Is the FICO score the best predictor or indicator of borrowers defaulting on their loan?

Victor Armero Bedoya – Queens College, NY – MA Data Analytics and Applied Social Research
DATA 790 – Spring 2020

Abstract

Are FICO scores the best predictor of credit default? Lenders have been using FICO scores as the principal indicator of creditworthiness and loan defaults for over 50 years. Machine Learning may provide better tools for assessing borrowers' spending habits. Other metrics such as revolving utility and debt-to-income ratio may be better indicators of a borrower defaulting or not. This research uses a Random Forest Classification model to make inferences on the likelihood of defaults. Moreover, the model provides a breakdown of the relevancy of each variable in the dataset. Test data was fed into the model to make predictions. It was found that revolving utility along with annual income had the biggest impact on the prediction of defaults; the FICO scores were kept constant some variables were inflated and deflated to asses sensitivity. Additionally, the 5 most important features for predicting defaults were: the number of days with a credit line, interest rates, income, revolving utility, and debt-to-income. FICO scores were not in the top 5. FICO scores provide some insight on the borrowers' financial past; however, they are not always a good nor bad indicator of the future finances of the borrowers.

Introduction

Lenders need to assess the risk being taken after issuing loans. Traditionally, the borrowers' creditworthiness in the form of a score is an indicator of successful repayment or default. Based on the perceived risk taken by the lender, some loan features are produced. For instance, for riskier loans, the lender may impose higher interest rates to compensate for the risk premium assumed.

The Classic Way of Evaluating Credit Worthiness and Default. Fair, Isaac, and Company (FICO) was founded in 1956 by an engineer and a mathematician. They developed a scoring system, which was sold for the first time in 1958, that would enable lenders to make better decisions by using data intelligently. The FICO score is made up of 5 components. 1) Prepayment history (35%). 2) Amounts owed (30%). 3) Length of credit history (15%). 4) New Credit (10%). 5) Credit mix (10%). [1]The higher the score, the lower the risk and the lower the interest rate. The probability of default is heavily influenced by two metrics, debt-to income ratio and the credit score. Along with these two metrics, lenders also use some other borrowers' characteristics such as, employment length, interest rate, income, and the age of credit history among many others. To assess the likelihood of default, lenders take all these metrics and pour them in a Regression model pot. Regression models are statistical tools that aim to quantify the effect of a set of variables (metrics) on a given target variable, in this case, default or non-default.

New Technologies and Machine Learning. One concern about FICO scores is that they are backward looking numbers. The scores are good at describing the borrowers' financial past; however, they might not be a positive or negative indicator of the borrower's future ability to pay their mortgage of payoff their loans. Machine Learning may help determine the borrowers' habits, which may be a better indicator of default or not. Some metrics that may provide insights on the borrowers spending patterns are revolving utility and revolving balance. Deloitte, which provides financial advisory and risk advisory services, researchers have found that "ML algorithms outperform traditional model in terms of predictive power for various applications, such as predicting defaults."[2]

For this research, I've used publicly available data from LendingClub.com. LendingClub.com operates with a peer to peer loan business model. The company connects investors with borrowers in an online marketplace.[3] Basically, investors buy notes, which correspond to a fraction of a loan, with the expectation of earning 5% return, according to their website.[4] I aim to find how much influence has the FICO score over a default/non-default outcome. A Random Forest Classification model will be used for this. In addition, I will run such model multiple times with modified test data sets. I will inflate and deflate interest rates, debt-to-income, installment amounts, revolving utility, and revolving balance. To quantify the performance of the runs, I will calculate the Expected Loss, given the predicted defaults, at the end of each run. The dataset being used does not provide a loan dollar amount variable; therefore, I will use a hypothetical $10,000 average loan amount. The

---

[1] FICO SCORE EDUCATION, viewed May 21, 2020, <https://ficoscore.com/education/#VitalPart>
[2] Artificial Intelligence for Credit Risk Management 2020, Deloitte, viewed May 21, 2020, <https://www2.deloitte.com/cn/en/pages/risk/articles/artificial-intelligence-for-credit-risk-management.html>
[3] How it Works 2020, Viewed May 21, 2020, <https://www.lendingclub.com/investing/peer-to-peer>
[4] Sample: How the Math Might Work 2020, LendingClub.com, Viewed May 21, 2020, <https://www.lendingclub.com/investing/peer-to-peer>

variables and methodology will be described below. Finally, I will use a Natural Language Process model to analyze the current market sentiment on credit. This analysis will be driven by new articles scraped from the web. A list of those articles will be provided in the references section at the end of the paper.

Tools for analysis. I will use Python for performing all the data analyses. The libraries used are Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, and NLTK among others.

Data

Source. The dataset was obtained directly from LendingClub.com. It contains loan information from 2007-2010. It also contains 9578 observation and 14 features.

Variable	**Description**
credit.policy	1 if the customer meets the credit underwriting criteria of LendingClub.com, and 0 otherwise.
purpose	The purpose of the loan (takes values “credit_card”, “debt_consolidation”, “educational”, “major_purchase”, “small_business”, and “all_other
int.rate	The interest rate of the loan, as a proportion (a rate of 11% would be stored as 0.11). Borrowers judged by LendingClub.com to be more risky are assigned higher interest rates.
installment	The monthly installments ($) owed by the borrower if the loan is funded.
log.annual.inc	The natural log of the self-reported annual income of the borrower.
dti	The debt-to-income ratio of the borrower (amount of debt divided by annual income).
fico	The FICO credit score of the borrower.
days.with.cr.line	The number of days the borrower has had a credit line.
revol.bal	The borrower’s revolving balance (amount unpaid at the end of the credit line billing cycle).
revol.util	The borrower’s revolving line utilization rate (the amount of the credit line used relative to total credit available).
inq.last.6mths	The borrower’s number of inquiries by creditors in the last 6 months.
delinq.2yrs	The number of times the borrower had been 30+ days past due on a payment in the past 2 years.
pub.rec	The borrower’s number of derogatory public records (bankruptcy filings, tax liens, or judgments).
not.fully.paid	1 for not fully paid/default, and 0 otherwise.

Target Variable & Independent Variables. The target variable is ‘not.fully.paid’. it contains two classes only, 1 for default and 0 otherwise. All the other variables are independent variables used in the model.

Data Preparation. All the NA values were dropped from the dataset using Pandas’ ‘dropna()’ method.

Methods of Analysis

Exploratory Data Analysis.

Target Variable

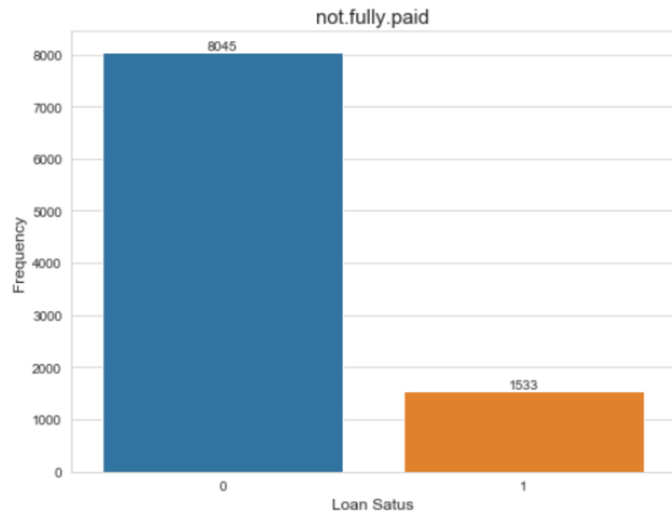

Figure 1 : Count Plot of the Target Variable

There are 9578 observations. 1533 observations show default, meaning 16% of the borrowers in the dataset defaulted.

