**Financial Analytics Assignment**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Prepared By** | **SID** | **TAID** |
| **Name** | Rajib Mandal | 118310 | TA17002 |
| **Submitted Date** | 20-Sep -2018 | | |

# Executive Summary

Lending Club (<https://www.lendingclub.com/>) is a US [peer-to-peer lending](https://en.wikipedia.org/wiki/Peer-to-peer_lending) company, headquartered in [San Francisco, California](https://en.wikipedia.org/wiki/San_Francisco,_California).[[3]](https://en.wikipedia.org/wiki/Lending_Club#cite_note-BusinessInsider2014December04McBrideSarah-3) It was the first peer-to-peer lender to register its offerings as [securities](https://en.wikipedia.org/wiki/Security_(finance)) with the [Securities and Exchange Commission](https://en.wikipedia.org/wiki/U.S._Securities_and_Exchange_Commission) (SEC), and to offer loan trading on a secondary market. Lending Club enables borrowers to create [unsecured personal loans](https://en.wikipedia.org/wiki/Unsecured_debt) between $1,000 and $40,000.

The standard loan period is three years. (https://en.wikipedia.org/wiki/Lending\_Club)

It has transformed the banking system to make credit more affordable and investing more rewarding. But it comes with a high risk of borrowers defaulting the loans. Hence there is a need to classify each borrower as defaulter or not using the data collected when the loan has been given.

# Project Objective

The goal is to build a prediction model that will classify if the borrower will default the loan using borrower’s finance history. we need to predict the target variable as 1 -> Defaulter or 0 -> Non-Defaulter

# Data Set Information

The data is collected from the lending club website(<https://www.lendingclub.com/info/download-data.action>) for the years 2014. The dataset consists of **235,629** observations and **145** features. Out of the 145 features in the dataset, many of them were empty. The dataset information is given **Appendix A.**

Out of 145 columns, 31 (Refer **Appendix B**) do not have any value. These 31 columns are removed from the dataset.

The dataset consists of one borrower detail per line. The summary is given below.

|  |  |  |
| --- | --- | --- |
| Loan Status | Amount | % |
| Defaulter | 41,372 | 17.55 |
| Non-Defaulter | 194,257 | 82.45 |
| Total | 235,629 | 100 |

# Pre-processing

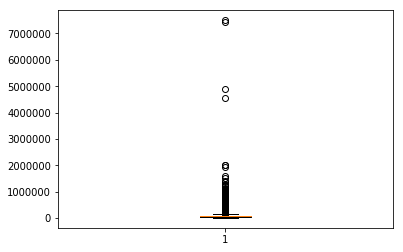
The following steps are performed during data pre-processing. we chose 22 best features related to our objective

**Data Transformation**

* Text fields are converted to integer
* Categorical values (Grade, loan\_status, home\_ownership etc ) have been transformed to numerical
* Redundant fields are dropped from the dataset during analysis.

**Outlier Detection**

* Identify the outliers from the dataset (fields -loan\_amnt, annual\_inc) and remove them from the dataset. We found Outlier found in annual\_inc fields. The corresponding records are removed from the dataset.



**Check for Normality**

### Perform Shapiro-Wilk Test to evaluate the data sample and quantifies if the data was drawn from a Gaussian distribution. We found following value.

|  |  |
| --- | --- |
| Statistics | 0.858 |
| p-value | 0.000 |

Since our p-value is much less than our Test Statistic, we have good evidence to reject the null hypothesis at the 0.05 significance level.

