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live)

[Welcome to this module on Deploying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=0) [scikit-learn Models to Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=3.78) [Scikit-learn is a very popular Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=6.993) [library and open-source framework for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=9.112) [classic ML. In this module, we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=12.144) [explore its use in building and deploying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=14.697249999999999) [models on ML Engine. We will use the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=17.269399999999997) [automobile training dataset as we did with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=20.611285714285717) [XGBoost. Once again, we will have a model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=23.742166666666677) [which seeks to predict the price of an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=25.590249999999994) [automobile given its various](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=27.65362500000001)[characteristics. We will also explore the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=30.05925) [use of on-premise models by building a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=32.565124999999995) [model in Kaggle's Playground and then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=34.97414285714286) [deploying it on ML Engine. And we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=37.38025) [also check out a very cool feature of ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=40.282) [Engine, which is role-based access control](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=42.604857142857135) [for individual models. Let's plunge right](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=45.6528) [in. Scikit-learn is famous enough and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=48.98325) [popular enough that it really requires](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=51.67066666666667) [very little introduction. One high-level](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=54.084999999999994) [construct which it's worth remembering is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=56.09133333333333)[that of a scikit-learn pipeline. As the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=58.05362500000001) [name would suggest, a pipeline refers to a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=61.5424) [sequential set of transformations applied](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=63.95074999999999) [to data. The typical pipeline has several](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=66.6065) [steps which end with the construction of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=70.00266666666666) [an estimator object. If you have used](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=71.8865714285714) [TensorFlow's estimators, those are modeled](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=74.738) [on the estimators in scikit-learn. In this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=77.503) [module, we will be constructing a pipeline](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=80.65925000000001) [with a single step. That one step will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=82.57719999999999) [involve the creation of a linear](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=85.43950000000002)[regression model. That is all we need by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=87.507) [way of preamble, so let's plunge right into a series of demos.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=0&mode=live&start=90.03699999999998)

[Implementing a Simple Regression Model Using Scikit-learn in Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live)

[Let's get started right away. We will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=2.078) [write and train a simple regression model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=4.34325) [in scikit-learn and then train that model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=6.8757142857142854) [on the cloud. Because of how similar this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=9.467) [process is to the XGBoost training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=12.411) [process, we will use this as an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=14.913) [opportunity to explore some new features](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=17.208125) [of ML Engine which we did not do while](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=18.99224999999999) [working with XGBoost. Let's start from the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=22.18533333333334) [code editor in Cloud Shell. We previously](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=25.08855555555556) [had a directory called auto\_xbg. Let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=28.284285714285716)[ahead and create a new folder there called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=31.248250000000002) [auto\_sklearn\_pipeline. Within this folder,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=34.86666666666666) [let's go ahead and create another](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=37.043000000000006) [directory called sklearn\_trainer. Remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=39.3225) [that this trainer directory needs to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=41.636) [adhere to the norms for the standard](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=44.566) [Python package. Once again, inside this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=48.234666666666676)[directory, we need to create a task.py,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=51.00560000000001) [and this simple task.py will encapsulate](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=53.802) [our model code. Now the actual model code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=57.218) [is going to be very similar to that of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=59.702) [XGBoost, so we'll spend relatively little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=61.585) [time going over it. We begin with a bunch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=64.216) [of import statements. Notice how we import](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=67.41)[the Google Cloud Storage client, as well](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=70.445) [as various modules from within the sklearn](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=73.04942857142858) [package. Exactly as before, we get a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=76.80080000000002) [handle to the bucket. We load the data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=79.56920000000001) [that we want for our training process.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=82.02822222222224) [This is the same automobile data which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=83.83966666666667) [load into a pandas DataFrame. Then we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=87.16655555555555) [create a list of those columns which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=90.527) [actually want to use in our model. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=92.92566666666666) [strip out or separate the x and the y](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=96.0999090909091) [variables. We invoke the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=99.76209090909096)[preprocessing LabelEncoder in order to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=102.0922) [convert the string values for the number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=104.35699999999999) [of cylinders into numeric values. Then we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=106.95614285714288) [carry out one-hot encoding for various](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=109.66714285714285) [categorical columns, namely the car make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=112.75379999999998) [and the body style. And finally, we go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=115.83950000000002) [ahead and create our pipeline. As](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=119.40955555555553) [discussed, a pipeline in scikit-learn is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=121.82714285714285) [used for various steps in model building.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=123.886) [The usual idea of a pipeline is to have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=126.96) [multiple steps chained together, and then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=128.951)[the final step in the pipeline creates an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=131.72699999999998) [estimator for our model. In this case,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=134.367) [because our model is so simple, we can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=137.138625) [skip with all of the preliminary steps and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=138.8845) [just go directly to creating the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=141.4372857142857) [estimator. This is a linear regression](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=144.34033333333335) [model in which we try to predict the price](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=146.23874999999995) [of the automobile from all of the features](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=148.33349999999996) [about its make, number of cylinders, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=151.15671428571426) [so on. Notice how we first instantiate the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=154.1239999999999) [pipeline and then go ahead and invoke the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=156.69) [fit method on it. In the fit method, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=159.473) [need to pass in the features, that is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=162.67957142857145) [x variables, as well as the target. That's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=164.6411) [the y variable. At this point, we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=166.98725) [ready to serialize our model. As always,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=169.64599999999996) [the naming conventions must be followed.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=171.878) [Here we are making use of the joblib](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=174.532) [serialization library, and so our model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=176.412) [must be saved as model.joblib. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=179.107) [what Cloud ML Engine is expecting. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=182.93175000000002) [saved model is now ready to be uploaded to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=185.44255555555557)[Google Cloud Storage, and as before, we go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=188.3072857142857) [ahead and do so by invoking the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=190.82766666666663) [subprocess.check\_call method. This calls](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=192.503) [gsutil, cp. All of this code, once again,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=196.88400000000001) [is identical to the code in the XGBoost](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=199.386) [example. Next, let's go ahead and create a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=202.7175555555555) [setup.py file for our package. This needs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=205.61637499999998) [to be one level up in the directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=208.61711111111114) [hierarchy. And again, remember that this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=211.3435) [is not strictly required because we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=213.2391428571429) [going to be using gcloud to package up our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=215.1848571428572) [Python code, but it's always helpful for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=217.82237499999997) [us to understand what's actually going on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=220.06475) [under the hood just in case something](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=222.48266666666663) [breaks or just in case we ever need to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=224.52125000000007) [package up our app ourselves. We start by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=226.59849999999997) [importing setuptools. This is the package](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=229.91514285714285) [from within which we need the setup and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=232.01649999999998)[find packages modules. Next, we create a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=235.15500000000003) [list of all of the required packages for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=237.58633333333336) [our training package. Here we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=240.28099999999998) [included all of the packages required by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=242.593375) [our model including Google Cloud Storage,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=244.54233333333335) [scikit-learn, pandas, and NumPy. You](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=247.47033333333331) [should be aware that for standard](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=250.308) [dependencies like these, this step is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=251.598) [optional because ML Engine will be able to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=254.462) [install these using pip on the workers.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=256.504) [However, if we have specific or custom](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=260.454)[dependencies, this is where we ought to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=262.621) [include them. And now comes the really](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=264.66) [important line. This is the invocation of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=266.83) [setuptools.setup. This is what's going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=269.314) [create the package. It takes in several](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=271.968) [arguments. The first of these the name of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=274.77237499999995) [the package we want to create. Here,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=276.7946666666668) [that's automobile-pipeline. The second is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=278.91566666666665) [the version, which here is version 1.2.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=281.2305) [Then comes the packages argument. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=283.61128571428566) [a list of all of the Python import](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=286.8974) [packages that we wish to include in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=288.859) [distribution package. One option is to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=291.093) [list out all of those packages manually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=293.45322222222234) [like we did with the required packages,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=294.9282857142857) [but a cleaner way of doing that is simply](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=297.0861428571428) [invoking setuptools.find\_packages. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=300.3789999999999) [will automatically discover all the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=303.74916666666667) [packages and subpackages required. After](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=305.45) [the description and the author and author](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=308.30871428571425) [email and license comes the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=310.6811666666666) [install\_requires flag. And here is where](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=312.952) [we point to the required packages list we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=315.517) [created earlier in the file. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=317.31812500000007) [difference between the install\_requires](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=319.3198) [and the packages is that install\_requires](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=321.019) [specifies all of the dependencies that a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=323.578) [project minimally requires in order to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=325.576) [run. When a project is installed using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=328.7602857142857) [pip, this is the specification that is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=331.09871428571427) [going to be used to install its](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=332.81074999999987) [dependencies. That does it for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=335.33849999999995) [setup.py file. This will help gcloud a little while it's creating the package.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=1&mode=live&start=337.54142857142864)

[Scikit-learn Model: Train on the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live)