Correlations Among Variables

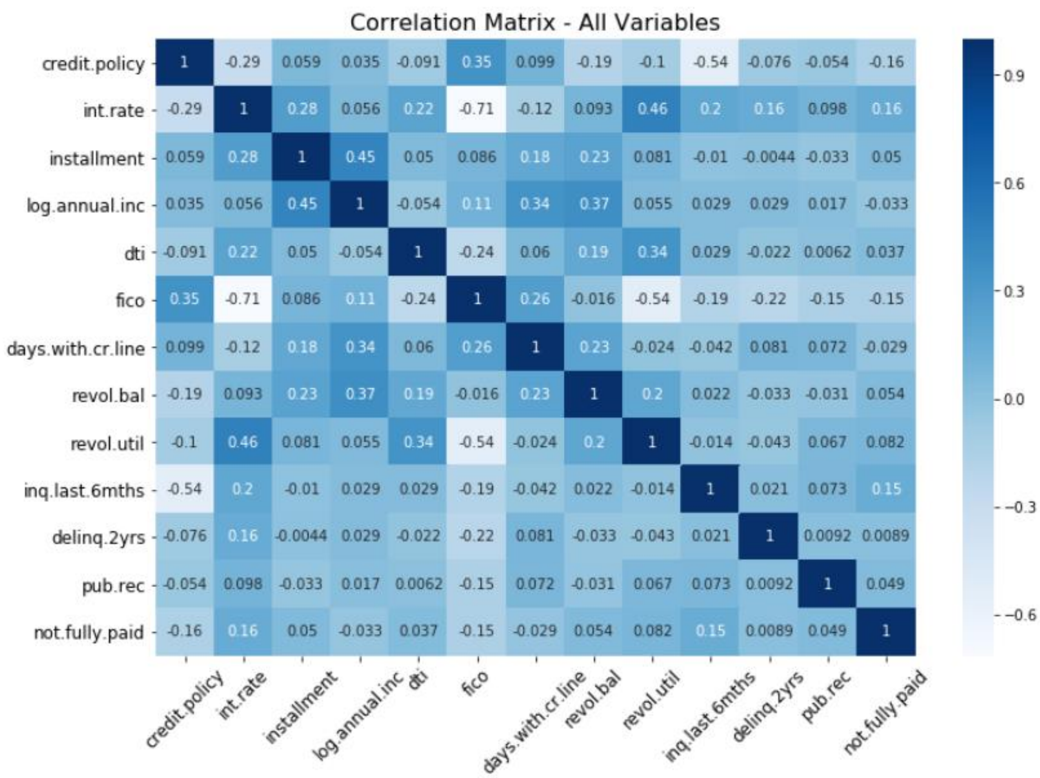

Figure 2: Correlation Matrix Using All Variables

The matrix shows 3 strong relationships relative to the other correlations. The strongest correlation is observed between the FICO score and the interest rate with a negative correlation of -0.71. The second strongest relationship is given by the correlation between the credit policy variable and the number of inquiries in the past 6 months with a negative correlation of -0.54. Finally, utilization rate and FICO scores also have a relative high correlation of -0.54. Figure 3 and 4 show the scatter plots of the 2 strongest correlations.

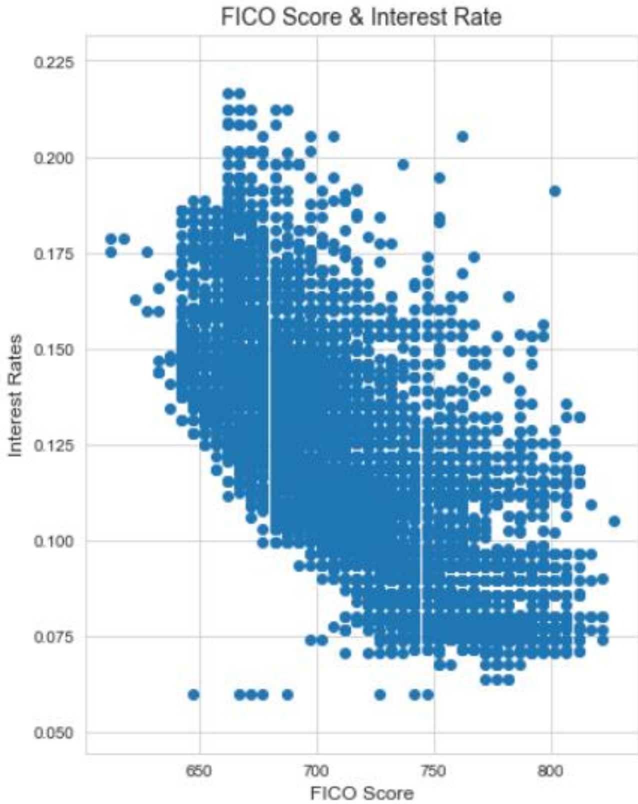

Figure 3: Scatter Plot of FICO & Int. Rates

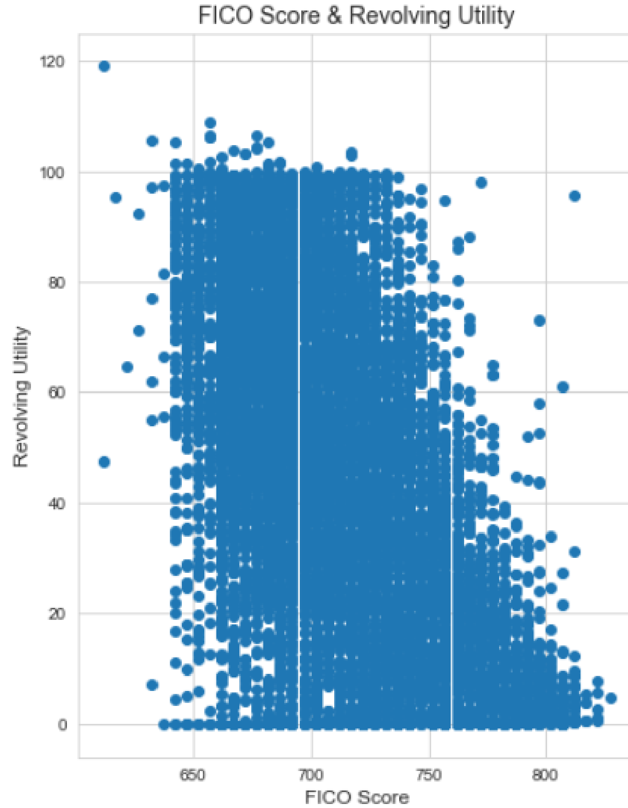

Figure 4: Scatter Plot of FICO & Revolving Utility

Figure 3 and 4 both show a negative correlation. The higher the FICO scores, the lower the interest rate being paid and the lower the revolving utility.

Distribution of FICO scores among the borrowers who met LendingClub's criteria versus the ones that did not.

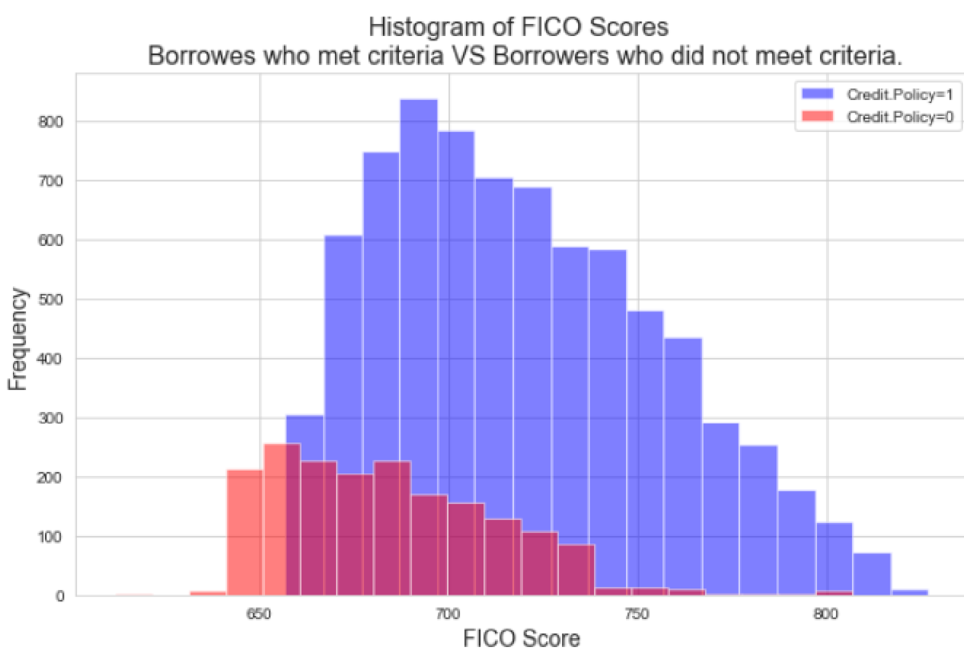

Figure 5: Distribution of FICO Scores Among Credit Policy

Figure 5 shows a visible noticeable difference in the distribution of FICO scores. Borrowers who met the criteria tend to be distributed on the higher end of the scores. Alternatively, borrowers who did not meet the criteria tend to group on the lower distribution of the scores. The distributions are positively skewed.

Distribution of FICO scores among the borrowers who defaulted versus the ones that did not.