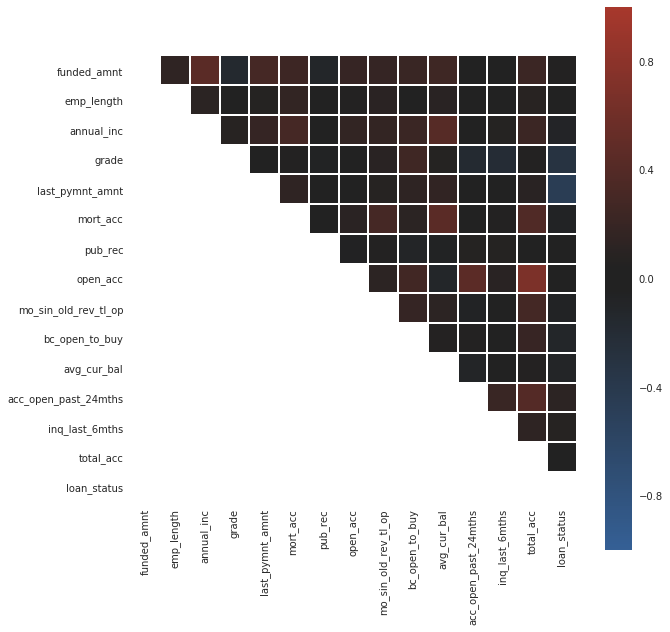
Hence, we conclude that sample is not Gaussian.

**Filling Missing values and Feature scaling**

* There are some important features which have some missing values. Impute NAN values with median values of corresponding columns(numerical)
* Standardize features by removing the mean and scaling to unit variance using standard scaler

**Heat map**

 Two-dimensional representation of data in which values are represented by colors.Below heat map provides an immediate visual summary of feature.



**Predictor Variables:**

On the above 22 features, we have implemented Recursive Feature Elimination (RFE) using Logistic Regression model to get the best 15 features.

* funded\_amnt
* emp\_length
* annual\_inc
* grade
* last\_pymnt\_amnt
* mort\_acc
* pub\_rec
* open\_acc
* mo\_sin\_old\_rev\_tl\_op
* bc\_open\_to\_buy
* avg\_cur\_bal
* acc\_open\_past\_24mths
* inq\_last\_6mths
* total\_acc

**Target Variable:**

The target variable in our dataset is ‘loan\_status’ which infers the status of the loan. It has 2 different values – ‘Charged Off’, ‘Fully Paid’

Fully Paid: Loan has been fully repaid.

Charged Off: Loan for which there is no longer a reasonable expectation of further payments.

**Prepare Training Data & Testing Data**

We have prepared training & testing data set from the given data set with 70:30 ratio

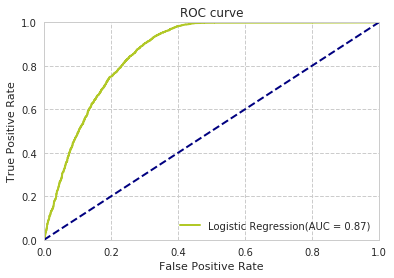
# Model building

We have used supervised learning algorithms like K-nearest neighbors, support vector machines logistic regression ,Random Forest for building the model

**Logistic Regression:**

The adjusted R -squared value for the model is 0.394. It indicates 39.4% of the variance in the [dependent variable](http://statisticsbyjim.com/glossary/response-variables/) (i.e. loan\_status’) that the [independent variables](http://statisticsbyjim.com/glossary/predictor-variables/) explain collectively

The accuracy for Logistic Regression Model is **80.22%.** The plot of the True Positive Rate(TPR) against the False Positive Rate(FPR) i.e. ROC curve is given below. The area under the curve of the model is 0.87

****

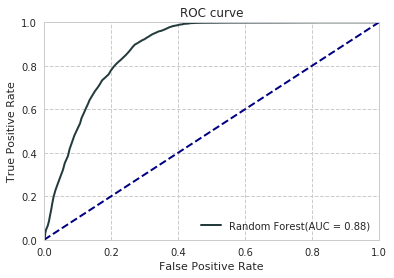
**Random Forests Classification:**

The following parameters are used during model building

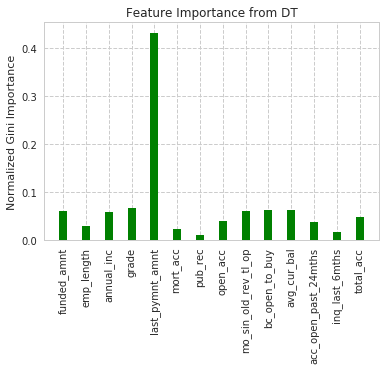
1. n\_estimators=100(The number of trees in the forest.)
2. criterion = "gini"( he function to measure the quality of a split)