[Now, let's go ahead and create a YAML file](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=2.374) [called config.yaml to specify all the job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=5.215) [configuration properties. While working](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=8.652) [with XGBoost, we had specified every](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=11.29) [property in the command line invocation.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=13.310142857142855) [Here we'll just include them in a YAML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=15.772) [file. You should note that JSON config](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=17.972888888888896) [files are acceptable as well. Let's use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=20.018749999999994) [this to specify the scale tier, so here we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=22.842571428571432) [would like to go with a CUSTOM scale tier](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=26.07925) [and a masterType, a master machine type of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=28.663599999999988) [large\_model. The CUSTOM scale tier allows](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=31.707) [us to specify the exact machine types for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=34.897) [workerType, masterType,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=37.973)[parameterServerType, as well as the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=39.864) [workerCount and the parameterServerCount.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=42.215) [Recall that the standard of preconfigured](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=45.496) [scale tiers make all of these choices for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=47.639) [us. That's basically it. We can now go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=51.40877777777779) [ahead and run our gcloud command. Let's cd](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=54.46580000000001) [into the right directory. Let's run the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=57.57366666666667) [list command and ensure we have the YAML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=59.714875) [file, setup.py, and our trainer package.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=62.511) [And then let's go ahead once again and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=65.995) [define the same environment variables as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=67.405) [the last time. So we have the trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=69.651) [package part. Let's use the touch command](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=71.905) [to create an empty init.py. Then let's use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=74.635)[the ls -d command to ensure that it does](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=79.32633333333332) [indeed exist. Next up, we define another](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=82.62719999999999) [environment variable for the main trainer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=85.23100000000002) [module, that's sklearn\_trainer.task. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=87.54416666666667) [specify the BUCKET\_ID as before. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=91.5195) [time, we include a PACKAGE\_STAGING\_PATH.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=94.2822) [While working with XGBoost, we had omitted](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=97.773) [this variable, which was fine. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=99.795) [staging bucket is where the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=102.54014285714285) [archives are going to be stored. It's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=104.467) [required only if some file upload needs to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=106.84542857142857) [happen during the training and if none of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=109.6615) [the other flags implicitly specify a Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=112.39428571428572) [Storage bucket. So if we omit the staging](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=115.786) [path, the job\_dir will just be used for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=118.29033333333331) [staging. Then we specify the JOB\_NAME in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=121.21283333333331) [which incorporate the current date and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=124.00085714285714) [time, we create the job directory exactly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=127.86133333333336) [as before, we specify a different region](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=130.1741428571429) [this time around. This is just to show](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=133.30514285714284) [that it is possible. This will slow down](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=135.35466666666676) [the training process because the compute](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=137.40428571428563) [engine instances being used to host the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=139.85933333333335) [training job are going to be in a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=142.38757142857142) [different region than our underlying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=144.7644444444444) [bucket. So this isn't usually recommended,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=147.49575000000002) [but you should know that it is possible.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=150.252) [Moving on, let's specify some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=152.93) [RESOURCE\_LABELS. This is a way of making](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=154.715) [sure that our job is identified and tagged](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=157.048) [correctly. These are simply key-value](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=160.78666666666663) [pairs. Tthe first label has key team, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=163.32519999999997) [value spikey\_ai. And the second has the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=167.5631111111111) [key lead and the value alice. We are now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=170.32155555555562) [ready to invoke gcloud ml-engine jobs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=174.44612500000005) [submit. Just a few flags here defer from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=178.71249999999998) [the XGBoost version, so let's highlight](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=181.43775000000005) [those. The labels, the staging-bucket, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=184.51133333333334) [the config.yaml file. We can now go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=188.37475) [and run this job as before. We can also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=192.19400000000002) [explore this job in the UI. We did this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=194.78599999999997) [the last time around, so we'll cover this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=197.1427142857143) [portion relatively quickly. We can now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=199.8804285714286) [view logs. We wait for our job to finish](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=202.385) [running, and in a few minutes, the UI](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=204.821625) [informs us that our task has completed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=207.0466666666667) [successfully. Let's navigate into the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=210.0728333333333) [Storage browser and click on our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=212.01528571428577) [spikey\_ml\_engine bucket. As before here,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=215.0218) [we have an output directory that was just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=217.38399999999996) [created by ML Engine. If we click on this,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=220.2235714285714) [we can find the saved model. Model.joblib](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=223.44890909090915) [is the name of our model. If we go back to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=226.1502857142857) [ML Engine and check on our job, we can see](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=229.7934545454545) [that it took 5 minutes and 25 seconds to execute, and it consumed 0.24 ML units.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=2&mode=live&start=231.86879999999996)

[Scikit-learn Model: Deploy Using the gcloud Command Line Utility](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live)

[Let's quickly dive into another demo where](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=2.374) [we will see how we can deploy our model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=4.48) [using the gcloud command line utility. At](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=6.95) [the end of the last demo, we had a saved](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=9.927666666666667) [model. Let's go ahead and give that model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=12.414333333333335) [a spin. For this, let's create an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=15.086555555555554) [import.json file in which we will specify](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=18.337571428571437) [the x variables that we wish to pass in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=21.314750000000004) [for prediction. As always, we need these x](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=24.481999999999996) [variables to be expressed in the form of a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=26.878333333333323) [list separated by commas. Let's use this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=29.55449999999999) [file to make sure that our model works](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=31.750500000000002) [correctly before pushing it out to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=34.22742857142856) [production. So let's create a MODEL\_DIR](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=36.532799999999995) [directory and INPUT\_FILE variable pointing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=38.98557142857141) [to our input.json and a FRAMEWORK](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=41.21060000000001) [environment variable which contains the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=44.58740000000001) [value SCIKIT\_LEARN. And now let's use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=46.80160000000001) [gcloud ml-engine local predict. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=50.5955) [going to use our saved model from the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=54.202) [Cloud Storage bucket without actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=56.44916666666666) [going through the ML Engine published](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=58.48) [model. This is an important step which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=61.92142857142856) [should be careful to undertake before](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=64.062125) [pushing our models out into production.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=65.91457142857142) [Note that this is different from just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=68.99) [running our Python code locally using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=71.144) [Python rather than ML Engine because here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=73.48) [the prediction is still going to make use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=76.123) [of the saved model which was trained on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=77.97) [the cloud. So let's go ahead and run this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=80.45479999999995) [code, and we find that the returned value](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=82.94937499999996) [is 41, 373. This is the predicted price,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=86.78839999999998) [the output of our regression model. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=90.00087500000002) [try a little experiment. Let's open up our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=93.196) [input.json file and reduce the engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=95.93428571428574)[capacity. This should also have the effect](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=99.1925) [of reducing our predicted price. If we now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=101.61385714285713) [try and rerun the exact same command with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=105.53333333333333) [the same input file, the output prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=108.042) [has reduced from $41, 000 to $36, 800. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=110.93499999999999) [at least we know that our model is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=114.7036) [responsive. The output is changing based](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=118.34920000000001)[on the inputs. This gives us enough](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=121.10062500000002) [confidence to go ahead and deploy this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=123.5435) [model. Let's define the model name as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=126.09174999999999) [sklearn\_automobile\_model and then invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=128.184) [gcloud ml-engine models create. Notice](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=131.989) [that we also use the same labels that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=135.8775) [had specified while training the model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=138.659) [This seems to have gone through](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=141.155) [successfully. Let's confirm that using the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=142.58042857142866) [gcloud ml-engine models list command, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=144.69966666666662) [let's also filter on the team:spikey\_ai.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=148.08516666666668) [And indeed, we do get back the model that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=150.01957142857142) [we had just created. Notice how at this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=154.4587142857143) [point there's no version associated with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=157.27649999999997) [this model, so let's go ahead and change](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=159.05914285714286) [that. Let's switch back to the web](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=161.6565) [console, navigate to the Models tab in ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=164.16642857142855) [Engine, and there we see our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=167.02375000000006)[sklearn\_automobile\_model. No versions have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=168.397) [been created yet. Instead of clicking on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=172.20350000000002) [that Create a version link, let's switch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=174.61971428571428) [back to the command line and use gcloud to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=176.40714285714282) [do this. Let's create a variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=180.13999999999996) [VERSION\_NAME with v1 and then invoke the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=182.8865) [gcloud ml-engine versions create command.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=185.49271428571436) [We've got to specify the version name, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=189.56) [model name, and the origin, that's the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=192.1148333333333) [Google Cloud Storage bucket where the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=194.15099999999998) [original model was saved after training,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=196.1292857142857) [then the runtime, the framework, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=198.82049999999998) [here is scikit-learn, and then the version](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=202.59285714285716) [of Python, which here is 3.5. If we do not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=205.3132857142857) [specify a version of Python, 2.7 will be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=209.42500000000004) [used by default. Finally, we can also add](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=212.112) [all of the labels which we created a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=214.71899999999997) [moment ago. Let's run this command, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=217.74566666666666) [the version will get created for us. If we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=220.32542857142852) [now switch back to the web console and hit](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=223.75488888888887) [refresh, we can see that version v1 does appear.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=3&mode=live&start=225.92155555555553)

[Scikit-learn Model: Share Model with Other Users Using Role Based Access Control](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live)