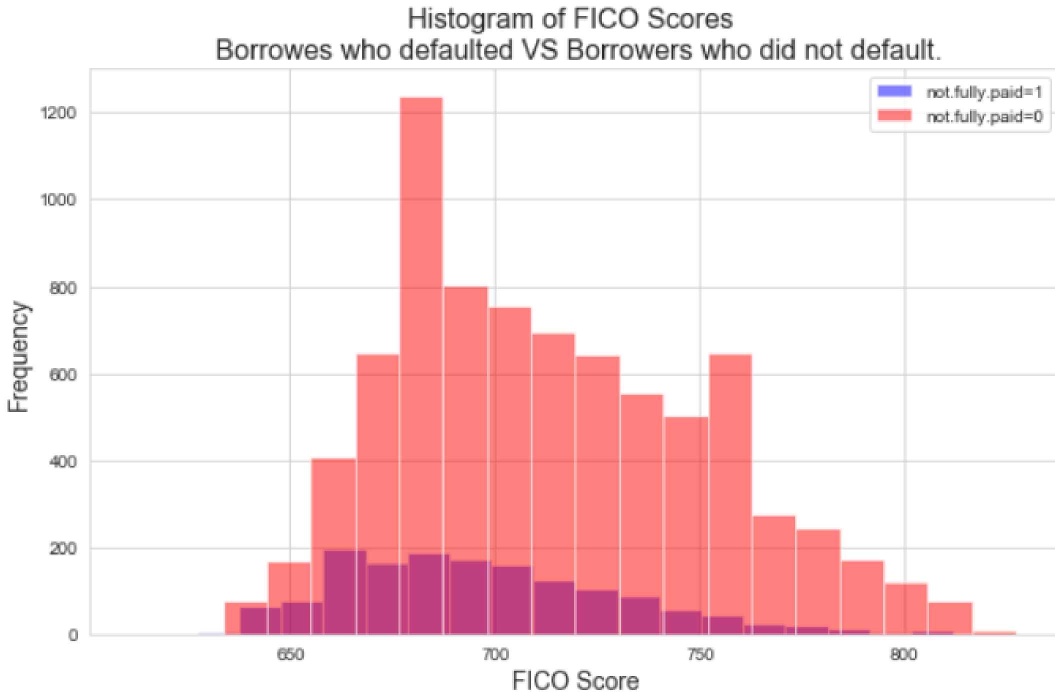

Figure 6: Distribution of FICO Scores Among the not.fully.paid Variable

Figure 6 displays a skewed distribution for both defaults and non-defaults. Most borrowers who default seem to have a FICO score in the rage between 675 and 725. Interestingly, defaulting borrowers with scores around 600 are less than the ones that scored between 625 and 725.

Distribution of Interest Rates among the borrowers who defaulted versus the ones that did not.

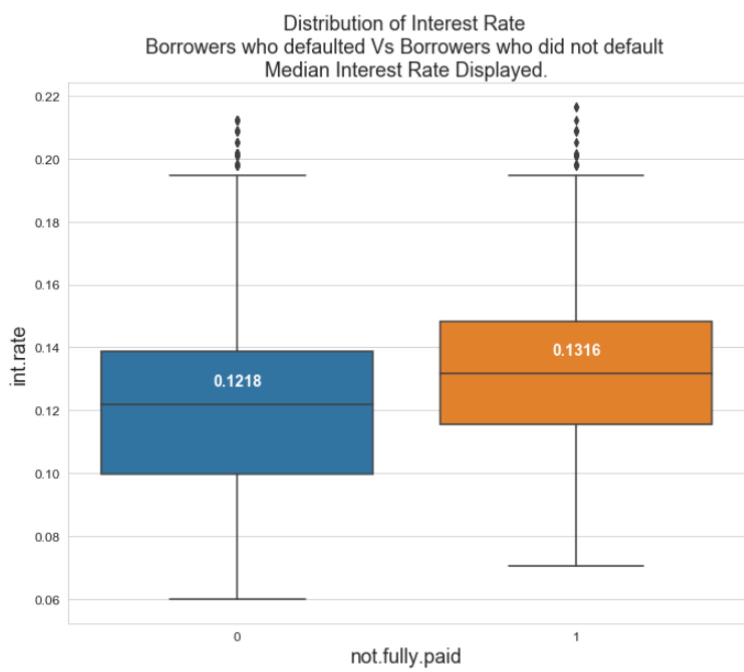

Figure 7: Distribution of Interest Rates Among the not.fully.paid Variable

The median interest rate among the borrowers who defaulted is higher. The median rates paid by defaulters is 8% more than the median rate paid by non-defaulters.

Distribution of Debt-to-Income Ratio among the borrowers who defaulted versus the ones that did not

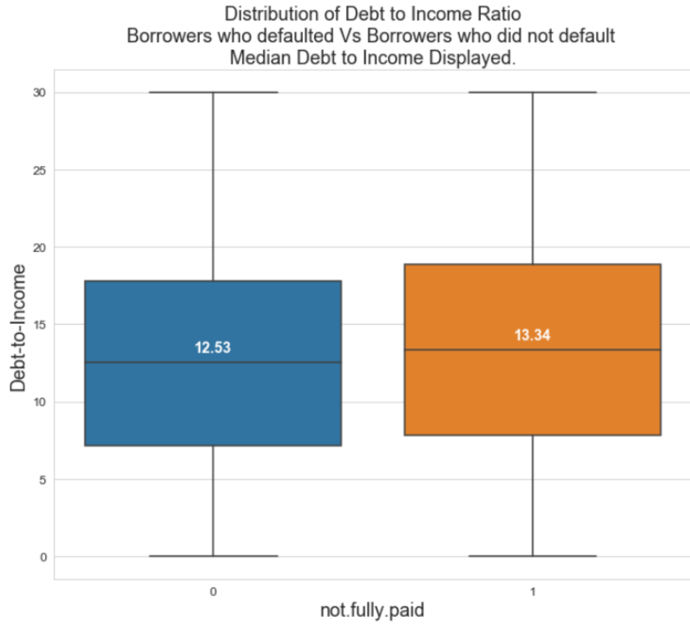

Figure 7: Distribution of Debt-to-Income Among the not.fully.paid Variable

The median debt-to-income ratio for defaulting borrowers is about 6.5% higher than the non-defaulting borrowers.

Distribution of the installment amounts among the borrowers who defaulted versus the ones that did not

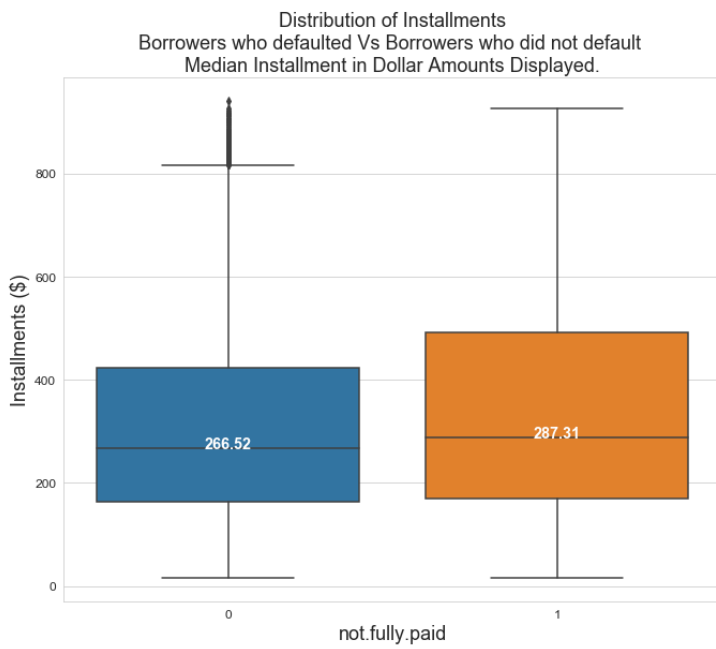

Figure 8: Distribution of Installments Among the not.fully.paid Variable

The median installment amount for defaulting borrowers is about 7% higher than the non-defaulting borrowers.

Purpose of the loan across defaults and non-defaults

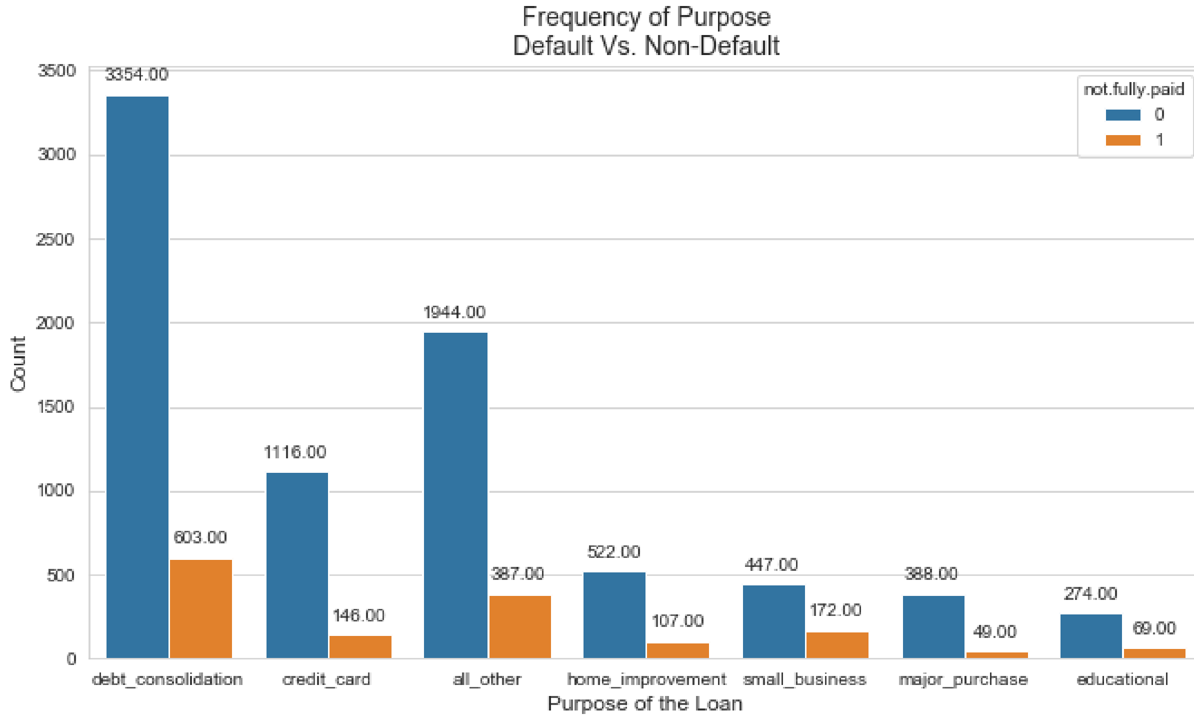

Figure 9: Count Plot of the purpose Variable

Most of the borrowers in the dataset got their loans to get out of debt. Consolidation of debt is a common strategy people pursue to improve their credit score. Essentially, all their debt is transferred to a one single creditor.

