Team #3

Joel Bravo

Stephanie Gardner

Kendra Danforth Osburn

Penematafele Seloti

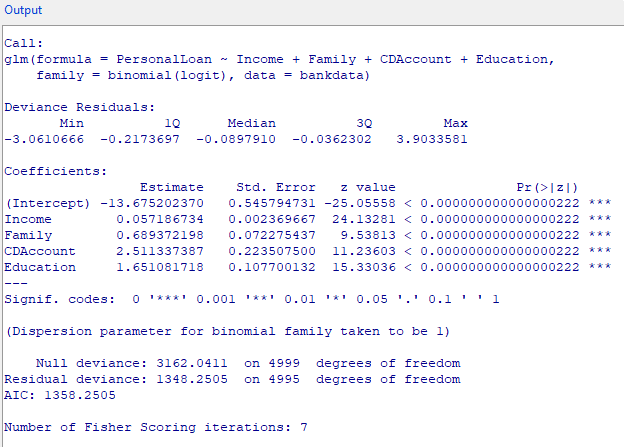
Kenneth Smith

Homework #4

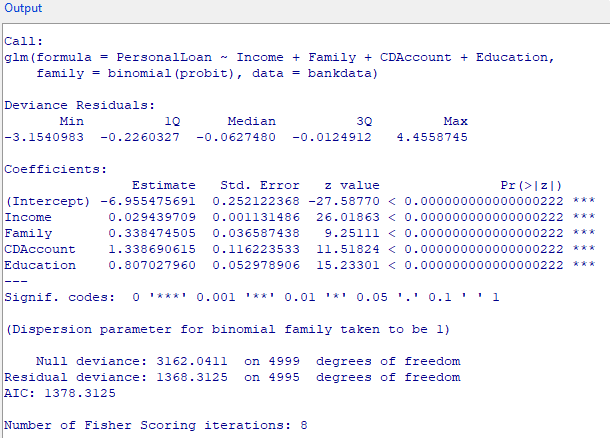
1. Perform a logit and probit analysis of the variables that affect whether a customer takes out a loan. Consider only main effects. Which variables are significant? How do the significant variables influence the likelihood of taking out a loan? Copy screen snapshots of your analysis in R to your report. (20%)

**The four variables considered are: Income, Family, CD Account, and Education. Intuitively, we expect that as these variables increase in value, the probability of customer taking a loan increases. The exception is CD Account since this variable is binary (1=Yes, 0=No). We expect that the probability of a customer taking a loan is higher if they have a CD Account with the bank than if they didn’t have one.**

**Logit Analysis: The analysis confirms our assumptions. All of the coefficients show positive values, which suggests that as the variables increase the probability of a customer taking a loan also increases. Also, all four variables are statistically significant with p-values less than 0.05.**

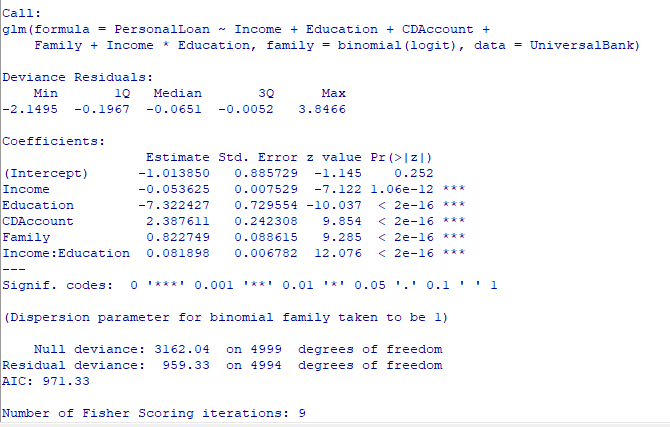


**Probit Analysis: This analysis is similar to the Logit Analysis. All of the coefficients show positive values, which suggests that as the variables increase the probability of a customer taking a loan also increases. Also, all four variables are statistically significant with p-values less than 0.05. There are some differences, for example, the values of the coefficients and intercept are not the same. However, the basic assumption still stand and these differences between the two analysis are to be expected.**

