Report for Milestone3

Containerization:

Automated model updates:

First of all, we pull the data from the database for analysis. After we get the data, we have to form the preprocessed data. Process the data by gender can get the analysed data. Treat the analysed data as meta data, we put it into the weights and biases. After that, we can write a YML file for the GitHub actions. So the basic idea for how we track the version of the model is we track the GitHub Sha of the branches and pull requests. Save the data and model with their Sha inside the weights and biases. So the model we deposited is always with the Sha. From the Sha we can get the information of the model, which branch is it from, which pull request is it related to and when we deployed the model. So at the prediction stage, we can get the prediction model from weights and biases using the Sha and predict.

Monitoring:

Monitoring has already been set up during the previous milestones.

Releases:

Provenance:

Provenance is designed to identify the model version, the used pipeline version, and the used training data.

Individual contributions to the Team: