

# Production Summary

How would you implement this in a cloud (AWS or GCP) architecture?

Depending on the finance team's use case there are several options in how to implement this project on a cloud architecture. One simple way would be to deploy the project on a scalable server instance. From the instance, the project can be scheduled to retrain and can write the serialized model or predictions directly to the database, or embedded into a web application.

Which IaaS/PaaS components would you use? Why?

I would likely use a PaaS such as AWS Elastic Beanstalk which could take care of complex design decisions such as load balancing and autoscaling. From there, it would be easy to interact with other platform offerings such as Amazon S3 and Redshift to make a seamless application. These components allow for more separation and focus on the development of the algorithm itself.

How would you monitor this service in production to accommodate for reliability?

In order to make sure that the production model does not degrade overtime, it is important to have a feedback loop to track actual performance and behavior changes. To do so, we can store snapshots of information such as training data distributions, model feature importance values, and testing accuracies within the database and feed this to a web app which allows for easy tracking and visualizing of trends. To further ensure reliability, alerts can be set up to notify the team of any dramatic shifts in the captured metrics so that the team can further investigate causes and mitigations.