For the neural network, we use the multi-layer perceptron classifier, we tune the hyper parameters to achieve the best performance. Here is the hyperparameters that we used. we use the Adam solver, the network has 6 perceptrons followed by 3 perceptrons. The activation function is Relu. and we set the maximum iteration to be 1000. and alpha to be 10 to the negative 4. The model accuracy could achieve 90.1%. Next, in order to ranking the feature importance, we introduce the permutation feature importance concept from the interpretable machine learning. It measures the incerease in the prediction error of the model after we permuted the features's value which breaks the relationship between the feature and the true outcome. The theory state that a feature is more important if shuffling its value increase the model error and vice versa. So first of all, we compute the original model error, for each feature we randomly shuffle the column and fix the other columns unchanged. We calculate the error for each features and then we do a little bit transformation to make the result more clear. Lastly, we ranking the error. Note that we use Mean square error for calculation. Here is the code which shows how we calculate the error. Note that for each feature error, we loop it for 10 times and summing up and averaging the error to make the result more stable. Here the transformation shows that how much percent each feature error exceed the original error. Remember that a feature is more siginificant if the error exceed the orginal error a lot. Here is the ranking of feature important. The result shows that the most siginificant feature is pagevalue which is the average value that user visited before transaction. It exceed the original error by around 60% which is a lot. The top five siginificant features are pagevalue, the month of November, information duration, month of december and administrative duration.