System/DevOps Engineer I Interview Questionnaire

Project Overview

You are given a pre-trained, `sklearn` model that has been trained to predict housing prices in Boston according to several features, such as average rooms in a home and data about highway access, teacher-to-pupil ratios, and so on. You can read more about the data, which was initially taken from Kaggle, on the data source site. This project tests your ability to operationalize a Python flask app—in a provided file, `app.py`—that serves out predictions (inference) about housing prices through API calls. This project could be extended to any pre-trained machine learning model, such as those for image recognition and data labeling.

Project Tasks

[This project currently only works on x86 architecture machine due to Sklearn's incompatibility with ARM architecture machines]

Directory containing the application is attached to the email with this Questionnaire. Your project goal is to operationalize this machine learning microservice using Docker. In this project you will:

1. Initialise the code directory as a Git repository
2. Complete a Dockerfile to containerize this application
3. Deploy your containerized application using Docker and make a prediction
4. Complete the CI/CD pipeline using Jenkinsfile present in the repo. [Note: Just completing the Jenkinsfile is enough, actually publishing image to Dockerhub is not a requirement]
    a. Pipeline should be divided into stages.
    b. Build Image stage is responsible for building the Docker image from Dockerfile completed in step 1.
    c. Publish Image stage is responsible for pushing the Docker image to Dockerhub.
5. Upload the completed repo to your Github account and reply to our email with the link to your Github repo.
    a. We will watch for the quality of your commits.
    b. So make sure to use correct Git workflow.

The final implementation of the project will showcase your abilities to operationalize production microservices.