Random Forest Classification.

Random Forest is a supervised learning algorithm based on classification and regression. It is composed of multiple Decision Trees. The algorithm predicts classes, in this case default/non-default based on the training data that it used for the learning process. Additionally, it can also provide the mean probability of the outcome of all the trees in the model.

Transforming Variables and Imbalanced Target. Before building the model, the ‘purpose’ variable had to be transformed to multiple dummy variables. Every class became its own binary variable. For instance, the new dataset has a binary variable for ‘debt_consolidation’. This new variable has two classes, 1 if the purpose is debt consolidation and 0 otherwise. Additionally, I had to deal with the imbalanced classes in my target variable, ‘not.fully.paid’. Machine learning models work better when the samples in each class/category are close to equal (balanced). The data I am using for this model has a large imbalance in the 'not.fully.paid' vector (see Figure 1). There are a handful of methods to deal with imbalanced data, most of them rely on resampling techniques.[5] Given that I have too many "Good" observations, meaning I have an imbalanced number of non-defaults in the target variable, I decided to under sample the majority class to match the minority class. In order to accomplish this successfully, I need to make sure I split my data first.

Preliminary Model. The data was split 70%/30%, training set and testing set respectively. The preliminary model was run while using 100 estimators, 100 trees. This model correctly predicted a default when it was an actual default about 29% of the times. It correctly predicted a non-default when it was an actual non-default about 29% of the times. About 21% of the times the model showed a Type 1 error, the model predicted default when it was not. Finally, about 21% of the times the model showed a Type 2 error, the model predicted a non-default when it was a default. Please see Figure 10 below for the Confusion Metrix. A simple way to analyze performance is with accuracy (number of correct predictions/number of predictions). To calculate the accuracy, I will use the score() method in the SciKit-Learn package. The method is applied to

[5] Tara Boyle 2019, Dealing with Imbalanced Data, towardsdatascience.com, viewed May 21, 2020, <https://towardsdatascience.com/methods-for-dealing-with-imbalanced-data-5b761be45a18>

the trained model and it provides a percentage of correct predictions. The accuracy score for the preliminary model was 60%.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Number\ of\ Predictions}$$

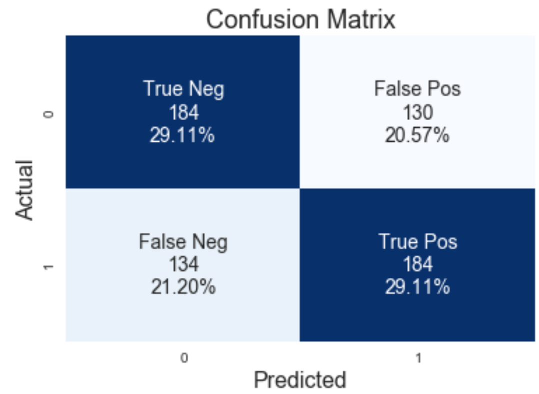

Figure 10: Confusion Matrix - Peliminary Model

Tuning Number of Estimators. The number of trees that we use in the model impact the accuracy score and model performance. Therefore, it is very important to find the optimal number of trees the model needs. For this, I use the GridSearchCv package from Scikit-learn. The best estimator is 250. This number of trees yields the highest means_test_score, which is the correct mean accuracy of the model.

Final Model-After Tuning. The final model uses the optimized number of trees, 250. The accuracy score increased by close to 1% only. There was no drastic difference from the preliminary model. There was a slight improvement in the confusion matrix as well. Please see Figure 11 below.

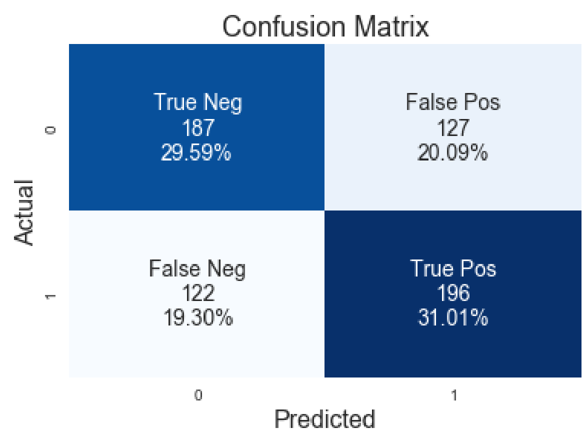

Figure 11: Confusion Matrix - Final Model

Predicting the Probability of Default at a 60% Threshold. To predict the probability of default for every observation in the test set, the final model was fed with the test data while using the 'predict_proba()' method. If the predicted probability of default for a given observation is greater than 60%, the observation is classified as a default. 60% is an arbitrary number I selected.

Model Performance. To visualize the performance of the model, I created a Receiver Operating Characteristics (ROC) plot. I plotted the false positive rate (fallout,x) and true positive rate (sensitivity,y). To calculate the area under the curve (AUC) I used the 'roc_auc_score()' method. The RCO curve shows how much the model is capable of distinguishing between classes. The higher the AUC, the better is the model at predicting the correct class, default/non-default.[6]

[6] Sarang Narkhede 2018, Understanding AUC – ROC Curve, towardsdatascience.com, <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>

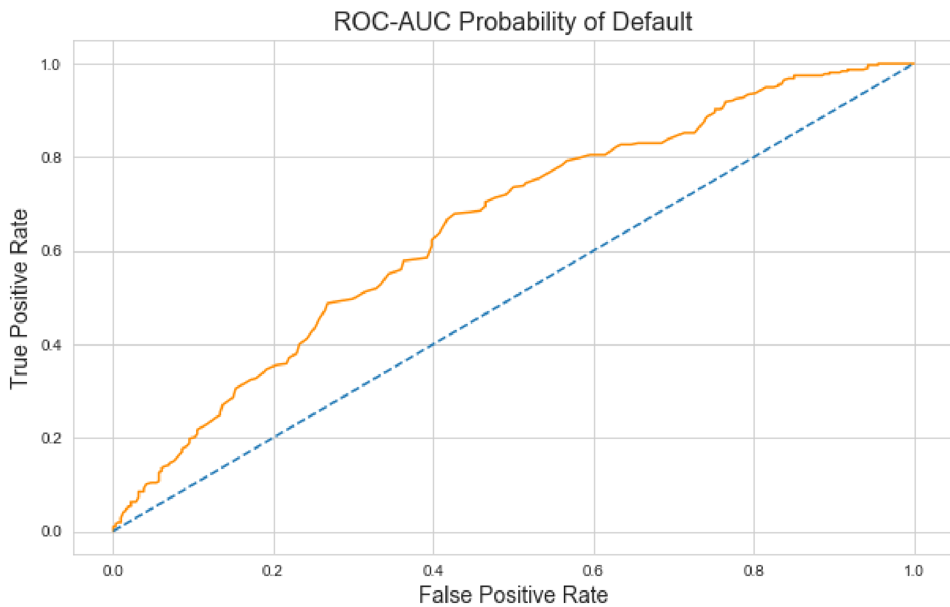

Figure 12: Receiver Operating CHaracteristics and Area Under the Curve Plot

The dotted blue line represents a random prediction and the orange line represents the model prediction. The further away the orange line moves from the random prediction, the larger the area under the curve gets. The AUC is the calculation of the area between the model predictions and the random predictions. Finally, another way to asses the performance of the model is by looking at the roc_auc_score. This score is a direct indicator of how well the model is capable of distinguishing between default and non-default. The higher the number the better the model. This model has a roc_auc_score of 0.651.

Estimating Expected Loss

$$EL = Predicted\ Defaults * Avg\ Loan\ Amnt * Default\ Recall$$

The calculated EL is based on a $10,000 average loan amount represents the expected loss if we were to invest in the test dataset. Once the model was completed and the performance assessed, I will run test data in the model to analyze the predicted defaults and the predicted expected loss. 5 scenarios were model. For each scenario one variable at the time was inflated/deflated using factor of 0.9 or 1.1 while all the other variables kept equal.