[Let's jump into a really fascinating](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=2.445) [little demo in which we will see how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=4.156714285714285) [role-based access control can be used to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=6.774749999999997) [share models with other users. At the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=9.1855) [start of this demo, let's check what user](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=12.081666666666665) [we are currently signed in as by clicking](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=14.567) [on the user icon in the top right. And we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=16.699555555555555) [can see that we have built this model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=20.12933333333333) [using the Spikey Sales Google identity.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=21.997142857142858) [Let's now return to ML Engine and share](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=25.448) [access to our model with a specific user.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=27.589) [So we click on Models, select the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=31.439) [sklearn\_automobile\_model, which we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=34.47633333333333) [just created, click on the SHOW INFO PANEL](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=36.116125) [in the top right. There we will find the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=38.62200000000001) [PERMISSIONS tab, which will allow us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=41.594714285714296) [edit or delete permissions and assign](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=43.361) [access to users. Currently there's just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=45.742999999999995) [the one user. That is the ML Engine Model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=48.5155) [Owner, which is Spikey Sales. Let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=51.626571428571445) [ahead and add a member by clicking on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=54.10588888888889) [ADD MEMBER button there. We are now going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=56.519000000000005) [to go ahead and give access rights to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=58.82099999999999) [another member. That is spikey.engineer.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=61.172) [Now remember that this is role-based](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=64.29) [access control, so we need to specify a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=65.97000000000001) [role as well. So we select the Machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=68.52500000000002) [Learning Engine set of curated roles and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=70.89142857142858) [pick ML Engine Model User. There's a tool](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=73.92424999999999) [tip there which tells us exactly how much](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=77.73875) [access we are allowing this particular](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=80.592875) [user. We are only allowing this user](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=83.13950000000001) [permissions to read the model and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=86.517) [versions and use them for prediction. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=88.74633333333335) [have not given any other roles such as ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=91.80722222222222) [Engine Admin, Developer, or Model Owner.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=94.08800000000001) [So ML Engine Model User is all that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=97.538) [have allowed spikey.engineer to access.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=99.96620000000003) [This identity now shows up in the list of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=104.114) [roles. Let's navigate into the identity](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=106.90850000000005) [and access management console just to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=108.86328571428572) [ensure that we do not give spikey.engineer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=111.33642857142858) [any higher levels of privilege at the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=114.31333333333335) [product level. So let's go ahead and add](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=117.18028571428572) [spikey.engineer, and the role that we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=119.98828571428568) [give at the project level is simply](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=121.84099999999997) [browser. This will only give](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=124.76919999999998) [spikey.engineer the ability to browse the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=126.87833333333336) [resources in this project. That's the bare](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=129.16899999999995) [minimum which that user will need in order](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=131.64671428571427) [to access our model. Now let's give this a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=134.307)[trial. Let's sign out of our Spikey Sales](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=137.12466666666668) [identity and sign back in as Spikey](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=140.01724999999993) [Engineer. Once we do that, we'll be able](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=143.4755714285714) [to test how much as we have to the model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=145.23459999999994) [Notice immediately how the project pane](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=148.837) [changes. If we navigate into ML Engine and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=152.28100000000006) [check on the Jobs, we are informed that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=154.4471111111111) [do not have sufficient permissions to view](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=156.73133333333337) [this page, and the same goes for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=158.81066666666663) [Models tab as well. So we can't even tell](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=161.36277777777784) [what models or versions exist. But despite](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=164.07188888888885) [all of this, if we switch into the code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=166.70085714285716) [editor and code up a little client, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=169.61444444444453) [will be able to successfully access our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=171.98700000000002) [model for prediction. Let's try that out.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=174.81050000000002) [Let's switch to the code editor, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=177.32449999999994) [create an input.json file just as before,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=179.70149999999998) [and then let's go ahead and try and invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=182.624)[our model. So we'll have the same model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=186.40842857142854) [name, sklearn\_automobile\_model version v1,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=189.01185714285722) [we set up the INPUT\_DATA\_FILE, and invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=193.072) [gcloud ml-engine predict. And the magic is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=196.2628) [that even though we were completely barred](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=199.14711111111112) [from viewing the details of the model, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=201.17533333333333) [were indeed able to invoke this model for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=204.05177777777772) [prediction. And it's important to note](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=207.601) [that this role that we were granted was](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=209.71) [restricted only to this specific model, so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=211.89124999999999) [if we now try to use our previously built](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=214.00044444444447) [model from XGBoost, that is if we tried to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=217.097) [invoke gcloud ml-engine predict, we would](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=220.27737500000003) [get an error. This is a 403 error telling](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=223.26762499999995) [us that access to this model has been denied.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=4&mode=live&start=226.61000000000004)

[Build and Train a Scikit-learn Model on Kaggle](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live)

[The beauty of Cloud ML Engine is that it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=2.689) [can be used to port models which we build](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=4.862) [on premise onto the Google Cloud Platform.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=7.571) [To demonstrate this, let's first build a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=10.923) [model using the Kaggle Playground. In the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=12.79) [next part of this demo, we will go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=15.780600000000002) [and host that model on the GCP. So let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=17.7555)[pick up the action by typing in kaggle.com](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=21.256555555555558) [into our browser window. Let's try and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=23.3866) [sign in with our Google identity. That's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=26.755125) [Spikey Sales. Because this is the first](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=29.009) [time, we need to create our Kaggle](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=31.00377777777778) [account, so we specify a username, a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=32.69700000000001) [display name, and go through various other](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=35.18085714285715) [formalities. Once this account comes into](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=39.105999999999995) [existence, we can access various datasets,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=41.76499999999999) [and not coincidentally, this includes the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=44.712) [automobile dataset which we were using for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=47.428) [both our XGBoost and scikit-learn models.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=49.919) [As an aside, just check out how cool](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=54.044) [Kaggle is. See how easy it is for us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=56.391199999999984) [visualize the different columns. There's a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=58.69400000000001) [good reason why Kaggle is so popular, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=60.66700000000001) [it's a great resource in case you're](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=62.92900000000001) [looking for datasets. In any case, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=65.39971428571428)[now head back into the Kaggle Playground.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=67.71266666666666) [This is effectively a Python kernel,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=69.75233333333334) [something like a hosted Jupyter Notebook.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=72.25966666666667) [We can go ahead and import libraries,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=75.05985714285714) [interact with the operating system, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=78.1802) [even save files into the current](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=80.78814285714284) [directory. The Playground will then give](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=83.88033333333334) [us a way to download those files. We shall](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=85.52488888888888) [give this a spin in just a moment. Before](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=88.47619999999999) [that, also pay attention to the fact that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=90.78125) [all input data files are going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=92.70955555555555)[available in the input directory that's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=94.458) [one level up from our code. We can access](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=96.97285714285715) [and browse this in the tab on the right.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=100.2065) [So we can see the Automobile\_data.csv file](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=102.35342857142857) [right there. Let's turn our attention back](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=105.45) [to the main window. We can execute Python](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=108.67966666666668) [code there by clicking run or by pressing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=111.16614285714284) [Shift+Enter. We will now go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=115.07516666666668) [rebuild very quickly our regression model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=117.13899999999998) [right here in the Kaggle kernel. All of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=118.9942857142857) [these steps are the exact same ones that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=121.70350000000002) [we executed in the Cloud Shell window](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=124.026) [previously, and by the end of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=126.29328571428573) [process, we will have a sklearn model. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=129.04520000000002) [as before, we instantiate the pipeline, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=132.0572857142857) [call fit on the pipeline to the features](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=134.62033333333332) [in the target, and then we can go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=136.679) [and serialize our model using pickle. Here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=139.44740000000007) [we are going to save this to a file called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=142.5228) [model.pkl. We hit Commit up top, and all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=144.709) [of this code is now going to be committed.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=148.91154545454546) [At this point, we can browse into the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=152.487) [Output tab, and there we will find the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=154.509) [model.pkl file which we just serialized.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=157.101) [So we are going to download it to our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=160.055) [local machine. And in just a moment, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=162.59954545454545) [are going to host the same model on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=164.15079999999998) [Google Cloud Platform using Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=5&mode=live&start=166.36425)

[Deploy On-premise Model Using ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live)

[Let's go ahead and end this model on a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=2.327) [high note. We are going to deploy the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=3.8848181818181806) [on-prem model which we just serialized on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=6.466000000000003) [Kaggle on the Google Cloud Platform. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=8.833500000000003) [first step is to port our model onto](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=11.651777777777777) [Google Cloud Storage into a location where](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=14.138) [it will be accessible to the ML Engine. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=16.576) [let's find our spikey\_ml\_engine bucket and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=20.00433333333333) [create a folder there. Let's call this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=23.092000000000002) [sklearn\_kaggle\_model. We create this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=27.633250000000004) [folder and then upload our saved model.pkl](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=29.5575) [file into it. And really, that's all that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=32.9222) [there is to it. Once we have model.pkl](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=36.35133333333333) [available on Cloud Storage, this can be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=38.871142857142864) [picked up by Cloud ML Engine. Just like](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=41.39599999999999) [any other model, it was trained on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=44.00075) [cloud. We navigate to the Models section](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=45.884750000000004) [and create a new model. And now all of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=48.32816666666667) [these steps ought to seem pretty familiar.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=51.6131111111111) [We specify a model name, save the model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=53.51) [At this point, no versions exist, but we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=56.748000000000005) [can change that. Click on the NEW VERSION](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=59.495999999999995) [button. This is version 1 of our model. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=62.43566666666666) [go with Python 3.5. The framework used was](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=66.65780000000001) [scikit-learn, and then it's time for us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=71.46579999999997)[browse to the place where we've uploaded](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=73.54169999999998) [our pickle file. so we click on the BROWSE](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=75.58114285714285) [button. Inside the spikey\_ml\_engine, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=78.72814285714288) [find the directory we just created and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=80.56633333333332) [select the model.pkl file. Again, just to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=83.41899999999998) [be clear, this is a file which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=86.56299999999997) [downloaded from the Kaggle Playground onto](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=89.34766666666664) [our local machine. We then uploaded it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=92.03500000000003) [from our local machine into Google Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=95.00350000000002) [Storage, and we are now hosting this file](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=97.064125) [using Cloud ML Engine. We give it a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=100.04671428571429) [moment, and version 1 comes into](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=103.17833333333333) [existence. Let's now switch back to Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=104.93057142857143) [Shell and use this model to complete the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=108.00625) [loop. This time, the model name is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=110.728375) [spikey\_onprem\_model. It's version v1 as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=113.013) [usual. And now we can go ahead and set up](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=116.71599999999998) [an input.json file with the x variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=119.06389999999998) [values. We paste these in from the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=123.55499999999999) [original automobile data, save the file,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=126.15814285714282) [and go ahead and use it in prediction. By](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=128.864) [now, we are quite comfortable use gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=132.02212500000002)[ml-engine predict, so we won't dwell on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=134.039) [all of the parameters. And once we run](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=136.22300000000007) [this command, we get back the prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=139.00249999999994) [We have successfully used and deployed our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=141.843) [on-prem model from the Google Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=6&mode=live&start=144.402)