Random forests select a subset of features in each of its decision trees thereby reducing the bias (because of high importance of single feature) of the model. The final output will be the mode of the outputs of all its decision trees which has better results than decision trees

The accuracy for Random Forest is **81.46%.** The plot of the True Positive Rate(TPR) against the False Positive Rate(FPR) i.e ROC curve is given below. The area under the curve of the model is 0.88

****

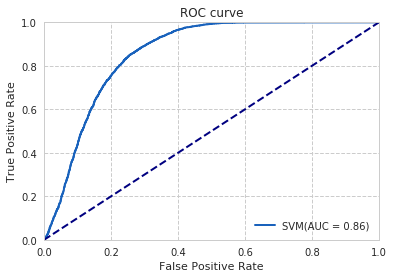
The feature ‘**last\_pymnt\_amnt’** which has an importance of more than 40%(Refer below graph).



**Support Vector Machine:**

We consider SVM classifier because, once a hyperplane is found, most of the data other than the support vectors (which are points closest to the boundary) become redundant. This means that small changes to data cannot greatly affect the hyperplane and hence the SVM.We use radial basis function (rbf) kernel in this model.

The accuracy for **SVM Model is 80.24%.** The plot of the True Positive Rate (TPR) against the False Positive Rate (FPR) i.e ROC curve is given below. The area under the curve of the model is 0.86

****

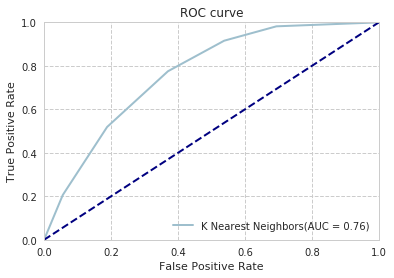
**K Nearest Neighbours(KNN):**

The following parameters are used during model building

n\_neighbors**: Number of neighbors 5**

weights='uniform' i.e All points in each neighbourhood are weighted equally.

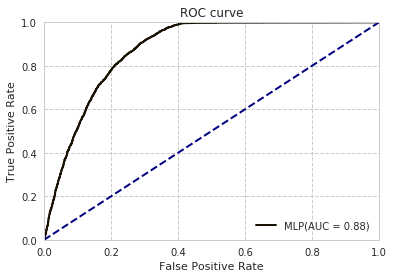
The accuracy for **KNN Model is 70.59%.** The plot of the True Positive Rate(TPR) against the False Positive Rate(FPR) i.e ROC curve is given below. The area under the curve of the model is 0.76

****

**Multi-Layer Perceptron Classifier**

MLP utilizes backpropagation for training. Its multiple layers and non-linear activation function help us distinguish data that is not linearly separable.

The accuracy of the model is **81.00%.%.** The plot of the True Positive Rate (TPR) against the False Positive Rate (FPR) i.e. ROC curve is given below. The area under the curve of the model is 0.88

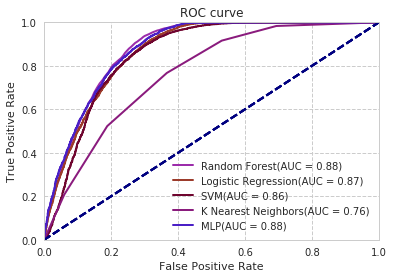
****

**Comparison among different models**

The below table shows accuracy of the different model. Out of 5 models, Random Forest has the highest accuracy (81.91%)

|  |  |
| --- | --- |
| **Model** | **Accuracy** |
| Logistic Regression | 80.72% |
| Random Forest | 81.91% |
| Support Vector Machine | 80.36% |
| K Nearest Neighbours (KNN): | 70.59% |
| Multi-Layer Perceptron | 81.35% |