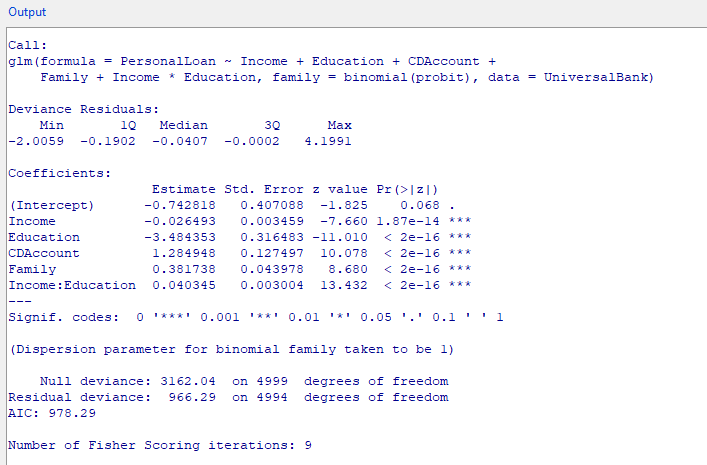


1. Add moderating effects (interactions of variables). Which interactions make sense conceptually? Which interactions are statistically significant? How do you interpret the coefficients on these variables? Copy screen snapshots of your analysis in R to your report. (20%)

**Moderating effects were added to the interaction of Income and Education. This makes sense conceptually since the more education one has most likely they are in a high paying field and the potential for a loan is higher. The analysis in R for the logit moderating effects also shows that all variables are statistically significant with extremely low p-values less than 0.05. Both intercepts are negative suggesting that the lower the income and less education, the probability for a loan decreases.**

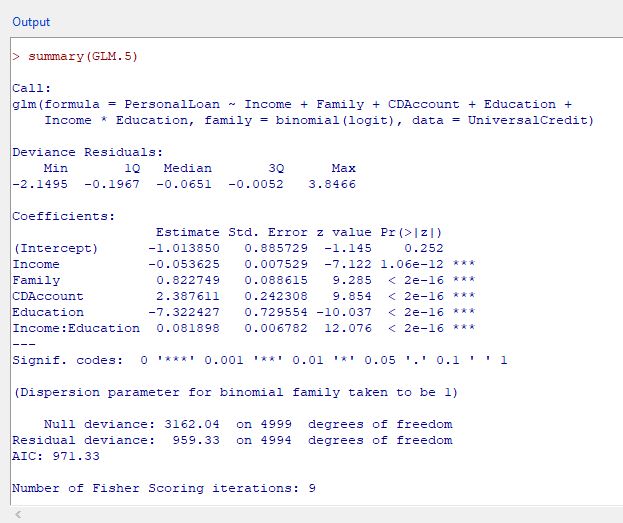


**Moderating the same variables of Income and Education using the probit analysis, we can also see that the coefficients are negative showing a very similar relationship as logit as a higher probability for a loan with increasing Income and Education. Again the p-values are extremely low showing to be statistically significant.**