[Summary](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live)

[That gets us to the end of this module.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=2.414) [Let's summarize the territory we covered.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=4.018) [This module was all about scikit-learn,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=7.206) [which is an extremely popular classic](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=8.774) [machine learning library in Python. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=11.511) [built scikit-learn models and deployed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=14.331666666666667) [these to the ML Engine. We also saw how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=16) [easy it was to build a model on-premise or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=19.053199999999997) [on Kaggle and then deploy it to ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=21.647428571428577) [Along the way, we also explored role-based access control for ML models.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=55495336-6e2b-42b1-8788-40aff6214107&clip=7&mode=live&start=24.903)

[Deploying TensorFlow Models to Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live)

[Module Overview](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live)

[Welcome to this module on Deploying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=0) [TensorFlow Models to Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=4.022) [While scikit-learn and XGBoost were only](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=7.078) [recently added to the Cloud ML Engine,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=9.686) [TensorFlow is the oldest framework](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=12.989) [supported on this service and is also the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=14.554) [most tightly coupled with it. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=17.315) [mechanism for that coupling is a specific](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=19.546142857142858) [environment variable called TF\_Config.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=21.883) [This is a JSON environment variable which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=24.891) [is used to pass information from ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=27.247) [into TensorFlow. In this module, we will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=30.387) [undertake a series of demos in which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=33.332) [will leverage this link and correctly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=35.47871428571428) [structure our TensorFlow package so that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=37.8256) [we can implement both distributed and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=40.42485714285714) [local training using ML Engine. Then we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=42.3565) [will move on to the topic of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=45.52400000000001)[hyperparameter tuning, and we will see how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=46.819) [quickly the amount of compute consumed by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=49.292) [our jobs goes up once we start tuning our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=52.116)[hyperparameters. And we will round out the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=56.645700000000026) [module by implementing both batch and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=58.532) [online prediction. TensorFlow is the only](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=61.41366666666668) [framework which currently supports both these modes of prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=0&mode=live&start=63.723333333333336)

[ML Engine and Tensor Flow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live)

[In this video, we are going to spend some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=1.976) [time understanding the coupling between ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=3.181) [Engine and TensorFlow. In the ML workflow,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=5.8974) [which we discussed earlier on in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=8.842999999999998) [course, we spoke about how model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=10.864499999999998) [development could be performed in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=12.55342857142857) [TensorFlow, Keras, XGBoost, or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=13.897) [scikit-learn. Keras uses TensorFlow as a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=16.5354) [back end on the Cloud ML Engine, so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=18.804999999999996) [really, it's three packages that we can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=21.6242) [use, and of these, TensorFlow is by far](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=23.826714285714285) [the most popular and most important.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=25.910333333333345) [TensorFlow needs little introduction. It](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=28.761) [is currently the most popular deep](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=30.97442857142857) [learning library, and it's really picked](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=32.573)[up in popularity in recent years. In this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=34.758714285714284) [course, we'll be using the estimator API.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=38.69585714285715) [This is a high-level part of TensorFlow,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=40.867285714285714) [which makes it really simple to build](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=43.466428571428565) [complex models. This course is certainly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=46.617000000000004) [not meant to be your first TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=49.07500000000001) [course, so if you're not familiar with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=51.01128571428571) [TensorFlow or how estimators work, I would](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=53.278000000000006) [highly recommend your spending some time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=55.99833333333333) [with one of the many excellent TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=57.814499999999995) [courses here in Pluralsight's catalog. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=60.015) [this module, we will be giving a thorough](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=63.766375) [workout to the capabilities of Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=65.98316666666666) [Engine. Remember that there are training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=68.47957142857142) [and prediction services, and then there](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=70.591) [are further subtypes. There is local and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=73.023) [distributed training, and on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=75.76333333333335) [prediction side, there is both online and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=77.44187499999998)[batch prediction. We are going to go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=79.874) [and implement all four of these in this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=82.11699999999998) [module. Of these, in particular,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=84.87899999999996) [distributed training for TensorFlow is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=86.75) [technically very complex. TensorFlow can](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=88.974) [be run in distributed mode, even on a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=92.86342857142857) [native cluster, but no one in their right](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=95.89133333333332) [mind would do it if they don't have a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=98.27777777777777) [managed service like Cloud ML Engine. It's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=100.64875) [just too complicated. Under the hood,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=103.27725) [Cloud ML Engine is going to assemble](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=105.102)[multiple nodes into a training cluster.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=107.37185714285714) [Our training job is going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=110.695) [replicated, and it's going to run on each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=113.42149999999998) [node in this cluster. Each of these copies](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=115.57690909090908) [is called a replica. There are three types](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=118.31100000000002) [of replicas, masters, workers, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=121.33933333333334) [parameter servers. Let's understand these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=124.297) [three real quick. Every cluster has](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=126.36749999999998) [exactly one replica designated as the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=128.6905) [master, which manages all of the others](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=130.92985714285714) [and reports the status. In addition, there](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=133.51085714285713) [are workers, which undertake specific](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=136.06000000000003) [parts of the parameter optimization. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=137.731) [finally, there are parameter servers,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=141.61583333333334) [which coordinate shared model state.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=142.795)[Parameter servers are required because](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=146.174) [neural networks require a complex](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=148.032) [optimization process, which is not that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=150.929) [easy to parallelize. And that's why the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=153.529) [parameter servers need to maintain shared](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=156.38342857142854) [state between the workers. Now remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=158.41657142857142) [that all of these are just cloud VM](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=160.69525) [instances, which could restart or crash at](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=163.06785714285712) [any point, and so we've got to make sure](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=165.47633333333332) [that our code is resilient to this. For](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=167.94877777777776) [this reason, it's important for us to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=170.436) [ensure checkpointing in our TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=171.976) [code. In our ML Engine workflow, we had](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=174.82266666666666) [discussed the model training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=178.41075000000004) [evaluation step as one where ML Engine and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=180.42714285714285) [the framework need to work especially](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=182.83975000000007) [closely together. This coupling is very](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=185.27859999999998) [tight, and it is implemented using an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=187.50299999999996) [environment variable called TF\_CONFIG.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=189.99739999999997) [TF\_CONFIG is a JSON environment variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=192.564) [which is set by ML Engine and is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=195.85) [accessible to our TensorFlow application.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=198.34777777777785)[Estimators and other important TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=201.144) [classes will automatically refer to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=203.009) [value of TF\_CONFIG during their execution.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=205.612) [Remember that TF\_CONFIG is JSON, and for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=209.093) [that reason it's keyed. The keys in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=212.20799999999997) [TF\_CONFIG include cluster, task, and job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=215.18050000000002) [information. We will not go through these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=217.91774999999998)[lines one by one. But at a high level, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=220.023) [cluster spec, job spec, and the roles of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=223.44724999999997) [each individual task are defined in TF\_CONFIG.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=1&mode=live&start=226.16575000000006)

[Hyperparameter Tuning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live)

[Let's briefly talk about hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=1.988) [tuning and how it's accomplished on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=4.13) [Cloud ML Engine. Simply put,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=6.054) [hyperparameter tuning is the process of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=8.502) [trying different model configurations,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=10.741600000000002) [evaluating all of those candidate models](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=12.4694) [using some metric, and then picking the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=15.794599999999999) [best one. If a model is properly tuned,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=18.82666666666667) [the improvements in performance can be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=21.95) [dramatic, and so it's important for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=23.709) [hyperparameter tuning to be supported by a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=25.595)[service like the ML Engine. Please note](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=27.8) [that hyperparameter tuning is a different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=30.73242857142857) [type of job than either training or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=32.957) [prediction, and it is only supported for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=35.60877777777778) [TensorFlow. This is a really](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=39.21922222222223) [computationally intensive process because](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=41.06966666666666) [effectively the ML Engine is going to run](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=42.708999999999996) [multiple trials, that is multiple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=46.213) [independent versions of different training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=48.413500000000006) [jobs, one for each candidate model. Each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=50.60433333333333) [of those trials is a complete execution of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=53.365625) [the training job and will therefore incur](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=56.346285714285706) [roughly as much billing as the original](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=58.715999999999994) [stand-alone training job. So we've got to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=61.33919999999999) [keep a careful eye on our costs.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=63.56750000000001)[Hyperparameter tuning can quickly get](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=66.065) [pretty expensive. As we shall see shortly,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=67.894) [to implement hyperparameter tuning, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=71.871) [need to structure our application in a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=73.79375) [specific way. We need to define one](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=75.668) [command-line argument for each tuned](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=78.357) [hyperparameter, and we also need to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=80.408) [specify how we would like that particular](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=82.9322857142857) [hyperparameter to be tuned in a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=85.206) [config.yaml file. The responsibility for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=88.62975000000003) [coming up with intelligent guestimates for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=90.67199999999998) [those hyperparameter values lies with ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=93.06249999999999) [Engine and not with the application](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=95.98233333333334) [developer. But inside that YAML file, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=98.6995) [will specify some information about each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=101.72449999999999) [hyperparameter to be tuned. This will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=104.26639999999998) [include the type and the range of values](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=106.26387500000001) [to be trialed. Optionally, we can also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=108.55250000000001) [specify the type of scaling, which could](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=110.39350000000002) [be linear, logarithmic, or reverse](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=112.55219999999998) [logarithmic. In addition to all of this,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=115.76175) [we also need to specify an evaluation](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=118.1642) [metric. This is a metric which ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=120.59900000000002) [will use to find the best trial, in other](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=122.94200000000004) [words, to find the best candidate out of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=126.47249999999997) [all of the trained models. This metric is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=128.97744444444444) [often called the hyperparameter metric,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=132.27271428571427) [and it's often a human understandable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=134.947) [metric, something like accuracy or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=136.726) [precision or recall, in contrast to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=138.532)[loss metric that's used during the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=141.851) [training process. The typical choice for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=144.5374999999999) [classification problems for the loss](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=146.42020000000002) [metric is cross entropy. That has little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=148.31749999999997)[intuitive significance, but it is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=151.56940000000003) [mathematically suitable for optimization.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=154.23557142857146) [In evaluation, we tend to use a human](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=156.367) [understandable metric like accuracy or precision or recall.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=2&mode=live&start=158.542)

