Joshua Granville, Nicholas Lichtsinn

Andrew Marcos, Jermaine Whitehead

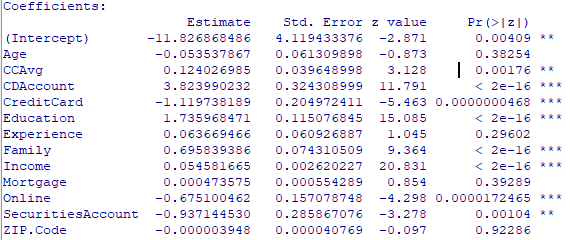
12/7/21

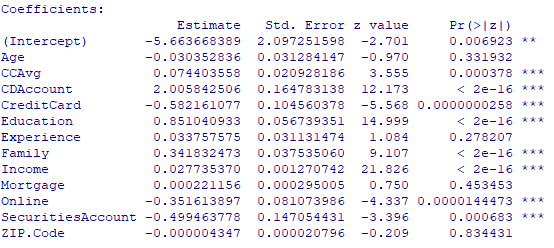
Homework 4

1. Logit and Probit Analysis

For our Logit and Probit analysis of the variables, most of them are significant with the exception of Age, Experience, Mortgage and Zip Code. Of the significant variables CreditCard, Online and Securities Account have a negative coefficient meaning an increase in these variables leads to a decrease in the probability that a customer takes out a loan. The other significant variables (CCAvg, CDAccount, Education, Family, and Income) have positive coefficients, meaning an increase in these variables leads to an increase in the probability of a customer taking out a loan.

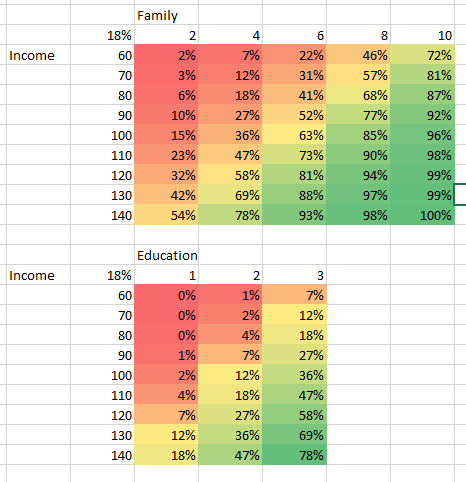
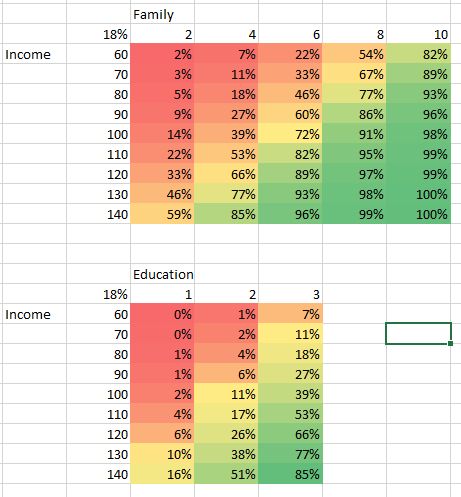
These variables increase the likelihood of a customer taking out a loan based on their coefficients, a lower (or higher) coefficient indicates how much an increase (or decrease) in that variable will affect the probability of a customer taking out a loan.

Logit:

Probit:

Here are the sensitivity analysis charts on Family/Income and Education/Income for our Logit and Probit analysis predictive models. In both you can see when Income, Family size and Education go up, the probability of the customer taking out a loan increases as well which we would expect given their coefficients.

Logit: Probit:



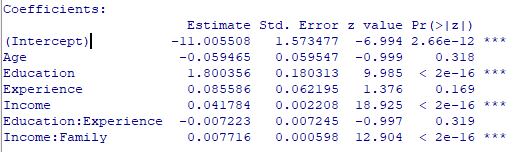
2. Moderating Effects

We added moderating effects to multiple variables, in most cases if the P value for the original variable was higher than 0.05, the same was true of the moderating effect and some where both had P values lower than 0.05 the moderating effect had a high P value. We kept Education\*Experience because education and experience compound with each other in most cases. You would think that someone who has more experience and more education would get a loan because education has a high coefficient already. We also tested Income/Family because each family shares one income, and we wanted to see how family size per income would affect the customer’s loan probability.

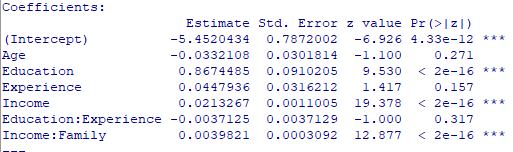
In this model Age and Experience were still not statistically significant along with Education\*Experience. Education, Income and Income/Family were statistically significant.