The area under the curve of all the five models is given blow

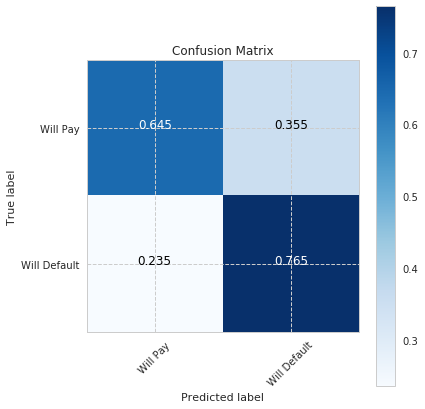
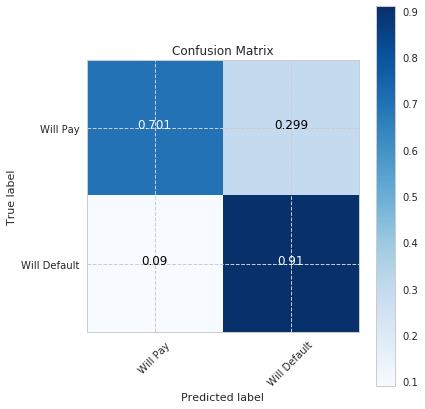
From the below graph, MLP & Random Forest models have highest roc\_auc\_score.****

**Conclusion**

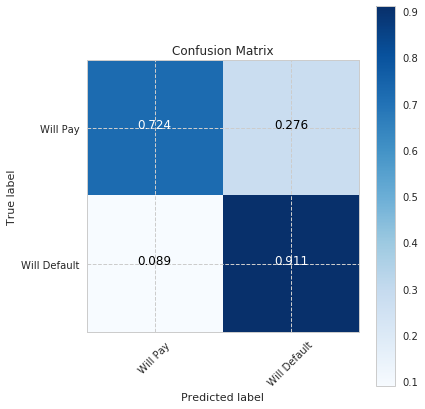
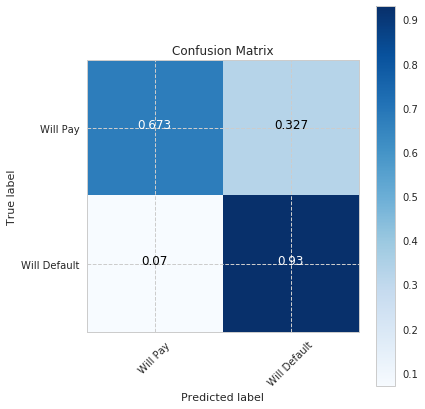
Loan default prediction problem, False Negatives Rate (to predict non-defaulter, when the customer is actually defaulter) is the best metric to evaluate the model. Lower the number of false negatives, better the model is. False negative is when model predicting “a borrower will not default a loan even though he will “. So higher False Negatives will lead to negative impact. So, we evaluated our models using the number of False negatives and accuracies

From the below confusion matrix for all the five models, we can infer that **Random Forest has least False Negative Rate (0.276).**

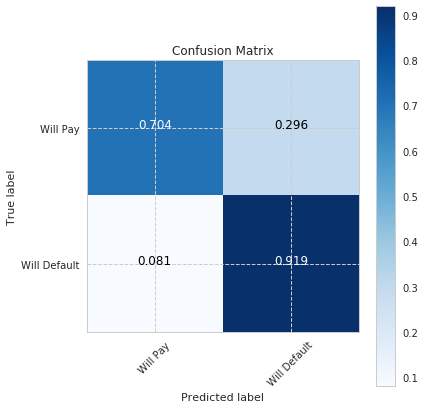
**KNN Confusion Matrix Logistic Regression Confusion Matrix**