Scenarios:

- Interest Rate – 10% decrease in interest rates across all observation in the test dataset.
- Days with Credit Line – 10% increase in the number of days with a credit line.
- Installment Amount – 10% decrease in installment amounts.
- Annual Income – 10% increase in log annual income.
- Revolving Balance – 10% decrease in revolving balance.
- Revolving Utility – 10% decrease in revolving utility.

Results

Predicting defaults is not an easy task. There are an infinite number of variables that influence the default or non-default outcome. In this research was constrained by the limited number of variables used. The Random Forest Classification model provided some insights as to how influential these variables are. Figure 13 (below) show a breakdown of the importance each variable has in the model; this is an indicator of how much they influence the target variable, in this case, the 'not.fully.paid'. The ranking was obtained by applying the 'feature_importance_' method to the final random forest model.

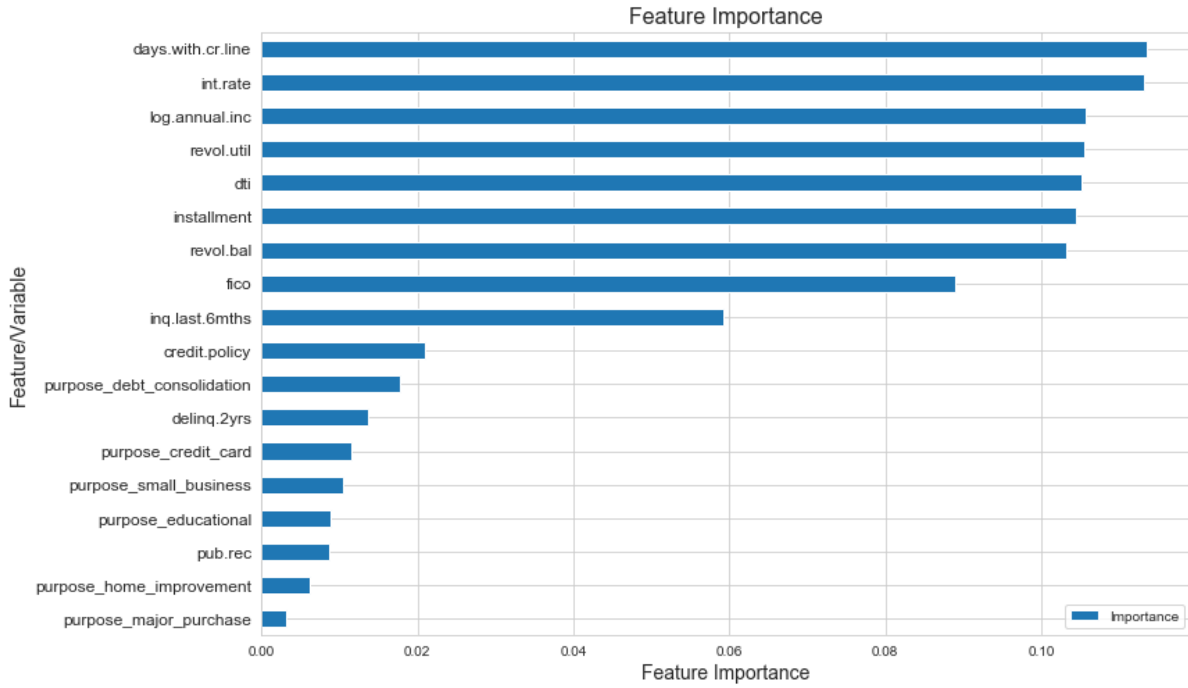

Figure 13: Horizontal Bar Plot of The Feature Importance Ranking

The top 5 important features are interest rate, the number of days with credit line, the installment amount, the log annual income, and the revolving utility. It is worth noting that the FICO score does not come up in the top 5, nor does debt to income ration.

Scenario Analysis.

Table 1

Scenario Analysis Output

	Factor	Predicted Defaults	Model Defaults	Pct. Chg Defaults	Expected Loss	EL Benchmark	EL-Difference
int.rate	0.90	149.00	149.00	0.00	431069.18	435754.72	-4685.54
days.with.cr.line	1.10	141.00	149.00	0.05	394622.64	435754.72	-41132.08
installment	0.90	140.00	149.00	0.06	378616.35	435754.72	-57138.37
log.annual.inc	1.10	100.00	149.00	0.33	198113.21	435754.72	-237641.51
revol.bal	0.90	149.00	149.00	0.00	445125.79	435754.72	9371.07
revol.util	0.90	142.00	149.00	0.05	397421.38	435754.72	-38333.34

The top 5 important features in Figure 13 were used for the scenario analyses. The “EL Benchmark” and the “Model Defaults” in Table 1 are the outputs of the actual model, before inflating or deflating the test dataset. Table 1 shows that decreasing the interest rate 10% would reduce the expected loss by $4,685.54. Similarly, increasing the log annual income by 10% would decrease the number of defaults by almost 32%. For this model, this translates into a $237,641.51 decrease in expected loss. It appears that the prediction model is most sensitive to the log annual income and the revolving utility. Since this research aims to explore indicators, other than FICO, for predicting default, the FICO variable was not changed.

The factors in Table 1 were selected with the intend of reducing the number of defaults, and therefore, reducing the expected loss. It is in the best interest of the lender to reduce the number of defaults and collect the repayment successfully. By leaving the FICO scores constant, we find that the other features of the model may provide some more information.

Natural Language Processing.

Natural Language Processing was used to do a text analysis on 15 current news articles. The articles were scraped directly from the news' websites using the beautifulsoup4 python package. The articles were retrieved from The Economist, The Barron's, Forbes, Market Watch, and CNBC among other (please see the Reference section for a list of the articles and their respective links). All the news are finance and business related; moreover, they were purposely selected based on their credit relevance. A corpus was built with all the articles. The purpose of this analysis is to asses the predominant topics related to credit and to make inferences on the current sentiment in credit. I used the NLTK Python library to analyze the corpus.

Unigrams

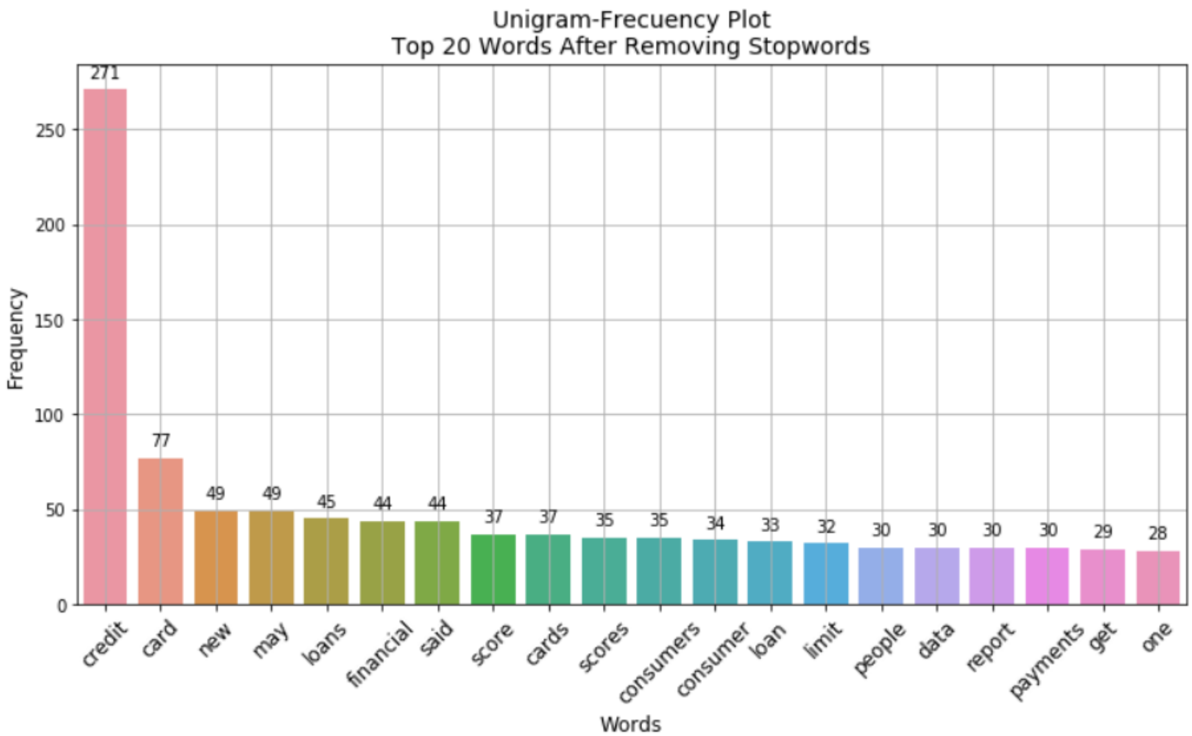

Figure 14: Frequency Chart - Top 20 Words in The Entire Corpus.

Figure 14 shows the top 20 unigrams, words, in the corpus. Not surprisingly the unigram with the most frequency is "credit" since all articles are related to credit news. "loans", "card", and "score" also appear multiple times across the corpus.

