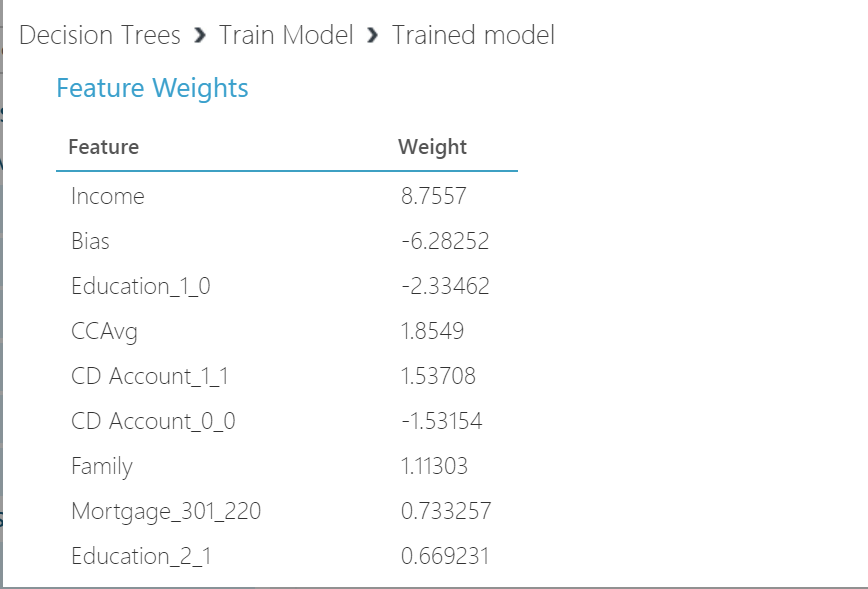
This assignment provides further opportunity to learn/apply logistic regression and CART models.

Using the Universal Bank dataset, your goal is to build models to forecast which customers will accept an offer of a personal loan from the bank. You will use the customer demographic and bank relationship data to build the models.

1. Split the data 70/30 (training/validation). Model loan offer acceptance using logistic regression and CART techniques. Show model output. Answer the following questions:

1. Which factors are important in the prediction of loan acceptance in each model? (10 pts.)

* In Logistic Regression model Income, CCAvg, CDAccount are the most important factors in prediction.



* In Decision Tree, these are the most important factors for prediction- Income, CCAvg and Education.



1. What is the best metric to evaluate model performance (5 pts.)

* The best metric to evaluate the model performance is Sensitivity (True Positive Rate). For Decision Tree model, True Positive Rate = TP/(TP+FN) = 133/133+11= 92.36% For Logistic Regression model, True Positive Rate = TP/(TP+FN) = 97/97+47= 67.36% Sensitivity is also help us to get a False Negative Rate (FNR= 1-TPR). This helps us to calculate the important factor from the decision tree model that means 9.7% of probability in which people will accept the loan, but the bank will lose the business if they don’t consider these potential customers.

1. Which is the better model? (2 pts.)

* As per the visualization, Decision Tree is the better model. After looking at the ROC curves, we can say that the curve of both the model (Logistic Regression & Decision Tree) are above the 45 degree line. Model which has the high True Positive Rate or on the other way a model which is close to 1 (upper top left corner).

[Attach curve]

1. Do the models make sense? (3 pts.)

* Yes, I believe that model make sense because they classified that the customers in the groups/individuals with the high Income, higher education and Credit Card Average are more likely to have a higher standard of living. So, these are the most important factors to predict the target variable (Personal Loan). Secondly, these factors have a positive correlation with the target variable and these customers are more likely to accept the loan and have the ability to pay back.

Business Case ?