** **

**Random Forest Confusion Matrix SVM Confusion Matrix**

 ****

**MLP Confusion Matrix**

****

**Precision, recall, F1 score for all models**

**Logistic Regression**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| **0** | **0.88** | **0.70** | **0.80** | **4680** |
| **1** | **0.75** | **0.90** | **0.83** | **4620** |
| **avg / total** | **0.81** | **00.82** | **0.81** | **9300** |

**Random Forest**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| **0** | **0.89** | **0.73** | **0.80** | **4680** |
| **1** | **0.77** | **0.90** | **0.83** | **4620** |
| **avg / total** | **0.83** | **0.82** | **0.81** | **9300** |

**Support Vector Machine**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| **0** | **0.90** | **0.66** | **0.76** | **4680** |
| **1** | **0.73** | **0.92** | **0.82** | **4620** |
| **avg / total** | **0.82** | **0.79** | **0.79** | **9300** |

K Nearest Neighbour

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| **0** | **0.73** | **0.63** | **0.68** | **4680** |
| **1** | **0.67** | **0.77** | **0.72** | **4620** |
| **avg / total** | **0.70** | **0.70** | **0.70** | **9300** |

**MLP**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| **0** | **0.89** | **0.70** | **0.78** | **4680** |
| **1** | **0.75** | **0.91** | **0.82** | **4620** |
| **avg / total** | **0.82** | **0.81** | **0.81** | **9300** |

**Limitations:**

**ANNEXTURE**

1. Code
   1. Assignment\_TA17002\_RAJIB\_MANDAL.ipynb
   2. Assignment\_TA17002\_RAJIB\_MANDAL.py
2. Data File (LoanStats3c\_final.csv)

**APPENDIX -A**

The definitions for all the data attributes given below.