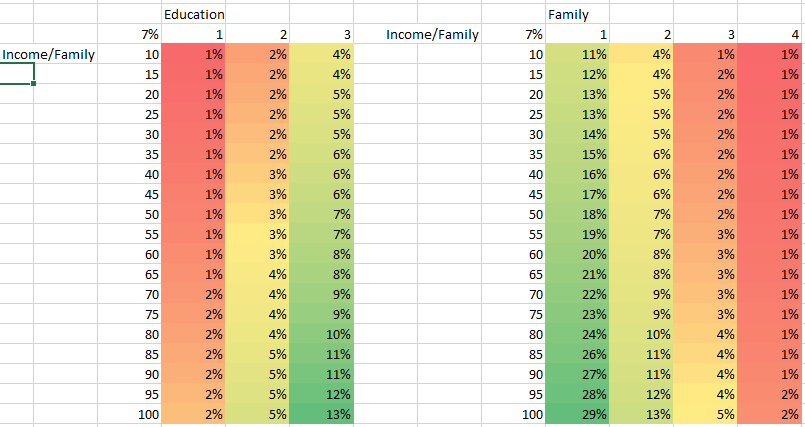
Age and Education\*Experience have a negative coefficient while the rest have positive coefficients meaning the older someone is and the more Education\*Experience someone has, the probability of them taking out a loan is lower with the opposite being true for Income/Family.

Logit:



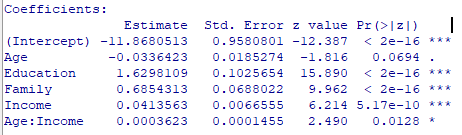
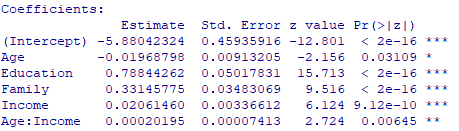
Probit:

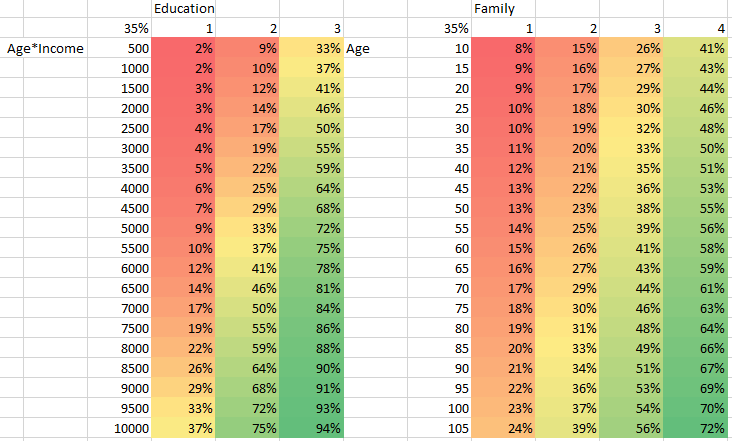


In the sensitivity analysis of our moderating effects we looked at the chart between Education and Income/Family and Family with Income/Family. In the first chart we can see that probability goes up with both Education and Income/Family which makes sense because you have more income for your family and you are more educated both with positive coefficients. Next we see that probability falls with family size and increases again with Income/Family, this is intuitive because larger families are less likely to take out a personal loan.

3. Final Regression

For the final regression model we used the variables that had more variability, some of the binomial variables are more intuitive and had higher coefficients but Age, Education, Family, Income and Age\*Income are more applicable to every person. All of these variables are statistically significant and they all have a positive coefficient except for Age. Education is the largest coefficient and Age\*Income is the smallest.

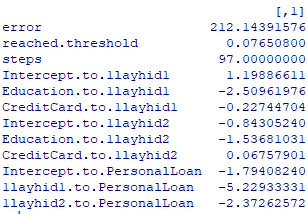
Logit: Probit:

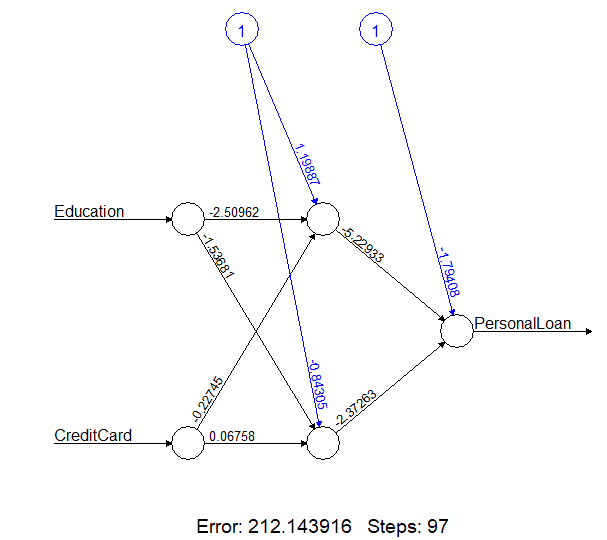
For the sensitivity analysis of the final regression model we looked at Education vs Age\*Income and Family vs Age. The first is intuitive because probability of a loan increases with both Education and Age\*Income. This is interesting because Age has a negative coefficient and Income is positive. The second chart isn’t so intuitive because probability increases with age and we do not think there are many 100 year olds taking out loans so there must be a break even point with these variables, they cannot always go up in probability.