[Implementing task.py for Distributed Training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live)

[Let's get started with TensorFlow on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=2.306) [Google Cloud Platform. We are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=4.298) [write and train a simple TensorFlow model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=6.8116666666666665) [As usual, we will begin by training it](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=8.290714285714285) [locally before moving to the cloud. Now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=11.285999999999998) [TensorFlow is a fairly involved package.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=14.26842857142857) [We are not going to spend much time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=16.491) [talking about the intricacies of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=18.130999999999986) [TensorFlow itself. The model that we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=20.164666666666662) [going to use is going to be the simplest](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=21.950875000000007) [and most generic possible, which is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=23.944750000000006)[iris classification dataset. Let's start](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=25.99) [with a fresh, empty directory. Let's call](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=29.753571428571426) [it iris\_tf. Inside this, we need to create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=32.5275) [a new subfolder for our trainer package,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=35.166250000000005) [which we will call tf\_trainer. Now in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=38.111) [case of both scikit-learn and XGBoost, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=42.04749999999999) [had just the one source file called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=43.9675) [task.py. We are going to need task.py in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=46.50971428571429) [TensorFlow as well. However, in addition](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=50.242399999999996) [to task.py, we will separately implement](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=53.3462) [our model in a model.py. We'll get to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=56.20522222222223)[implementing that in just a moment. First,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=59.770999999999994) [let's get the task.py out of the way. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=62.443444444444445) [start with a bunch of imports. The only](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=65.45442857142858) [important point here to note is that are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=68.43950000000001) [also importing the model.py, which exists](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=71.31433333333332) [in the same directory. We are going to get](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=74.17366666666668) [around to implementing that in just a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=76.40687500000001) [moment. Then let's go ahead and import](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=79.41399999999997) [tensorflow as tf and also the hparam](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=81.92928571428573) [module. This a part of the contrib](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=86.16757142857146) [package. Hparams is a class which will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=89.2724285714286) [hold a set of hyperparameters as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=91.0244) [name-value pairs. Next, we implement a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=94.079) [get\_args function. Here we make use of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=98.19999999999999) [argparse.ArgumentParser. This is a way of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=100.36) [taking in all of the hyperparameter values](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=103.89950000000003) [as command-line arguments. In case you](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=106.3642) [haven't come across it before, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=109.14850000000001)[argparse.ArgumentParser is a great way to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=110.92333333333333) [make your command-line interfaces user](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=114.04357142857143) [friendly. This is a way to parse the input](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=116.58642857142857) [arguments which are passed in via a list](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=119.55900000000005) [called sys.argv. The zeroth element of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=123.54277777777781) [this list is the name of the program](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=125.55877777777776) [itself. Notice here how we are parsing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=128.60725) [arguments for the training files, eval](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=131.02185714285707) [files, and also for various](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=133.0836666666666) [hyperparameters such as the learning rate,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=134.98066666666668) [batch size, and so on. Next is a function](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=137.57255555555554) [which is going to extract configuration](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=140.68949999999998) [information from the environment. Remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=142.78233333333336) [that there is this environment variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=145.55433333333332) [called TF\_CONFIG. This is a JSON](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=147.75633333333334)[environment variable which is going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=150.6778571428572) [set for us by Cloud ML Engine. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=153.03190909090904) [includes some recommended code which is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=155.86533333333333) [used to ensure that communication between](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=157.823) [the master and the workers is reliable and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=159.66299999999998) [resilient to VM crashes. This code also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=162.47899999999998) [includes a way to ensure that the master](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=165.50925)[only communicates either with itself or](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=167.926) [the parameter servers and likewise for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=170.0285714285714) [workers. Next is an important function](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=173.53025) [called train\_and\_evaluate. Notice how the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=175.74433333333326) [hparams object is passed in here. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=178.62300000000002) [hparams object is going to contain](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=181.5517142857143) [key-value pairs with one key for each](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=183.134)[hyperparameter. The first thing we do here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=187.22637500000008) [is to create the runtime config for our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=189.07279999999994) [TensorFlow estimator. This includes data](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=192.4133333333333) [such as the session\_config, which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=194.423) [obtain from the environment, that's the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=196.11900000000003) [function we just defined a moment ago, as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=197.95950000000002) [well as the checkpointing information and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=200.53400000000002) [the model directory. Notice how the model](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=203.04299999999998) [directory is simply set to be equal to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=205.56771428571432) [hparams.job\_dir. This was the runtime](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=210.32014285714294) [config for the estimator. Now let's create](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=212.34400000000005) [the training spec. Here we need to set up](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=214.8895) [a model.input function. This will point to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=218.5713333333334) [the training data, the batch size, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=220.77785714285713)[also insist that we shuffle the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=222.53799999999995) [data during the training process. Next, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=224.57600000000002) [use the tf.estimator .FinalExporter to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=227.3283333333333) [export the serving graphs, as well as the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=230.138) [checkpoints. The one bit worth noting here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=234.23142857142858) [is that we also export the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=235.9736666666666) [model.serving\_input\_receiver function.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=238.084) [This is a way of making sure that during](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=241.104) [prediction, the incoming data is subjected](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=243.325) [to the same transformations as work](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=246.185) [carried out during training. This avoids](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=248.651) [the problem known as training- serving](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=251.05885714285714) [skew. Just like we had to define a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=252.76711111111112) [training spec for the training process, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=256.376) [now need to create an eval spec for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=259.0652) [evaluation process. Once again, we specify](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=261.083) [the eval\_input function, the number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=263.81233333333336) [steps, and the exporters. Notice how we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=266.5080000000001) [shuffled the data during training, but we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=269.4078571428571) [did not do so during evaluation. Finally,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=271.7176666666666) [we are ready to actually create our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=274.4875714285714) [estimator using the model.build\_estimator.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=277.07)[Notice here that model refers to the class](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=280.718) [defined in model.py, which we will get](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=283.632) [around to writing in just a moment. We are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=286.231) [going to make use of a DNNClassifier. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=289.6243333333333) [is going to have three layers of neural](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=292.44375) [networks with 10, 20, and 10 neurons](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=295.47211111111113) [respectively. Because we're using this for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=299.48388888888906) [the classic iris classification problem](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=301.2248571428573) [where there are three possible values for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=302.75600000000003) [the output label, we specify num\_classes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=305.2102857142857) [equal to 3. It's also worth pointing out](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=307.6025) [that TensorFlow estimator objects know how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=310.13866666666655) [to read the TF\_CONFIG environment](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=312.60099999999994) [variable, and in this way, they will be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=315.8508000000001) [able to read the cluster specification](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=317.71820000000014) [task ID and other properties which have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=319.3935714285714) [been set by Cloud ML Engine. Finally, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=322.1713333333334) [can train and evaluate our estimator. Now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=325.61375) [here is a dire warning contained in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=329.0422) [docs. Always use tf.estimator](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=331.94780000000003) [.train\_and\_evaluate if you're planning to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=335.23974999999996) [do distributed training on the cloud. Do](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=337.27666666666676) [not use the simpler tf.estimator .train](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=340.03037500000005) [method. We are almost done with the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=343.29062499999986) [task.py. All that's left is for us to code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=345.8139999999999) [up the main. Here we get the args using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=347.946909090909) [the argument.arg parser, instantiate an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=351.1872499999999) [hparams object, and invoke the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=354.06280000000004) [train\_and\_evaluate method which we just spent all this time implementing.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=3&mode=live&start=356.67480000000006)

[Implementing model.py for Classification Using TensorFlow Estimators](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live)