|  |  |  |
| --- | --- | --- |
| **#** | **Name** | **Description** |
| 1 | id | A unique LC assigned ID for the loan listing. |
| 2 | member\_id | Maximum current balance owed on all revolving accounts |
| 3 | loan\_amnt | The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value. |
| 4\* | funded\_amnt | The total amount committed to that loan at that point in time. |
| 5 | funded\_amnt\_inv | The total amount committed by investors for that loan at that point in time. |
| 6 | term | The number of payments on the loan. Values are in months and can be either 36 or 60. |
| 7 | int\_rate | Interest Rate on the loan |
| 8 | installment | The monthly payment owed by the borrower if the loan originates. |
| 9 | grade | LC assigned loan grade |
| 10 | sub\_grade | LC assigned loan subgrade |
| 11 | emp\_title | The job title supplied by the Borrower when applying for the loan. |
| 12 | emp\_length | Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years. |
| 13 | home\_ownership | The home ownership status provided by the borrower during registration or obtained from the credit report. Our values are: RENT, OWN, MORTGAGE, OTHER |
| 14 | annual\_inc | The self-reported annual income provided by the borrower during registration. |
| 15 | verification\_status | Indicates if income was verified by LC, not verified, or if the income source was verified |
| 16 | issue\_d | The month which the loan was funded. |
| 17 | loan\_status | Current status of the loan |
| 18 | pymnt\_plan | Indicates if a payment plan has been put in place for the loan |
| 19 | url | URL for the LC page with listing data. |
| 20 | desc | Loan description provided by the borrower |
| 21 | purpose | A category provided by the borrower for the loan request. |
| 22 | title | The loan title provided by the borrower |
| 23 | zip\_code | The first 3 numbers of the zip code provided by the borrower in the loan application. |
| 24 | addr\_state | The state provided by the borrower in the loan application |
| 25 | dti | A ratio calculated using the borrower’s total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower’s self-reported monthly income. |
| 26 | delinq\_2yrs | The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years |
| 27 | earliest\_cr\_line | The month the borrower's earliest reported credit line was opened |
| 28 | inq\_last\_6mths | Number of credit inquiries in past 6 months |
| 29 | mths\_since\_last\_delinq | The number of months since the borrower's last delinquency. |
| 30 | mths\_since\_last\_record | The number of months since the last public record. |
| 31 | open\_acc | The number of open credit lines in the borrower's credit file. |
| 32 | pub\_rec | Number of derogatory public records |
| 33 | revol\_bal | Total credit revolving balance |
| 34 | revol\_util | Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit. |
| 35 | total\_acc | Total collection amounts ever owed |
| 36 | initial\_list\_status | The initial listing status of the loan. Possible values are – W, F |
| 37 | out\_prncp | Remaining outstanding principal for total amount funded |
| 38 | out\_prncp\_inv | Remaining outstanding principal for portion of total amount funded by investors |
| 39 | total\_pymnt | Payments received to date for total amount funded |
| 40 | total\_pymnt\_inv | Payments received to date for portion of total amount funded by investors |
| 41 | total\_rec\_prncp | Principal received to date |
| 42 | total\_rec\_int | Interest received to date |
| 43 | total\_rec\_late\_fee | Late fees received to date |
| 44 | recoveries | post charge off gross recovery |
| 45 | collection\_recovery\_fee | post charge off collection fee |
| 46 | last\_pymnt\_d | Last month payment was received |
| 47 | last\_pymnt\_amnt | Last total payment amount received |
| 48 | next\_pymnt\_d | Next scheduled payment date |
| 49 | last\_credit\_pull\_d | The most recent month LC pulled credit for this loan |
| 50 | collections\_12\_mths\_ex\_med | Number of collections in 12 months excluding medical collections |
| 51 | mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating |
| 52 | policy\_code | publicly available policy\_code=1  new products not publicly available policy\_code=2 |
| 53 | application\_type | Indicates whether the loan is an individual application or a joint application with two co-borrowers |
| 54 | annual\_inc\_joint | The combined self-reported annual income provided by the co-borrowers during registration |
| 55 | dti\_joint | A ratio calculated using the co-borrower’s total monthly payments on the total debt obligations, excluding mortgages and the requested LC loan, divided by the co-borrowers' combined self-reported monthly income |
| 56 | verification\_status\_joint | Indicates if the co-borrower’s joint income was verified by LC, not verified, or if the income source was verified |
| 57 | acc\_now\_delinq | The number of accounts on which the borrower is now delinquent. |
| 58 | tot\_coll\_amt | Total collection amounts ever owed |
| 59 | tot\_cur\_bal | Total current balance of all accounts |
| 60 | open\_acc\_6m | Number of open trades in last 6 months |
| 61 | open\_act\_il | Number of currently active instalment trades |
| 62 | open\_il\_12m | Number of instalment accounts opened in past 12 months |
| 63 | open\_il\_24m | Number of instalment accounts opened in past 24 months |
| 64 | mths\_since\_rcnt\_il | Months since most recent instalment accounts opened |
| 65 | total\_bal\_il | Total current balance of all instalment accounts |
| 66 | il\_util | Ratio of total current balance to high credit/credit limit on all install acct |
| 67 | open\_rv\_12m | Number of revolving trades opened in past 12 months |
| 68 | open\_rv\_24m | Number of revolving trades opened in past 24 months |
| 69 | max\_bal\_bc | Maximum current balance owed on all revolving accounts |
| 70 | all\_util | Balance to credit limit on all trades |
| 71 | total\_rev\_hi\_lim | Total revolving high credit/credit limit |
| 72 | inq\_fi | Number of personal finance inquiries |
| 73 | total\_cu\_tl | Number of finance trades |
| 74 | inq\_last\_12m | Number of credit inquiries in past 12 months |
| 75 | acc\_open\_past\_24mths | Number of trades opened in past 24 months. |
| 76 | avg\_cur\_bal | Average current balance of all accounts |
| 77 | bc\_open\_to\_buy | Total open to buy on revolving bankcards. |
| 78 | bc\_util | Ratio of total current balance to high credit/credit limit for all bankcard accounts. |
| 79 | chargeoff\_within\_12\_mths | Number of charge-offs within 12 months |