Bigrams

Bigrams are sequence pairs of words. Looking at a single word in the corpus may provide some information, however, single words lack context. Therefore, a bigram analysis was also performed. The context of the corpus is better assessed when the "stopwords" are still in the corpus and so they were not removed for this analysis. Here is the frequency of the top 20 bigrams in the corpus.

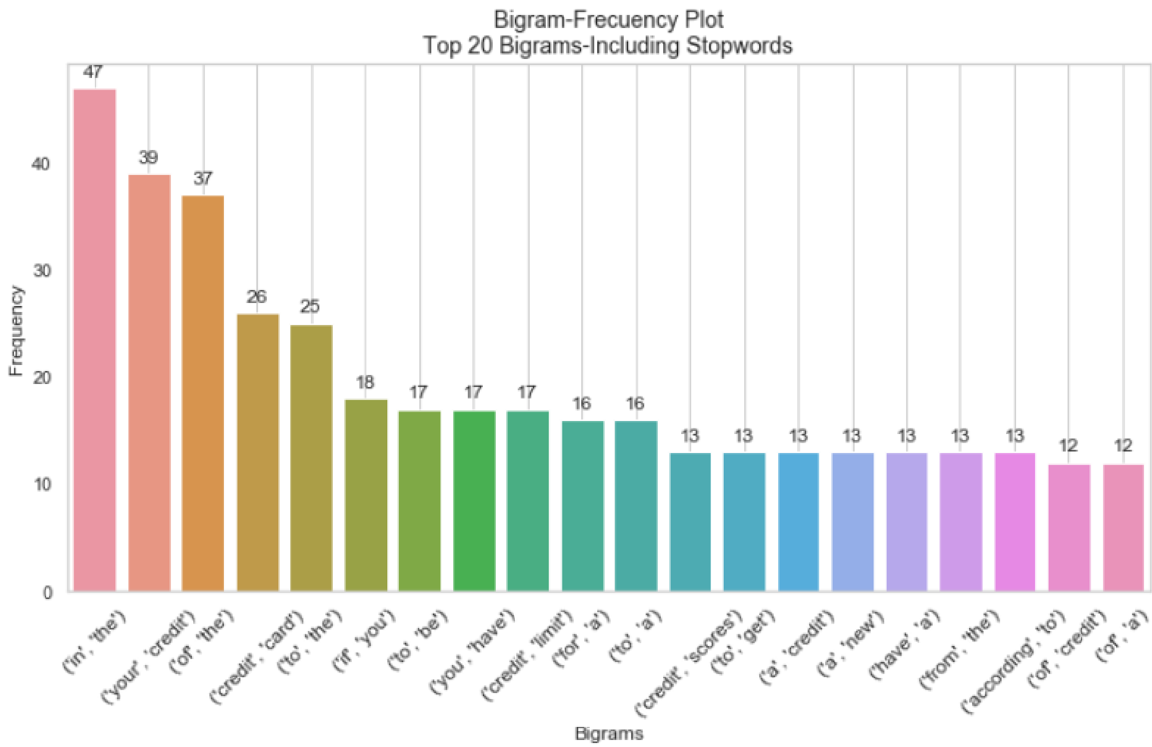

Figure 15: Frequency Chart - Top 20 Bigrams in The Entire Corpus.

Some bigrams do not provide a lot information such as, “in the”, “of the”, “to the”, and “to be”. The bigrams that may provide some information are: “your credit”, “credit card”, “credit limit”, and “credit scores”.

Trigrams

A trigram is composed of 3 consecutive words in a sentence. Trigrams may provide further context to the corpus.

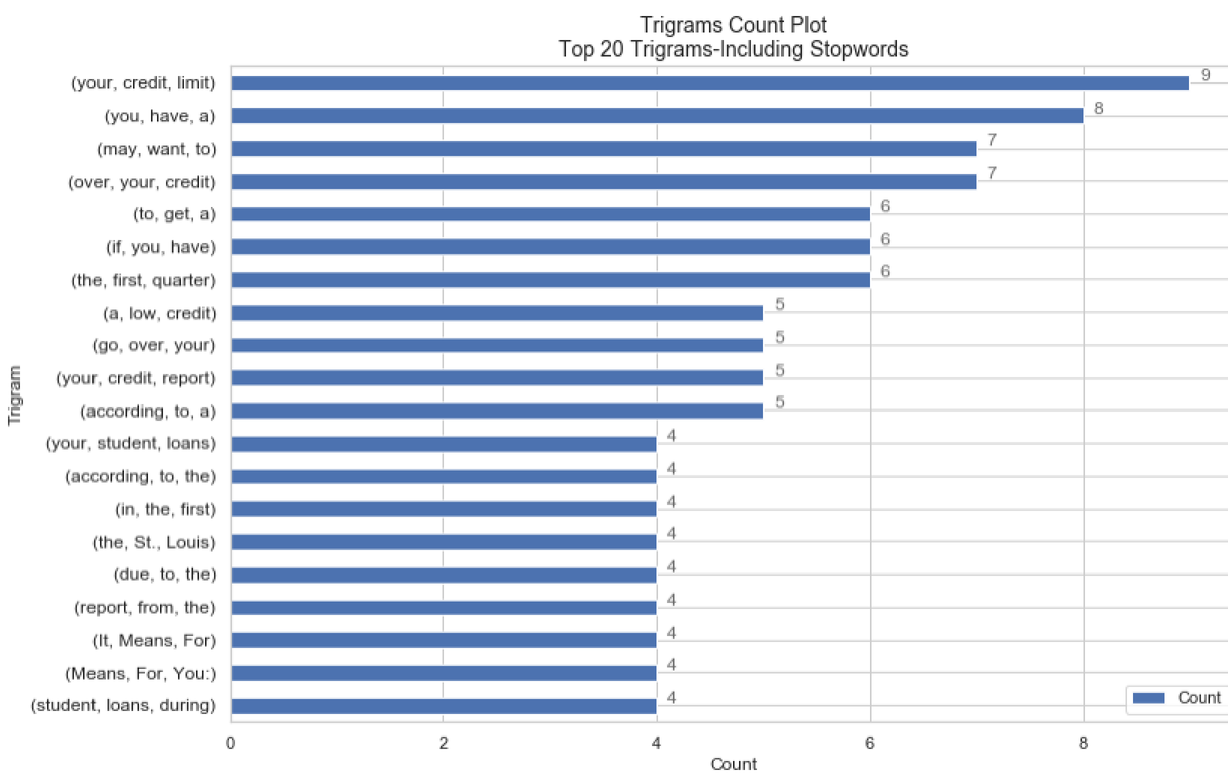

Figure 16: Horizontal Bar Plot of The Top 20 Trigrams

Figure 16 shows the top 20 trigrams in the corpus of news. At a high level, it is evident that more inferences could be drawn. Credit limit seems to be the topic that dominates in the corpus.

Topic Modeling

Topics can be modeled by using Latent Dirichlet Allocation (LDA). This is an unsupervised machine-learning model that outputs the topics in the corpus. In a nutshell, this process clusters the tokens, words, in a given document by their topic. It is not sensitive to the order of the words. For this analysis, the LDA model uses the entire corpus to model topics, instead of modeling the top on a per article basis. The LDA model takes in the number of topics *k* as one of its hyperparameters. This number of topics *k* was optimized by plotting the coherence score by a range of topics from 1 to 10.

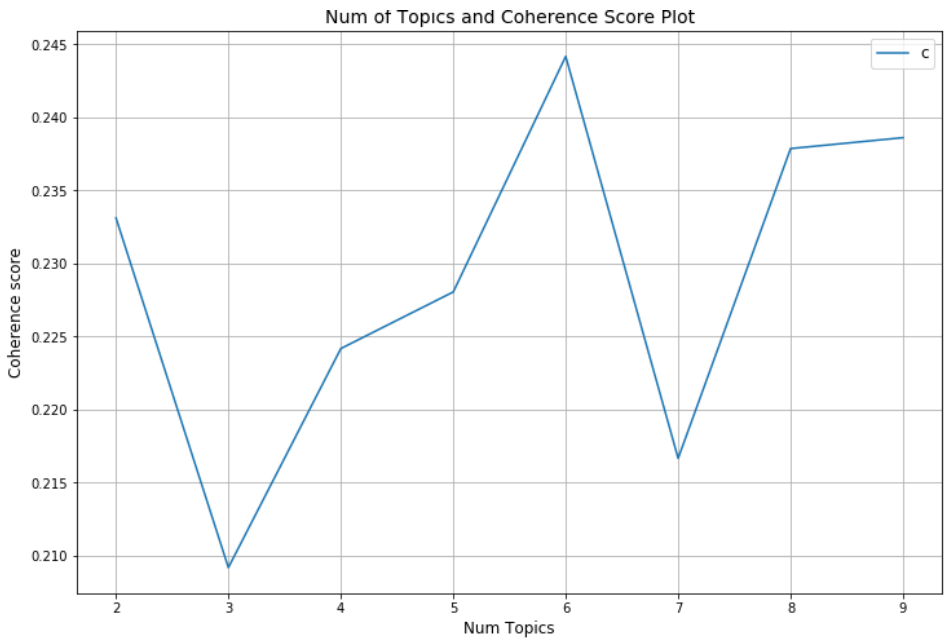

Figure 17: Number of Topics and Coherence Scores Plot.