[We are done with the task.py. Let's now go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=2.233) [ahead and implement the model.py which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=5.1049999999999995) [referenced inside our task module. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=7.580000000000001) [task.py was very specific to Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=11.086571428571428) [Engine. It was all about the handshake](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=14.563666666666665) [between the system and our code. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=16.58157142857143) [model.py is a lot simpler. This is a lot](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=20.167833333333334) [more just conventional TensorFlow code. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=23.307714285714287) [start with a bunch of import statements](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=25.250375) [including tensorflow as tf, and the rest](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=27.980999999999998) [of this file is going to be a fairly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=30.756000000000004)[straightforward implementation of iris](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=33.84839999999999) [classification. We start with the column](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=36.193333333333335) [names which we care about. These include](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=38.075624999999995) [the SepalLength, SepalWidth, PetalLength,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=40.86599999999999) [PetalWidth, and the Species. We define a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=43.465666666666664) [quick little lightweight](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=47.023750000000014) [get\_feature\_columns method, which is going](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=48.745666666666665) [to return a list of lists. Each list](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=51.18299999999999) [within this list of lists is going to be a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=53.41825000000001) [feature column, so for this, we use the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=56.25999999999999) [tf.feature\_column .numeric\_column method.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=60.2945)[Notice that we only do this for the first](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=62.784) [five elements of the column names. Moving](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=64.535) [on, we actually build our estimator, so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=67.80714285714286) [here we pass in the config, the number of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=71.05228571428572) [hidden units, the learning rate, and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=73.95500000000001) [number of classes. This build\_estimator](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=76.04474999999996) [method is what was invoked from task.py.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=77.57033333333334)[Notice how we have specified default](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=82.241) [values, but all of these are actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=84.038) [going to be populated when task.py invokes](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=86.07499999999997) [our build\_estimator method. Inside this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=88.77371428571432) [method, we instantiate an optimizer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=92.56433333333334) [object. This is an AdamOptimizer. There](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=94.73833333333334) [are many different types of optimizer](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=98.64028571428571) [objects available in TensorFlow, all of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=99.932) [which perform the same gradient descent](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=102.50585714285715) [process. This is an optimization process](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=104.366) [to find the best parameter values for all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=106.88271428571429) [of our neural network coefficients. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=110.48271428571428) [then we are able to return the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=114.06524999999999) [tf.estimator .DNNClassifier object. So](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=115.698) [this is going to be our model object. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=120.6631111111111) [is a high-level estimator, so this is an](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=123.30025) [API to build a neural network in which we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=125.76466666666667) [just pop in the optimizer, the config, the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=128.55962499999998) [values of the x and y data, and the number](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=131.283) [of classes. And that is basically it.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=134.68519999999998) [TensorFlow will take care of the rest and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=137.344) [train our deep neural network classifier](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=138.897) [for us. But there's still some important](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=141.719) [wiring up left for us to do. We have to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=144.174) [make sure that our estimator object is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=147.39700000000002) [able to make sense all of the training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=149.24455555555556) [data that's fed in. And for this, we need](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=151.8608) [to create an input parser, which in turn](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=155.0983333333334) [returns a function. That function is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=158.22957142857143) [decode\_csv function, and that function](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=161.75699999999998)[needs to return a pair. The first element](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=164.12650000000002) [in the pair is going to be a dictionary of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=166.89166666666668) [all the features, and the second element](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=169.37957142857144) [in the pair is going to be a list of all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=171.37111111111116) [the target or the y values. I'm going over](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=174.5275) [this code relatively quickly because this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=177.46299999999997) [is pretty specific to TensorFlow. This is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=179.3895714285714) [not really about cloud machine learning as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=181.94828571428573) [such. If you're not familiar with how](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=184.34714285714284) [estimators are wired up then I suggest you](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=186.562) [go back to a class like the Foundations of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=188.82855555555554) [TensorFlow class here on Pluralsight. In](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=191.4777142857143) [addition, we also need to set up a serving](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=194.4378888888889) [function. This is going to get the data in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=196.571) [during prediction, and it's important for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=199.63350000000003) [us to ensure that during prediction,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=201.95800000000003) [incoming data is subjected to the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=203.92399999999998) [processing as during training. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=207.51100000000002) [finally, we now need to set up the input](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=209.88366666666667) [function, which is going to feed in all of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=212.21812500000001) [the features and labels for training. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=214.80200000000002) [is fairly involved, but it's also fairly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=217.35314285714284)[standard for any time that you're using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=219.82314285714287) [estimator. We are now done with both our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=221.868375) [model.py and our task.py. Let's go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=225.5542) [and complete the package process by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=228.69571428571425) [creating the empty init.py file. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=231.01771428571428) [time, Let's also go ahead and create our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=234.4965) [training and eval data right here in this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=236.889) [data directory. We are doing so merely as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=240.30933333333334) [a convenience because we are going to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=243.54550000000006) [first go ahead and run this TensorFlow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=245.10977777777774) [code right here in this machine, and we do](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=247.073125) [not want to have to go all the way to a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=249.87140000000005) [Cloud Storage bucket for it. Later when we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=252.93750000000003) [invoke Cloud ML Engine, we will upload](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=255.40728571428568)[these files to the Cloud Storage bucket as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=257.25) [usual. So we set up our iris.csv and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=260.17799999999994) [iris\_eval.csv. These are two small files](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=264.389125) [in our data directory. Just for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=267.0025555555555)[completeness, let's also go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=269.205) [create out setup.py. Once again, this step](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=271.26175) [is optional. We do not need to perform](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=274.90166666666676) [this if we are using gcloud to package up our code.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=4&mode=live&start=276.64639999999986)

[TensorFlow Model: Train Locally and on the Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live)

[Now that we have our TensorFlow code](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=1.689) [ready, let's go ahead and embark on a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=4.666624999999999) [series of training processes. The first is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=7.021666666666668) [going to be just using TensorFlow on our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=9.579428571428572) [Cloud Shell machine. This will not even](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=13.084666666666669) [involve the use of Cloud ML Engine at all.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=15.355555555555554) [As always, let's define a set of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=19.646) [environment variables with our training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=21.299) [file and our eval file. Then let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=23.248) [ahead and save the date and time. As](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=26.8707) [always, we'll use these to construct our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=29.758125) [job directory path. Let's make sure we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=31.6) [in the right directory, and then let's go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=34.367) [ahead and invoke Python. Notice here that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=36.376) [we are not making use of Cloud ML Engine.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=39.64866666666667) [This is just good old Python. Our job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=42.43433333333334) [directory is local, and so the checkpoints](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=45.726499999999994) [are going to be saved locally, that is on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=47.830999999999996) [the VM hosting our Cloud Shell session.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=51.01279999999999) [Once this runs through successfully, let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=53.817) [switch back to the code editor window and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=55.622375000000005) [find our job directory. And within that,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=58.3866) [we can find all of the model checkpoints](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=61.10833333333334) [that have been saved there. All of these](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=63.24414285714286) [have been saved here by TensorFlow. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=66.31166666666665) [take a look at the directories that were](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=68.591) [created. There's an export directory,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=70.46700000000001) [within which there is a subdirectory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=73.197) [called exporter, inside which we can see](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=74.64557142857143)[the variables of our model, as well as the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=77.39585714285712) [saved model itself. Two points worth](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=80.504) [commenting on here, notice that the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=83.255) [variables and the saved model are in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=85.012) [separate files, and the saved model has a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=87.024) [.pb extension. That means it's a protocol](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=89.929) [buffer file. Let's explore that variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=92.865) [subfolder, and we can see that there are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=96.39) [multiple files in there. These are the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=98.274) [model parameters. All of these represent](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=102.025) [the final model outputs, in addition to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=104.59271428571428) [which are all of the various checkpoints](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=107.04749999999999) [saved in the same directory. Now that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=109.69637499999999) [are satisfied that this code runs fine,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=112.52175000000001) [let's go ahead and upload and train on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=114.15625) [cloud. The first step is to upload our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=117.03366666666668) [training and evaluation data on the cloud,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=119.59842857142857) [so we navigate to our trusty Cloud Storage](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=121.97850000000001) [bucket and upload the eval and training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=124.79285714285713) [files. These are the same two files which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=127.15025) [we just used for the local training. So we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=129.75833333333335) [are going to redefine our training and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=132.92) [eval file names to point to the BUCKET\_ID.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=135.05344444444444) [This is in preparation of invoking Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=139.017) [ML Engine. We construct the job directory](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=142.23622222222218) [as usual by combining the date, along with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=144.73975000000007) [our model name. Make sure that the job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=147.78374999999994) [directory does not already exist for some](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=150.15437499999996) [reason. That would cause problems in the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=152.38379999999998) [training. And then two new flags, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=154.6742857142857) [define the number of training steps to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=156.49999999999997) [1000 and the number of evaluation steps to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=158.70699999999994) [be 100. Now we are ready to invoke Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=161.89899999999994) [ML Engine. We are still going to do so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=165.3946666666667) [locally. Note that this is the first time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=168.464875) [we are using the gcloud command-line tool](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=170.0737777777778) [with TensorFlow, and there is one little](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=172.138) [wrinkle that we should be aware of. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=174.548) [convention while invoking TensorFlow is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=177.47316666666666) [that we first need to specify all of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=179.689) [command-line arguments for ML Engine, then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=182.147) [specify an empty dir argument, which is a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=185.98000000000002) [single backslash, and then specify all of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=189.56780000000003) [the command-line arguments for the user](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=191.82085714285708) [application. This is quite an annoying](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=194.51760000000002) [little restriction, but this is what the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=196.3555) [docs tell us to do, so let's go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=198.24774999999994) [do that. Let's go ahead and kick this off.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=201.69359999999992) [Along the way, some interesting points.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=204.179) [Notice first how Cloud ML Engine is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=206.287) [setting the values in the TF\_CONFIG](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=208.75975000000003) [environment variable. Second, remember](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=211.45585714285718) [that even though this is Cloud ML Engine,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=213.05785714285716) [we were running it in local mode, and so](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=215.35750000000002)[the checkpoints and the saved model are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=217.64149999999998) [still running in our Cloud Shell home](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=219.99325000000005) [directory and not being saved to a Google](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=222.6945) [Cloud Storage bucket. And that's why we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=226.1954) [see another job directory has appeared](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=228.8205) [here in our Cloud Shell code editor. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=230.94544444444443) [change that. Let's create yet another](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=234.991)[JOB\_NAME with the latest date and time and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=236.72471428571427) [then create a new JOB\_DIR, this time](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=239.20328571428573) [pointing to a Google Cloud Storage bucket](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=241.85057142857144) [location. Next, let's go ahead and invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=244.8013333333333) [the gcloud ML Engine jobs submit training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=247.51524999999995) [command. This time, we are actually](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=250.9991666666667) [running training on the cloud. Notice that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=252.9962857142857) [we have not specified a value for the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=255.30674999999997) [scale tier, which means that by default,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=258.105) [we will just use the BASIC scale tier with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=259.7588750000001) [a single node. Because we are running this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=262.9113333333332) [on the cloud, we also want to be streaming](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=265.22012499999994) [our logs so that we can examine them from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=267.42566666666664) [the jobs viewer. So that's another thing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=270.66766666666655) [that we don't do while running it locally.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=272.6920000000001) [And then a bit of annoying fine print. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=275.483) [job-dir for distributed training needs to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=278.3698333333333) [be specified before that empty flag. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=281.03328571428574) [the empty flag is of course the single](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=284.3677777777778) [backslash, which is used as a separator](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=286.118) [between the arguments for Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=288.423) [and the arguments for the application. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=291.01) [reason for this, according to the docs, is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=293.704625) [that ML Engine is going to validate the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=295.6017777777778) [job-dir before it gets passed in to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=297.696) [application. Let's stick with best](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=301.22311111111105) [practice and run our training in the same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=302.866) [region as our Cloud Storage bucket and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=305.0615) [remember again that we should avoid using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=307.07800000000003) [multi-regional buckets with Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=310.3140000000001) [Engine because they can give rise to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=312.54187499999995) [issues of eventual consistency. We go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=315.33133333333336)[ahead and submit this job successfully. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=318.31624999999997) [can now switch to the UI and ensure that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=320.50890000000004) [it is indeed being processed. We can view](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=322.092) [the logs at any point in time, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=325.4563999999999) [eventually, this process completes. So the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=327.5448) [TensorFlow job took 7 minutes and 52 seconds, and it consumed 0.09 ML units.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=5&mode=live&start=330.7504444444444)

