Final Project Report

My final project is to predict customers' payment preferences and find out which algorithm has the highest accuracy. After finding the algorithm with the highest prediction accuracy, whether increasing or reducing training elements will have a huge impact on the accuracy

The data of this final project comes from kaggle. This set of data comes from an online shopping website, which contains different data such as customer purchase quantity, single product price, total price, discount, etc. All I need to do is to select the required data, train them and predict the payment method. Here I selected five algorithms to predict the payment method after training the data. The five algorithms are: linear regression, logistic regression, decision tee, Naive Bayes and SVM.To ensure fairness, all algorithm training elements are the same. The result is the screenshot below.

Decision Tree



Linear Regression

LDA:

QDA:

Logistic Regression



Naive Bayes



SVM

Linear



Gaussian



Polynomial



We can see the decision tree has the highest accuracy. Then I will increase and decease the training element of the too find new accuracy of decision tree.

Increase



Decrease



We can see if we decrease the number of elements,accuracy is very low, but if we increase the number of elements, the accuracy increase a little bit.

So, the decision tree is the best algorithm for my data-set