Connor Adams

DS 4220-001

Professor Armstrong

9/30/19

Homework 3

Looking at all of the logistic regression models that I created, I found that the best model is the one that uses all of the variables (student, balance, income). The reason why I think this is the best model is because of the accuracy. This model’s accuracy was the best accuracy at 0.9722, with a 95% confidence interval between 0.9667 and 0.9771 and this is why the model is the best.

Looking at the model’s coefficients they show that if a student is marked Yes this will have a negative affect on a student having student loans (-0.711), which probably goes to show that they don’t have loans because they are still a student. Their balance actually has a small positive affect on if a student has student loans or not (0.0058). A student’s income has such a little affect (0.000003) on if a student has loans or not that it wouldn’t be smart to take income into account.