[TensorFlow Model: Simulate Distributed Training, Run Distributed Training with Multiple Workers](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live)

[In the previous demo, we did run our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=2.49) [training job on the cloud. However, that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=4.583) [was not really fully distributed because](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=7.1946666666666665) [we had made use of the BASIC scale tier](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=9.116249999999999) [where there was just one single instance.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=11.986499999999998) [In this demo, we will take things further.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=14.422600000000001) [We will first simulate distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=17.255) [training on a local machine and then](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=19.205000000000002) [actually go ahead and do the real thing.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=21.275000000000006) [That is, we will run distributed training](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=23.529) [with multiple workers. This will seem to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=25.952) [be really easy, and it is, because of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=29.079999999999995) [Cloud ML Engine. So let's get started. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=31.52875000000001) [define a job name as usual, and then we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=34.035900000000005) [call gcloud ml-engine local train, so even](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=36.564) [though we have local in the invocation, we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=40.701249999999995) [have a distributed flag down below. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=43.541000000000004) [is a special mode called locally](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=47.221) [distributed mode in which cluster](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=49.087) [configurations, etc., are set up exactly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=51.344) [as if a cluster was in place, even though](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=54.02600000000002) [it's all actually happening on this same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=56.911222222222236) [one machine. So if we pay careful](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=60.362249999999996) [attention to the output, we can see that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=62.559124999999995) [the TF\_CONFIG environment variable is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=64.93085714285715) [being set as usual, and then something a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=66.8532857142857) [little different. Notice how there are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=69.78433333333334) [explicit entries for the parameter servers](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=71.73871428571428) [and the workers. These messages were not](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=74.42550000000001) [displayed in the plain old local mode. By](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=77.29242857142857) [default, in locally distributed mode,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=80.7825) [there are two parameter servers and two](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=82.435) [workers. Again, these are all simulated.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=84.83862499999996) [All of this is still happening right here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=87.704) [on this Cloud Shell VM. This runs through](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=89.584) [successfully, and now we are in a position](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=93.38357142857141) [to really run distributed training with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=95.75950000000002) [the scale tier STANDARD\_1. So we define](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=98.248) [this environment variable and invoke](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=101.41828571428573) [gcloud ml-engine jobs submit training. Now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=103.2085) [that we are using the jobs API, we go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=107.35662500000001) [ahead and stream the logs out, and we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=109.62899999999999) [specify our scale tier of STANDARD-1. This](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=112.64525) [is going to default run 1 master, 4](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=115.80224999999999) [workers, and 3 parameter servers. So by](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=118.89183333333332) [this scale tier, we are significantly](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=122.9435) [going to increase the amount of compute](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=125.006) [capacity that we are throwing at this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=126.877)[training problem. In a real enterprise](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=128.694) [setting, this is probably where you will](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=131.621) [experience the magic of distributed cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=134.307) [training for the first time. There are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=136.475) [numerous instances of models which are run](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=139.207) [in local mode and report accuracies on the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=141.37112500000003) [order of 50 or 60%. And then they are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=144.22575)[subjected to distributed training with](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=147.7174285714286) [enough compute thrown in, and that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=149.465) [accuracy rises to maybe 85 or 90%. If we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=151.8201111111111) [continue to pay close attention to the log](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=155.21275) [messages, we can see that there are now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=157.39955555555557) [multiple worker replicas and multiple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=159.51116666666667) [parameter server replicas. And there's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=162.342) [nothing particularly remarkable about](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=165.79440000000002) [anything else in this demo. The job logs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=167.37033333333332) [can also be viewed in the ML Engine UI.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=170.50279999999998) [Indeed, the whole magic of this demo is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=173.13575000000003) [just how much additional complexity we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=176.40414285714283) [have going on under the hood, and we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=178.9183) [completely abstracted away from it all.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=181.18) [Good luck trying to implement any of this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=184.222) [with a native Kubernetes distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=186.45) [cluster without the benefit of a managed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=188.302) [service like ML Engine in between. If we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=191.043) [switch back to the UI, we can see that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=194.43779999999998) [this job took 5 minutes and 59 seconds to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=196.6042222222222) [run, so it ran more quickly than our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=199.20455555555554) [previous invocation, which took 7-odd](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=201.369) [minutes. But this time, we consumed 0.99](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=203.87400000000002) [ML units. That's an order of magnitude](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=207.31650000000002) [more than we have consumed in any of the other demos in this course so far.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=6&mode=live&start=210.07199999999995)

[TensorFlow Model: Hyperparameter Tuning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live)

[In this demo, we'll see how hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=2.376) [tuning can be performed using the Cloud ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=4.055) [Engine. This will effectively cause a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=7.186555555555557) [number of independent training jobs to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=9.163714285714285) [executed. And then the best one of those](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=12.254666666666667) [will be chosen on the basis of our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=14.024666666666663) [evaluation metric, which here is accuracy.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=16.31171428571429) [This time, we'll need to create a config](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=19.121) [file, so let's go ahead and create a new](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=21.499) [file and call this hptuning\_config.yaml.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=24.076) [And we'll use this to configure our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=29.341) [hyperparameter tuning job. So we've got to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=30.98) [specify right up front our goal, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=34.67155555555554) [here is to maximize our hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=36.724) [metric tag, which is the accuracy.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=39.64071428571429) [Hyperparameter tuning can get really](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=42.921) [expensive, and so in order to keep control](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=44.673) [on the costs, we apply a maxTrials field,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=47.6) [which is set to 4, and we also specify](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=51.042) [that there should be at most 2 trials](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=53.20740000000001) [going on in parallel at the same time. Now](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=55.7261111111111) [the hyperparameters that we want to tune](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=59.66271428571428) [are the training batch size and the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=61.5255) [learning rate. The training batch size is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=64.455) [of type INTEGER. We would like to try a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=67.168) [minimum value of 8 and a maximum value of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=69.994) [64, and we would like the scaling to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=72.994) [carried out in a UNIT\_LINEAR\_SCALE.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=75.869) [Clearly, when we've specified the max](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=78.912) [number of trials is 4, it's incumbent on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=80.675) [the Cloud ML Engine to intelligently pick](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=83.596) [the values of these hyperparameters. It's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=86.323) [not going to be able to try every single](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=88.97044444444445) [value between 8 and 64 in units of 1. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=90.7383) [this is where the different hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=95.02533333333332) [tuning algorithms which we just discussed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=96.8918) [can come into play. Likewise, we also](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=99.33833333333332) [specify how we would want the learning](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=101.91899999999998) [rate to vary, this is of type DOUBLE, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=104.20366666666668) [we would like it to vary between 0.01 and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=106.33200000000001) [0.1 in a logarithmic scale. Let's go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=109.94212499999999) [and define an environment variable](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=114.3217142857143) [pointing to this config file, and then the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=115.71883333333332) [rest is pretty familiar. We set up a job](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=119.05757142857144) [name and then invoke gcloud ml-engine jobs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=121.65000000000003) [submit. The only noteworthy bit is that we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=124.50966666666666) [specify a config parameter which points to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=127.36137500000001) [our YAML file. We kick this off as usual](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=130.572) [and then switch over to the UI to the Jobs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=133.63633333333337) [pane where we now see that this indeed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=136.98349999999994) [does show up, but it's of a different](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=139.35087499999995) [type. Notice that the type of this job is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=141.67449999999997) [hyperparameter tuning, in contrast to all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=143.644) [of the training jobs which we were running](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=146.196) [so far. We keep monitoring the logs, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=148.133) [then at some point, the job finishes.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=151.655125) [Notice that this job takes quite a while](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=153.805) [to run. It takes 17 minutes and 46](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=156.22819999999996) [seconds, and what's more, it consumed 3.95](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=159.06514285714292) [ML units, so that's about 4 times as much](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=162.57724999999994) [compute capacity as the distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=165.29437500000006) [training job and about 40 times as much](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=167.3736666666667) [compute capacity as any other job that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=170.26800000000006) [we've run so far in this course. Clearly,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=172.08149999999998) [the moment we get into hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=174.35214285714287) [tuning territory, the amount of compute](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=177.13150000000002) [required just increases dramatically. As](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=179.87266666666667) [an aside, it's no coincidence that this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=183.19)[job consumed four times as much compute as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=185.031) [the previous one. To see why that's the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=187.49) [case, just scroll up. We can see the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=189.997) [training output has a completedTrialCount](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=192.33900000000003) [of 4. That's because we had specified](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=193.95066666666668) [maxTrials equal to 4. So effectively, ML](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=197.29433333333336) [Engine ran four independent training jobs,](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=200.4027499999999) [that is four trials, and it went ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=203.918) [chose the one with the best accuracy. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=206.229) [so this is why the amount of compute](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=209.76555555555555) [required was almost exactly four times](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=212.082) [what we had just used in the distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=214.6671428571429) [training job. The trial which gave the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=217.133) [best accuracy is going to be selected.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=220.06) [Let's now go ahead and use that best model for some prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=7&mode=live&start=222.823)

