Springboard

What Is A Good ML Project?

A project walkthrough is an important part of the interview. You need ONE beautiful project, that showed iteration, with trial and error.

1. Consider What Skills You Want To Demonstrate and Develop

In unit 2, you have learned about different roles in the MLE job family. Capstone is a great chance to show your strength. Shape your capstone in a way that helps you shine.

2. Pick An Area of Focus

Although all students need to go through the milestones, we Introduce 4 areas of focus: data, models, engineering, and product.

A. Data-Focused Engineer

In this area of focus you will spend 50-60% of your time working with data, 20-30% of your time working with models, and 20-30% on architecture and MLOps. You will demonstrate a compelling solution to the problem you have selected where data is the star of the show. If you want to be a Data-Focused Engineer, you may:

- Spend more time engineering your data pipeline: acquisition, cleaning, processing, etc.
- Build a fully automated pipeline for data ingestion
- Drive the increased success and accuracy of your model through better data processing. High data quality is the best way to increase the accuracy of most models. The alternative to this is to increase accuracy by focusing more modeling and tuning, to circumvent flaws in the data. If you would prefer to focus on this aspect of machine learning, one of the other areas of focus might be preferable.

- Build lots of tests to ensure data sanity. This might involve implementing anomaly detection feeding the data into the ML model.
- Use a real-time/streaming solution vs the tradition batch processing approach, and potentially build an online training approach as a result
- Develop and display your ability to collect and handle a very large amount of data

B. Modeling-Focused Engineer

In this area of focus you will spend 25% of your time working with data, 50% of your time working with models, and 25% on architecture and MLOps. You will demonstrate a compelling solution to the problem you have selected where the model is is the star of the show. If you want to be a Modeling-Focused Engineer, you may:

- Spend a considerable amount of time on feature engineering.
- Try many different models and carefully review and compare results to balance accuracy, generalization, overfitting, success metrics, and more.
- Perform hyperparameter tuning for every model.
- Go further than using a simple “out of the box” solution. You may achieve that by blending multiple models together, stacking, or any technique that will use a more elaborate approach. That might include testing a new deep neural network (DNN) structure for the problem at hand.
- Use some of the latest research and tools, and implement or adapt them to your problem.

C. Architecture-Focused Engineer

In this area of focus you will spend 30-40% of your time working with data, 20-30% of your time working with models, and 40% on architecture and MLOps. You will demonstrate a compelling solution to the problem you have selected where the engineering and software architecture is the star of the show. If you want to be a Architecture-Focused Engineer:

- Your primary goal will be a well-engineered solution from the data perspective, from model training to deployment and especially the post-deployment phase. This may require monitoring, retraining, and iterating on your solution based on performance.
- You will likely not use an out of the box cloud solution, but instead focus on deploying your own from scratch, writing your own API, building a Docker container, and deploying it on Kubernetes or another cloud solution.
- You will have a detailed workflow diagram for each logical area - data, modeling, and in production - as well as operational diagrams.
- You might use Airflow or other orchestration solution to trigger retraining or redeployment of your model.

- You will integrate many open source platforms/solutions together to create an enterprise-like-solution.
- Your deliverable will be operating live, and you will have developer-oriented documentation available for others to use or integrate your work.
- You will know exactly how long it takes to train your model, and how many inferences/s you are able to achieve, and demonstrate you can scale.

D. Product-driven Engineer

In this area of focus you will spend 30% of your time working with data, 20% of your time working with models, and 50% on architecture and packaging your model as a product. You will demonstrate a compelling solution to the problem you have selected where the product or app is the star of the show. The utility of your model is the major focus of your project, and the engineering work might be written up in a white paper on how you used your data and tools. If you want to be a ML-Driven Product Engineer, you'll:

- Show the clear utility of your solution to end-users
- Demonstrate how the product works to anybody
- Have a website that shows off the product, why you built it, how you built it, and which problem it solves
- Still spend a generous amount of time on data, modeling, and deployment of your project.
- Spend extra time in testing the product, even potentially building a feedback loop into your product in order to improve the original trained model
- Show both the synthetic and real life accuracy of your model
- Recognize what kind of biases your system might have, and how to address it

You do not have to make a decision now and can change as you know more about what datasets and models you would want to work with. This focus will help you to manage your time and direct you to allocate energy wisely throughout the course.

3. Is Your Project Too Simple Or Complex?

This goal of this course is to give you the skills to design and create an ML/DL application, and the skills to deploy that application to production using the latest engineering tools and techniques.

ML/DL contains a smorgasbord of techniques. There is no way to apply everything ML technique to the problem you have chosen, nor should you. The more complex the techniques you use, the more attractive your project will be to employers. However, it's important to balance complexity with a practical approach that produces a real, scalable application. Some questions to ask yourself:

- What technique best applies to this problem?
- How easily can this technique be deployed to production?
- Are the results of this technique “good enough” to meet the business requirements of the solution?

In the real world, it **can sometimes be better** to use a simple logistic regression approach that’s accurate enough, highly performant, and easy to deploy as a production application, than a intricate deep learning approach that’s intensive to deploy or maintain. You’ll have to make similar decisions and tradeoffs about your capstone project. When you’re interviewing for your MLE role, you’ll be expected to justify these decisions and tradeoffs.

Work with your mentor to determine a problem and approach that meets the requirements of this course and makes you both feel confident that you’ll be able to do well.

It is worth reading and thinking through these questions before you submit your ideas and plan!