| 80 | delinq\_amnt | The past-due amount owed for the accounts on which the borrower is now delinquent. |
| 81 | mo\_sin\_old\_il\_acct | Months since oldest bank installment account opened |
| 82 | mo\_sin\_old\_rev\_tl\_op | Months since oldest revolving account opened |
| 83 | mo\_sin\_rcnt\_rev\_tl\_op | Months since most recent revolving account opened |
| 84 | mo\_sin\_rcnt\_tl | Months since most recent account opened |
| 85 | mort\_acc | Number of mortgage accounts. |
| 86 | mths\_since\_recent\_bc | Months since most recent bankcard account opened. |
| 87 | mths\_since\_recent\_bc\_dlq | Months since most recent bankcard delinquency |
| 88 | mths\_since\_recent\_inq | Months since most recent inquiry. |
| 89 | mths\_since\_recent\_revol\_delinq | Months since most recent revolving delinquency. |
| 90 | num\_accts\_ever\_120\_pd | Number of accounts ever 120 or more days past due |
| 91 | num\_actv\_bc\_tl | Number of currently active bankcard accounts |
| 92 | num\_actv\_rev\_tl | Number of currently active revolving trades |
| 93 | num\_bc\_sats | Number of satisfactory bankcard accounts |
| 94 | num\_bc\_tl | No of bankcard accounts |
| 95 | num\_il\_tl | Number of installment accounts |
| 96 | num\_op\_rev\_tl | Number of open revolving accounts |
| 97 | num\_rev\_accts | Number of open revolving accounts |
| 98 | num\_rev\_tl\_bal\_gt\_0 | Number of revolving trades with balance >0 |
| 99 | num\_sats | Number of satisfactory accounts |
| 100 | num\_tl\_120dpd\_2m | Number of accounts currently 120 days past due (updated in past 2 months) |
| 101 | num\_tl\_30dpd | Number of accounts currently 30 days past due (updated in past 2 months) |
| 102 | num\_tl\_90g\_dpd\_24m | Number of accounts 90 or more days past due in last 24 months |
| 103 | num\_tl\_op\_past\_12m | Number of accounts opened in past 12 months |
| 104 | pct\_tl\_nvr\_dlq | Percent of trades never delinquent |
| 105 | percent\_bc\_gt\_75 | Percentage of all bankcard accounts > 75% of limit. |
| 106 | pub\_rec\_bankruptcies | Number of public record bankruptcies |
| 107 | tax\_liens | Number of tax liens |
| 108 | tot\_hi\_cred\_lim | Total high credit limit |
| 109 | total\_bal\_ex\_mort | Total credit balance excluding mortgage |
| 110 | total\_bc\_limit | Total bankcard high credit limit |
| 111 | total\_il\_high\_credit\_limit | Total instalment high credit limit |
| 112 | revol\_bal\_joint | Sum of revolving credit balance of the co-borrowers, net of duplicate balances |
| 113 | sec\_app\_earliest\_cr\_line | Earliest credit line at time of application for the secondary applicant |
| 114 | sec\_app\_inq\_last\_6mths | Credit inquiries in the last 6 months at time of application for the secondary applicant |
| 115 | sec\_app\_mort\_acc | Number of mortgage accounts at time of application for the secondary applicant |
| 116 | sec\_app\_open\_acc | Number of open trades at time of application for the secondary applicant |
| 117 | sec\_app\_revol\_util | Ratio of total current balance to high credit/credit limit for all revolving accounts |
| 118 | sec\_app\_open\_act\_il | Number of currently active instalment trades at time of application for the secondary applicant |
| 119 | sec\_app\_num\_rev\_accts | Number of revolving accounts at time of application for the secondary applicant |
| 120 | sec\_app\_chargeoff\_within\_12\_mths | Number of charge-offs within last 12 months at time of application for the secondary applicant |
| 121 | sec\_app\_collections\_12\_mths\_ex\_med | Number of collections within last 12 months excluding medical collections at time of application for the secondary applicant |
| 122 | sec\_app\_mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating at time of application for the secondary applicant |
| 123 | hardship\_flag | Flags whether or not the borrower is on a hardship plan |
| 124 | hardship\_type | Describes the hardship plan offering |
| 125 | hardship\_reason | Describes the reason the hardship plan was offered |
| 126 | hardship\_status | Describes if the hardship plan is active, pending, cancelled, completed, or broken |
| 127 | deferral\_term | Amount of months that the borrower is expected to pay less than the contractual monthly payment amount due to a hardship plan |
| 128 | hardship\_amount | The interest payment that the borrower has committed to make each month while they are on a hardship plan |
| 129 | hardship\_start\_date | The start date of the hardship plan period |
| 130 | hardship\_end\_date | The end date of the hardship plan period |
| 131 | payment\_plan\_start\_date | The day the first hardship plan payment is due. For example, if a borrower has a hardship plan period of 3 months, the start date is the start of the three-month period in which the borrower is allowed to make interest-only payments. |
| 132 | hardship\_length | The number of months the borrower will make smaller payments than normally obligated due to a hardship plan |
| 133 | hardship\_dpd | Account days past due as of the hardship plan start date |
| 134 | hardship\_loan\_status | Loan Status as of the hardship plan start date |
| 135 | orig\_projected\_additional\_accrued\_interest | The original projected additional interest amount that will accrue for the given hardship payment plan as of the Hardship Start Date. This field will be null if the borrower has broken their hardship payment plan. |
| 136 | hardship\_payoff\_balance\_amount | The payoff balance amount as of the hardship plan start date |
| 137 | hardship\_last\_payment\_amount | The last payment amount as of the hardship plan start date |
| 138 | disbursement\_method | The method by which the borrower receives their loan. Possible values are: CASH, DIRECT\_PAY |
| 139 | debt\_settlement\_flag | Flags whether or not the borrower, who has charged-off, is working with a debt-settlement company. |
| 140 | debt\_settlement\_flag\_date | The most recent date that the Debt\_Settlement\_Flag has been set |
| 141 | settlement\_status | The status of the borrower’s settlement plan. Possible values are: COMPLETE, ACTIVE, BROKEN, CANCELLED, DENIED, DRAFT |
| 142 | settlement\_date | The date that the borrower agrees to the settlement plan |
| 143 | settlement\_amount | The loan amount that the borrower has agreed to settle for |
| 144 | settlement\_percentage | The settlement amount as a percentage of the payoff balance amount on the loan |
| 145 | settlement\_term | The number of months that the borrower will be on the settlement plan |