[TensorFlow Model: Online and Batch Predictions](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live)

[Let's round off this course with one final](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=1.652) [demo in which we will see how TensorFlow's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=3.873) [saved model can be used for both online](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=6.633) [and batch prediction. Let's go ahead and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=8.866777777777779) [create a model using gcloud ml-engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=11.734500000000004) [models create. This is our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=14.4055) [iris\_spikey\_model. Now the next step is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=17.907) [worth paying attention to. We need to make](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=19.80033333333333) [sure that we use the best trial. That is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=21.972090909090902) [the saved model from our hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=24.546571428571426) [tuning. To do this, first let's just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=26.8035) [examine the contents of the JOB\_DIR](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=29.377250000000007) [directory. So we carry out this gsutil ls](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=32.0088) [command, and we run this on the JOB\_DIR](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=35.26074999999998) [directory, /export/exporter. This is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=37.791125) [usual convention. This is where ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=41.36833333333333) [is going to have written the outputs. And](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=44.045666666666676) [indeed, this directory structure looks a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=46.990833333333335) [lot like the one on our local machine. We](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=49.02033333333333) [can see that there is one saved\_model.pb.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=51.47542857142857) [That is a variables subfolder. And all of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=53.6035) [this is inside a directory with a numeric](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=58.903999999999996) [system-generated name. So we need to go](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=62.415499999999994) [ahead and capture this directory and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=64.98488888888888) [assign it to a variable called](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=66.53628571428571)[MODEL\_BINARIES. Now that we know where the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=69.339) [saved model is located, we can go ahead](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=71.7176666666667) [and create our first version. So we use](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=73.92662499999997) [gcloud ml-engine versions create v1, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=77.14714285714282) [the most important parameter here is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=80.33514285714288) [origin, which is MODEL\_BINARIES. From here](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=83.82899999999998) [on in, the rest of the process is pretty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=86.99579999999999) [straightforward. Just do keep in mind that](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=89.54599999999999) [TensorFlow models can be used with both](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=92.31624999999997) [batch and online prediction. Let's start](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=94.8977777777778) [with an online prediction, since we done](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=97.13566666666667) [that before. Let's create a simple](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=99.4448) [input.csv in which we just have the values](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=101.89399999999999) [for our x variables separated by commas.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=105.14549999999997) [Let's now go ahead and invoke gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=109.182) [ml-engine predict, along with this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=111.143) [input.csv. Notice the name of the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=115.62966666666668) [parameter we pass in is text\_instances.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=117.69699999999997) [This is a way of telling Cloud ML Engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=119.38699999999999) [that our input data is comma-separated](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=122.58549999999998) [text. Remember that this is the iris](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=126.08433333333335) [classification problem, and so we are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=128.632) [going to get back a vector of three](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=130.69599999999997) [probabilities, which correspond to the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=132.701) [three classes. Here the prediction is for](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=135.261) [the iris Versicolor class, which is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=138.316) [represented by the numeric label 1. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=141.428) [online prediction was pretty simple. Let's](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=144.77183333333332) [now go ahead and upload that same](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=147.20025)[input.csv file with just one prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=149.015) [instance to Google Cloud Storage. Note](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=152.173) [that there is nothing preventing us from](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=155.19957142857143) [having as many instances in here as we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=156.8425) [require. This file will then be passed in](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=159.49450000000004) [for our batch prediction. Let's navigate](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=161.64599999999993) [to the UI and refresh our bucket and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=163.99399999999997) [ensure that our input.csv file does indeed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=166.80999999999995) [appear there. So let's initialize our](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=169.92966666666663) [JOB\_NAME and go ahead and invoke gcloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=172.32800000000006) [ml-engine jobs submit, this time with the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=175.62466666666668)[prediction keyword. Please pay careful](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=179.34314285714285) [attention to this. We are making use of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=181.3895) [the jobs service, and we are submitting a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=183.95699999999997) [prediction job, so this is pretty](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=186.6088571428571) [different syntax from every online](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=188.2354285714286) [prediction which we have performed so far](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=189.59933333333333) [where we simply used gcloud ml-engine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=191.97828571428568) [predict. And because this is a batch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=194.9706666666667) [prediction job, we need to specify the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=197.65250000000003) [data format, which is text, the input](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=199.7034285714286) [path, as well as the output path. So all](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=201.71366666666663) [of the predictions are also going to be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=204.40824999999995) [written back out into Google Cloud](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=206.45688888888887) [Storage. And because this is a job, its](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=208.73533333333336) [state is initially queued. We can follow](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=210.7993333333334)[up on it in the ML Engine Jobs UI, and we](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=212.80762500000003) [can see that this is yet another type of](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=216.36900000000009) [job. This is a prediction job. Contrast](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=219.18612500000003) [this with the hyperparameter tuning and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=221.478) [training jobs run before this. We monitor](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=223.76314285714284) [the logs, and eventually, our job runs](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=227.036) [through successfully. We can now switch](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=230.2152) [back to Cloud Storage and examine the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=232.37155555555557) [output path where all of the results have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=234.236) [been written. In this case, we have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=237.216) [uploaded the same input.csv which we had](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=239.2885) [created with one prediction instance, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=241.65642857142862) [so when we click on it, we can see just](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=244.08257142857144) [the one vector of probabilities. We have](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=247.91636363636368)[successfully demonstrated both online and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=250.61020000000002) [batch prediction with our published TensorFlow model.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=8&mode=live&start=252.582)

[Summary and Further Study](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live)

[This gets us to the end of this module and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=1.748) [the end of this course. We covered a lot](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=3.632) [of ground here. We studied TensorFlow on](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=6.097142857142858) [the Google Cloud ML Engine. These are](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=8.040571428571429) [fairly tightly coupled, and the mechanism](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=10.690000000000001) [for a lot of that coupling is the](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=12.6595) [TF\_CONFIG environment variable. We went](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=16.04657142857143)[through an extensive set of demos, which](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=18.3075) [demonstrated both local and distributed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=20.3794) [training. We also saw how hyperparameter](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=22.7225) [tuning is implemented and how it leads to](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=25.003857142857136) [a sharp increase in the amount of consumed](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=27.59775000000001) [ML units. We rounded off the module with a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=31.36550000000001) [look at both batch and online prediction.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=33.82455555555556) [Remember that online prediction is](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=36.923) [supported for XGBoost and scikit-learn, as](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=38.567) [well as TensorFlow, but batch prediction](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=41.65014285714286) [is only possible with TensorFlow. If](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=44.057) [you're interested in other machine](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=47.59466666666667) [learning options on the cloud, it might be](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=48.908) [worth your while to check out these two](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=51.315)[excellent courses on Pluralsight. The](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=53.3) [first is Designing and Implementing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=55.538) [Solutions Using Google Cloud AutoML, and](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=57.023) [the second is Deep Learning Using](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=60.45242857142858) [TensorFlow and Apache MXNet on AWS](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=62.243) [SageMaker. Google Cloud AutoML is a](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=66.5587142857143) [solution on the GCP which offers](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=68.87550000000003) [pre-trained models. SageMaker is a service](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=71.7036) [on AWS, which, of course, is a competing](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=74.62257142857142) [cloud platform. I hope you enjoyed this](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=77.88271428571429) [course. Thank you for your time and thank you for watching.](https://app.pluralsight.com/player?course=google-ml-engine-architecting-production-ready-models&author=vitthal-srinivasan&name=05afa667-5fc7-4348-9f83-fba177a285a2&clip=9&mode=live&start=79.94499999999998)