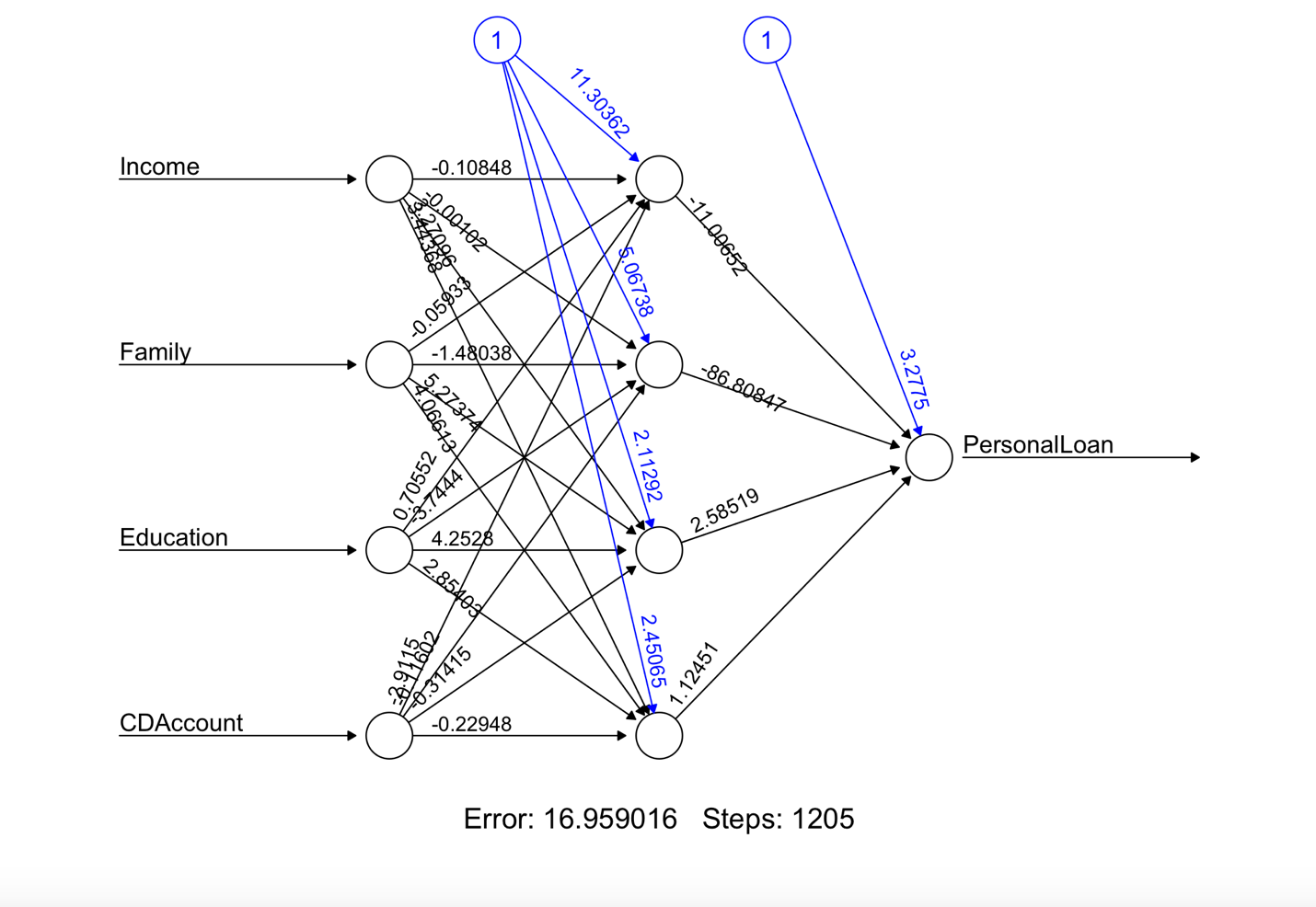


1. Create a final regression model with the variables that you feel are important (both main effects and interaction terms). Create a spreadsheet prediction of the model. Which variables have the greatest influence on the customers’ loan behavior (combined main effects and interaction effects)? Perform a sensitivity analysis as seen earlier in the semester. Copy screen snapshots of your analysis in R to your report. (20%)

**A final logit analysis was run using both the main effects identified (in this case Income, Family, CDAccount, and Education) and the moderating effects (Income:Education). As the size of a family and the presence of a CD account increase, so does the probability that the customer has a personal loan, though CDAccount has a greater influence compared to family size. Interestingly, as Income and Education increase, the probability of a customer having a personal loan decreases, which would be contrary to conventional wisdom. However, when Income is multiplied by Education as a moderating effect, the coefficient is positive, indicating that as it increases so do the odds of having a personal loan. This seeming contradiction is intentional, as the positive coefficient of our moderating effects counteract the combined negative coefficients of Income and Education. By far, Education, CDAccount, and Family have the greatest influence on the customers loan behavior. In addition, all values would be considered significant due to their very low p-values.**



1. Perform a neural network analysis of the variables found to be significant in the logit and probit analysis above. Copy screen snapshots of your final neural network model in R to your report. (20%)



1. Create a prediction model of the neural network. Using the prediction model, perform a sensitivity analysis for the neural network model similar to the logit and probit sensitivity analysis. (20%)

