The goal of this project is to predict the mutations of a human antibody protein by its 4 amino acids. Firstly we use one hot encoding for the sequences to ensures that machine learning does not assume that higher numbers of categories have a higher value but still kind of category dummy variables. Then we divide our train data into train set and validation set of 0.9 ratio and apply the multi-layer Perceptron classifier using stochastic gradient descent to optimize the loss. We first construct 3 hidden layers of 50 neurons in each using early stopping to terminate training when validation score is not improving, it will automatically set aside 10% of training data as validation and terminate training when validation score is not improving by at least 10 consecutive epochs. The activation function we choose here is ReLU. The L2 penalty (regularization term) parameter here is 0.0001 and we use constant learning rate. Finally we will construct a output layer with logistic activation function to do the classification. We use the f1 score on validation set which captures both precision and recall to evaluate our model. After confirming its performance, we then use the model to predict the test set.