Figure 17 shows 6 as the optimal number of topics for the LDA model. This is the number of topics used in the model.

Results

The NLP analysis provides an overview of the current topics in credit. As seen in Figure 16, the news outlets seem to be taking about the consumer credit limits. Given the poor and fragile state of the economy banks are cutting down credit limits to mitigate the risk. During this difficult time, unemployment is at the highest it has ever been. Households are having problems keeping up with their expenses and they may be forced to rely on their credit cards or credit lines to support their needs. Households are turning to banks to get cash support and as a result of the higher risk, lenders are tightening the approval standards. Additionally, another preoccupation in the public is student loans. In the trigram plot, student loans appear twice. There are 44.7 million borrowers with student loans. And the total debt in student loans amounts to almost \$1.6 trillion.[7] It is no wonder why it comes up in the top trigrams; lenders are concerned this huge debt won't be repaid.

The LDA topic model was built while using the optimal number of topics. The interpretation of words needs a linguistic approach. The study of words' connotations and semantics is a complex process and an interdisciplinary approach is needed. Nonetheless, for the scope of this research, a basic interpretation may be drawn from the LDA model. Here are the results of the top 3 topics.

---

[7] Zack Friedman 2020, Student Laon Debt Statistics In 2020: A Record \$1.6 Trillion, forbes.com, viewed May 21, 2020, <https://www.forbes.com/sites/zackfriedman/2020/02/03/student-loan-debt-statistics/#264b61c7281f>

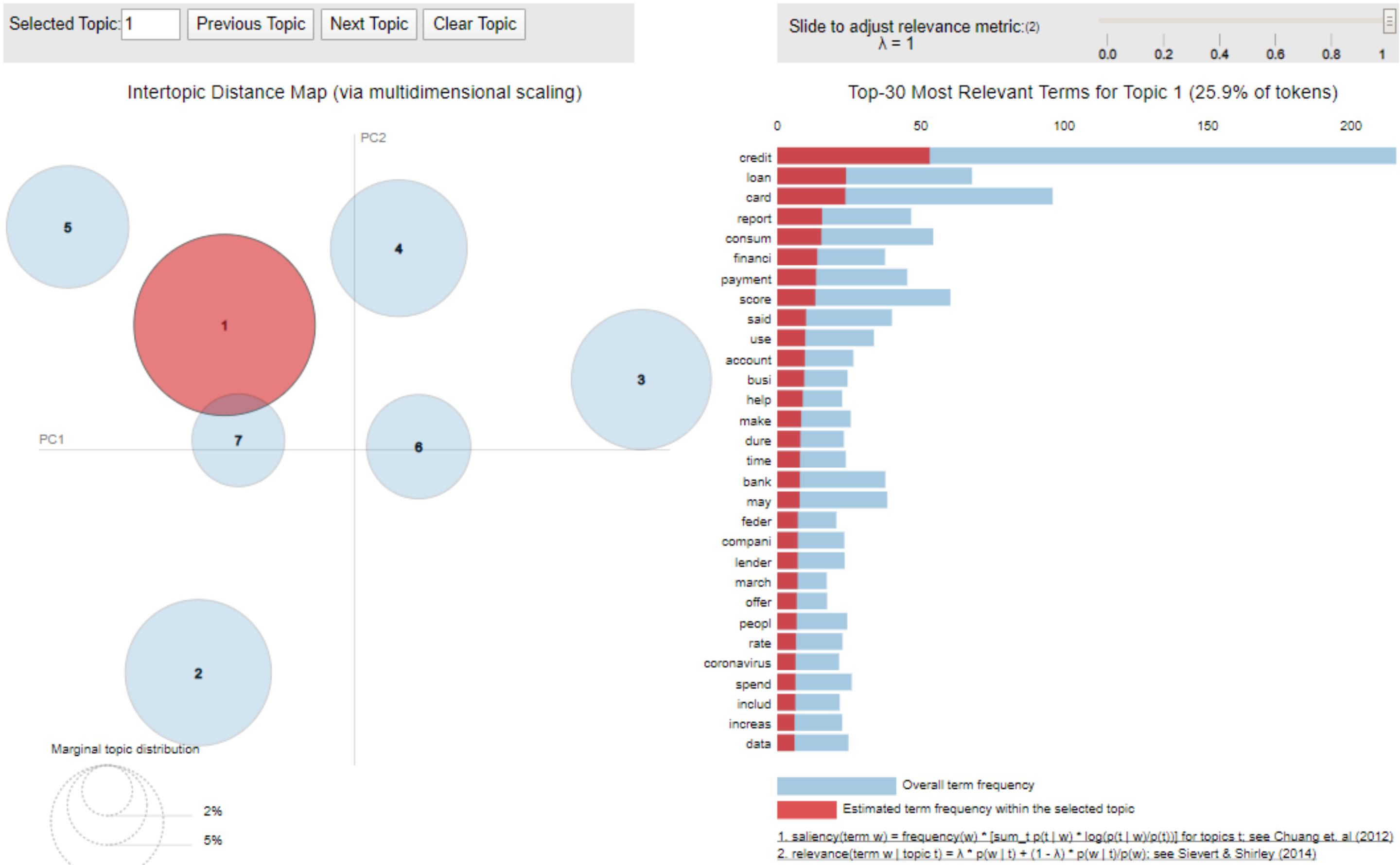

Figure 18: LDA Topic Analysis Visualization – Topic 1 – Top 30 Most Relevant Terms

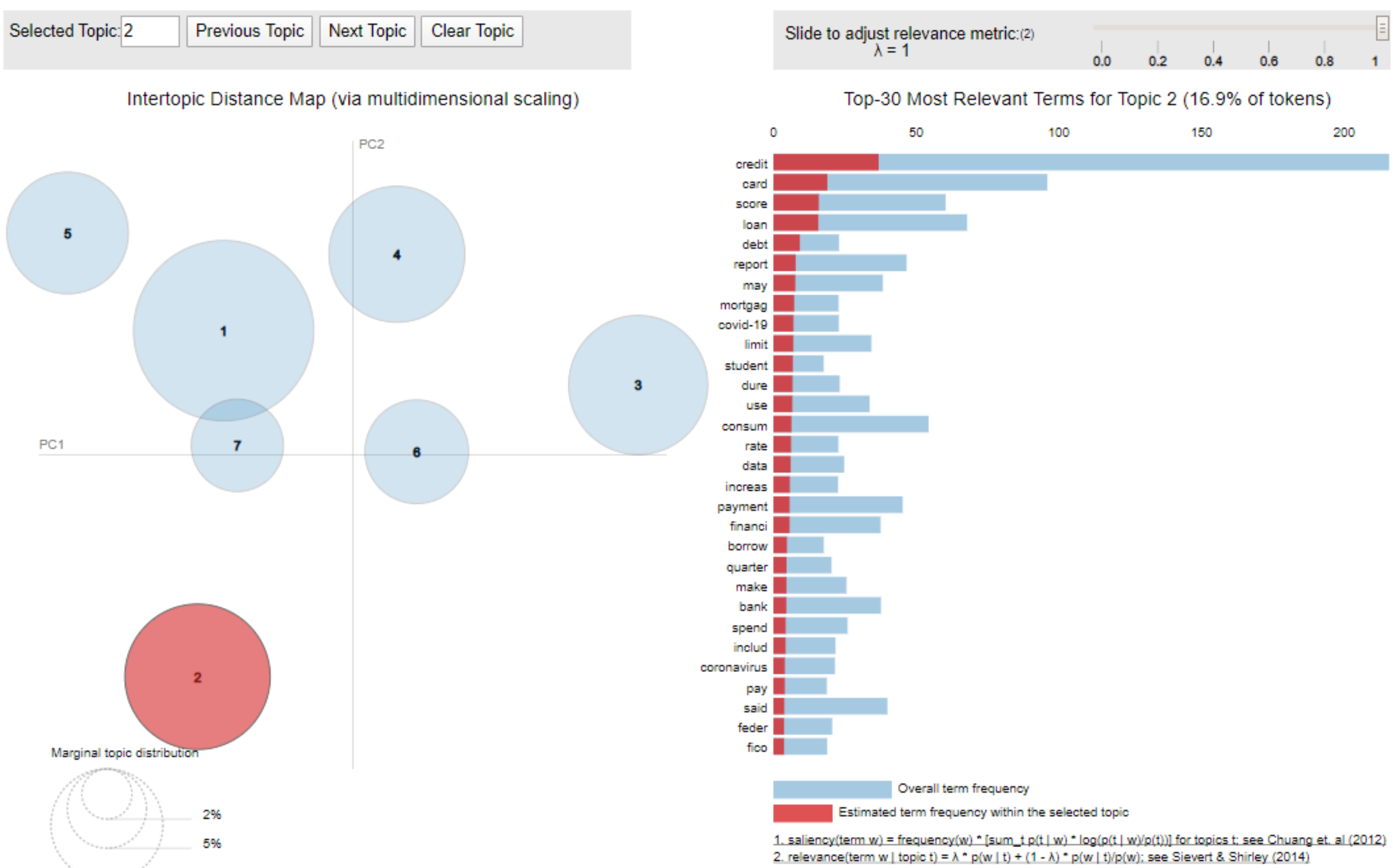

Figure 19: LDA Topic Analysis Visualization – Topic 2 – Top 30 Most Relevant Terms