4.Neural Network

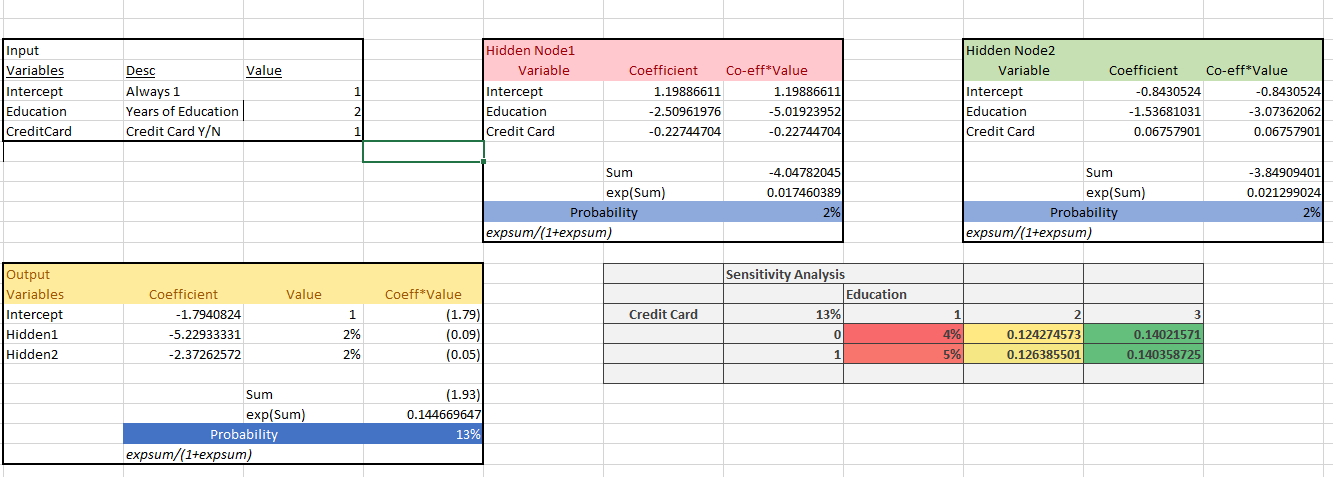
For our Neural Network we chose to use Education and Credit Card because they have the highest positive and negative coefficients respectively so we could see a volatile sensitivity chart. We had an Error of 212.143916 and it took 97 steps to complete the neural net.

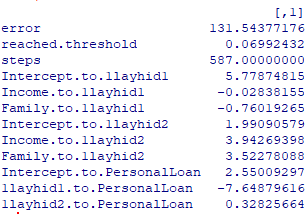
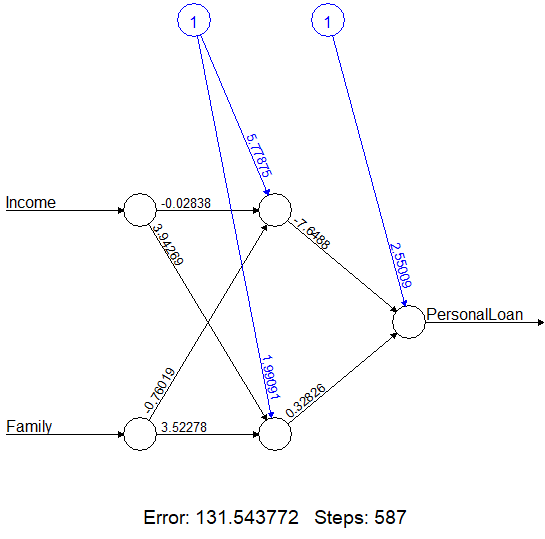
Error:



Plot:

5. Neural Network Prediction Model

We chose Education and Credit Card because they both had high coefficients and we wanted to see a volatile sensitivity chart but with Credit Card being a binomial variable the chart doesn't show us much, so we ran it again with Income and Family.

Error: Plot:

This Neural Net had an Error of about 131 so much lower than our first example and was completed in 587 steps, more than double that of our first example. This sensitivity shows us that probability increases with income and family size just as the previous analysis did.