**APPENDIX -B**

The following 31 fields do not have any value in the dataset.

|  |  |
| --- | --- |
| SR# | Name |
| 1 | id |
| 2 | member\_id |
| 3 | url |
| 4 | annual\_inc\_joint |
| 5 | dti\_joint |
| 6 | verification\_status\_joint |
| 7 | open\_acc\_6m |
| 8 | open\_act\_il |
| 9 | open\_il\_12m |
| 10 | open\_il\_24m |
| 11 | mths\_since\_rcnt\_il |
| 12 | total\_bal\_il |
| 13 | il\_util |
| 14 | open\_rv\_12m |
| 15 | open\_rv\_24m |
| 16 | max\_bal\_bc |
| 17 | all\_util |
| 18 | inq\_fi |
| 19 | total\_cu\_tl |
| 20 | inq\_last\_12m |
| 21 | revol\_bal\_joint |
| 22 | sec\_app\_earliest\_cr\_line |
| 23 | sec\_app\_inq\_last\_6mths |
| 24 | sec\_app\_mort\_acc |
| 25 | sec\_app\_open\_acc |
| 26 | sec\_app\_revol\_util |
| 27 | sec\_app\_open\_act\_il |
| 28 | sec\_app\_num\_rev\_accts |
| 29 | sec\_app\_chargeoff\_within\_12\_mths |
| 30 | sec\_app\_collections\_12\_mths\_ex\_med |
| 31 | sec\_app\_mths\_since\_last\_major\_derog |