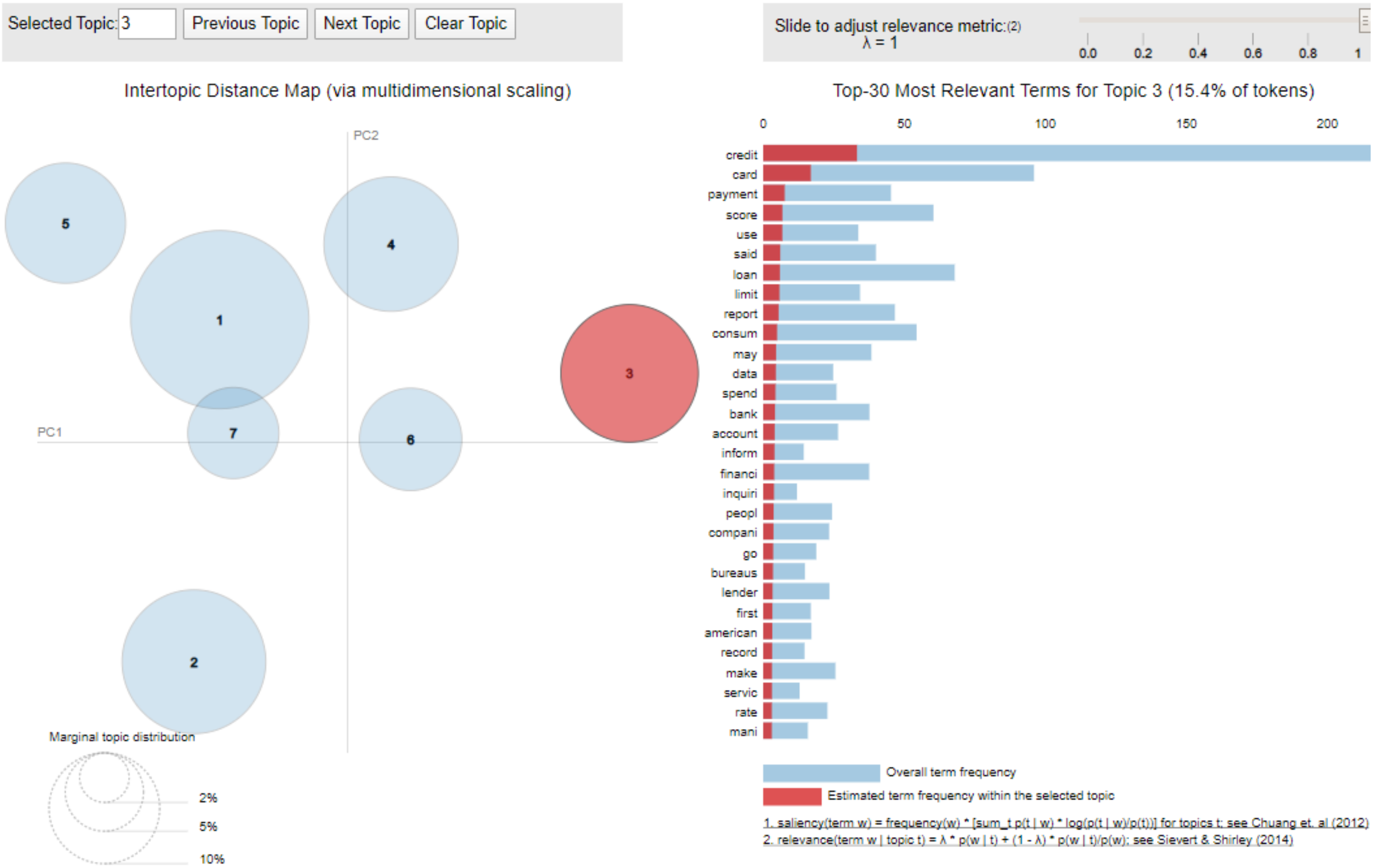

Figure 19: LDA Topic Analysis Visualization – Topic 3 – Top 30 Most Relevant Terms

There are some patterns in all 3 topics. The biggest one seems to be credit card loans, followed by credit score and payments. Additionally, 2 of the main topics are coronavirus related as well. At this time the covid-19 pandemic may be the principal cause of credit default talks, the poor state of the economy and the unnaturally high unemployment rate.

Credit scores may be affected by extreme natural disasters or by a pandemic, which is what is happening right now as some people are not able to make payments on their loans. [8]Nobody knows how severe these defaults may get as news outlets are forecasting a slow recovery.[9] In situations like this, other variables may provide a better indicator of the loan repayment likelihood, such as, employment, or the installment amount.

Conclusion

One could argue that lending companies should concentrate their efforts in analyzing the forecasted revolving utility, the installment amounts, or the future log annual income in order to weight more on these variables at the time of assessing future defaults. Those variables showed to be the most sensitive in the random forest runs. There is a spurious dilemma when trying to attribute causes to defaults. The exploratory data analysis showed how borrowers who defaulted were paying more interest rate and had greater installment amounts than the borrowers who did not default. But did high interest rates and installments lead to default? And, isn't it the low FICO score that determines how much interest the borrowers will pay in the first place? More research needs to be done to determine causation of the defaults. Previous studies show that FICO scores are the best tool we have available to predict creditworthiness and defaults. FICO scores are of high importance when forecasting defaults, but as new technologies emerge, such as machine-learning and AI, the weight on the FICO scores

[8] Andrew Soergel 2020, World's Banks Brace for Rise in Loan Defaults, usnews.com, viewed May 21, 2020, < https://www.usnews.com/news/best-countries/articles/2020-04-17/threat-of-bad-loans-looms-over-banks-during-coronavirus-pandemic>

[9] Dhara Ranasinghe, Ritvik Carvalho 2020, Alphabet soup: How will post-virus economic recovery shape up?, reuters.com, viewed May 21, 2020, < https://www.reuters.com/article/us-health-coronavirus-economy-graphic/alphabet-soup-how-will-post-virus-economic-recovery-shape-up-idUSKCN21R242>

should be reduced to account for other more sensitive variables, like consumer spending habits and forecasted revolving utility.

Limitations. The scope of the dataset from LendingClub.com only comprises 3 years of observation. This is not enough data to represent the current credit circumstances. The data was also adapted for this research, as many as 40 features were removed in order for this research to be fishable given the time constrain.

References

Please see foot notes for references.

Internet Articles Used for the NLP Analysis

1. https://www.economist.com/international/2019/07/06/a-brief-history-and-future-of-credit-scores
2. https://www.arkansasonline.com/news/2020/may/10/banks-tighten-lending-standards-amid-mo/
3. https://www.ocregister.com/2020/05/05/u-s-household-debt-reaches-another-record-on-home-loans/
4. https://www.barrons.com/articles/financially-distressed-households-face-lengthier-consumption-recovery-heres-why-that-matters-51588105801
5. https://www.forbes.com/sites/zackfriedman/2020/05/07/consumer-debt-financial-crisis/#158581486ded
6. https://www.forbes.com/sites/billhardekopf/2020/05/08/this-week-in-credit-card-news-more-issuers-cut-credit-limits-card-use-takes-huge-drop-during-pandemic/#43e5d1f25b13
7. https://www.forbes.com/sites/billhardekopf/2020/05/01/this-week-in-credit-card-news-credit-gets-even-tighter-contactless-payments-skyrocket-during-pandemic/#192723f91097
8. https://www.forbes.com/sites/forbesfinancecouncil/2020/05/08/beyond-credit-scores-how-fintech-companies-are-leveraging-tech-to-see-a-bigger-picture/#437fda031315
9. https://www.marketwatch.com/story/consumer-borrowing-drops-in-march-as-households-shy-away-from-using-credit-cards-2020-05-07
10. https://www.natlawreview.com/article/cfpb-report-shows-credit-applications-dropped-significantly-march-2020
11. https://newyork.cbslocal.com/2020/05/12/coronavirus-covid-19-comfort-spending-retails-shopping-online-paul-oster/
12. https://www.cnn.com/2020/05/13/success/available-consumer-credit-coronavirus/index.html
13. https://www.americanbanker.com/list/5-ways-the-cfpb-has-eased-industrys-coronavirus-burden
14. https://www.nytimes.com/2020/05/15/your-money/coronavirus-credit-reports.html
15. https://www.cnbc.com/select/exceeding-credit-limit/