What is your main role in doing this project?

My main role is the implementation of pre-processing and training of the clustering model.

What were the main difficulties encountered by the whole team? and by you individually?

The main difficulties is the balancing between the performance and the running time. Due to the large scale of the training data, the running time of the data will be large if we use a large number of feature. For my individual difficulties, I am unfamiliar with some of the code to implement the visualization and the training model.

What have you learned from this project?

I learnt the implementation details of the data analyzation and the machine training model. After gaining such experience, I am more familiar with the use of machine learning and how to deal with the dataset. Also, I have gained insights for the system design and presentation of the machine learning problems.