|  |  |
| --- | --- |
|  | description |
| collection\_recovery\_fee | post charge off collection fee |
| last\_pymnt\_amnt | Last total payment amount received |
| grade | LC assigned loan grade |
| debt\_settlement\_flag | Flags whether or not the borrower, who has charged-off, is working with a debt-settlement company. |
| Loan status (response) | Current status of the loan |

Table 1: Logistic Regression Attributes Description

|  |  |  |
| --- | --- | --- |
| Threshold | Train TP Rate | Train Accuracy rate |
| 0.5 | 68.49% | 95.30% |
| 0.3 | 68.93% | 95.08% |
| 0.2 | 71.63% | 93.95% |
| 0.1 | 86.88% | 81.90% |
| 0.08 | 91.76% | 76.15% |

Table 2: TP Rates with Different Thresholds

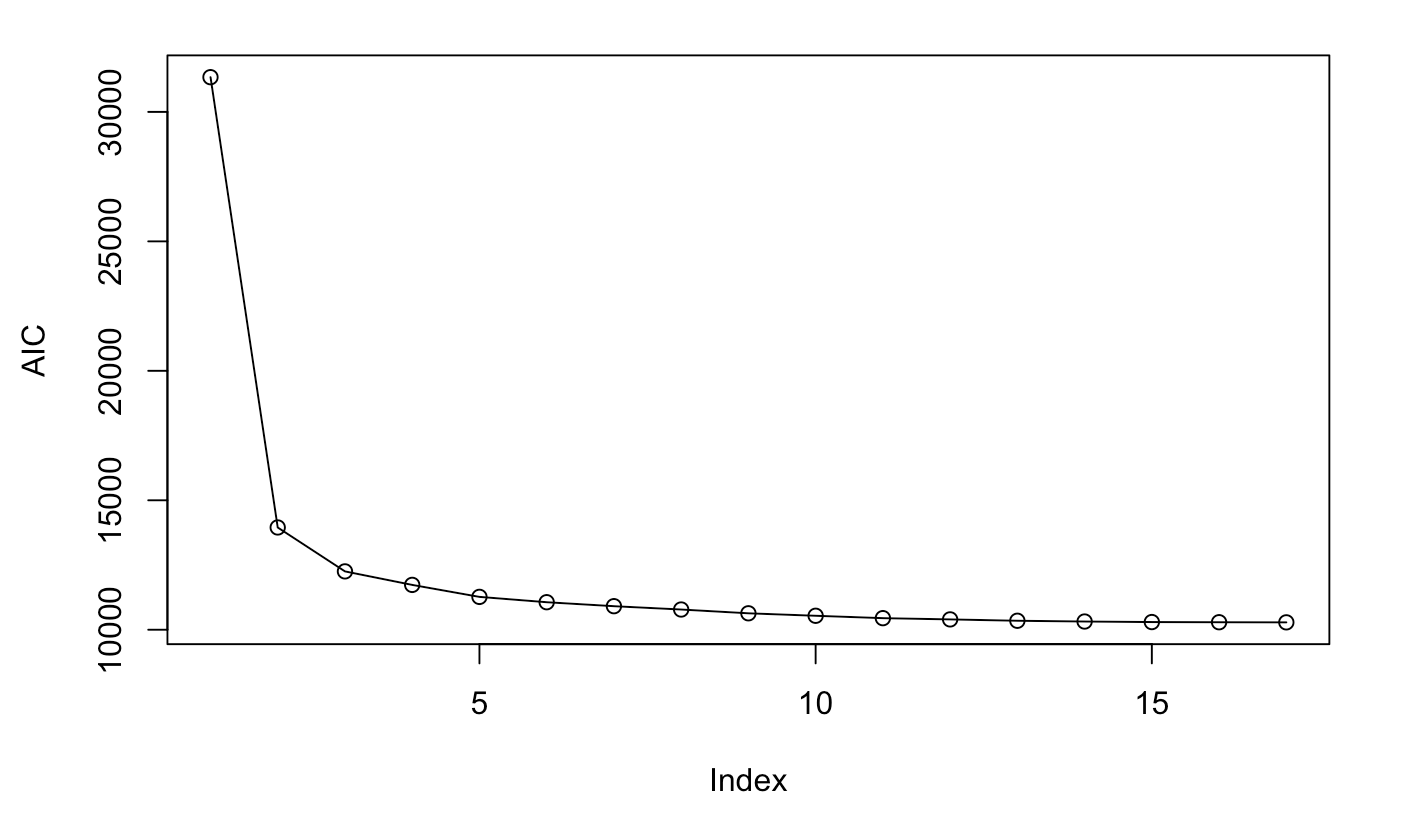


Fig 1: AIC Logistic Regression Plot



**Fig 2: CV Plot of Ridge Regression**

|  |  |
| --- | --- |
|  | description |
| Total payment (response) | Payments received to date for total amount funded |
| Funded amount | The total amount committed to that loan at that point in time. |
| Term | The number of payments on the loan. Values are in months and can be either 36 or 60. |
| Installment | The monthly payment owed by the borrower if the loan originates. |
| Annual interest | The self-reported annual income provided by the borrower during registration. |
| Purpose | A category provided by the borrower for the loan request. |
| Inquiries | The number of inquiries in past 6 months (excluding auto and mortgage inquiries) |
| Recovery fee | post charge off collection fee |
| Debt settlement flag | The most recent date that the Debt\_Settlement\_Flag has been set |

Table 3: Linear Regression Attribution Description



Fig 3: Feature Selection Plots



Fig 4: Principle Components



Fig 5: Residuals vs Fitted Values



Fig 6: BIC Plot

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | predictors | Df | SSE | MSE | F value | P value |
|  |  | 1 | 3281752 | 3281752 | 5497.6001 |  |
|  |  | 1 | 182450 | 182450 | 305.6409 |  |
|  |  | 1 | 445478 | 445478 | 746.2654 |  |
|  |  | 1 | 73852 | 73852 | 123.7174 |  |
|  |  | 1 | 6984 | 6984 | 11.6387 |  |
|  |  | 1 | 33505 | 33505 | 56.1279 |  |
|  |  | 1 | 99814 | 99814 | 167.2090 |  |
|  |  | 1 | 9378 | 9378 | 15.7095 |  |
|  |  | 1 | 9246 | 9246 | 15.4887 |  |
|  |  | 1 | 4167 | 4167 | 6.9801 |  |
|  |  | 1 | 9517 | 9517 | 15.9436 |  |
|  |  | 1 | 23800 | 23800 | 39.8697 |  |
|  | 5539 | 5539 | 3306466 | 597 |  |  |

Table 4: Analysis of Variance Table

|  |  |  |  |
| --- | --- | --- | --- |
|  | Adjusted | lillie.test p-val | bptest p-val |
| Do not delete influential cases. | 0.5338 | 0.03931 | < 2.2e-16 |
| Delete influential cases. | 0.5713 | 0.0002574 | < 2.2e-16 |

Table 5: Influential Cases

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | true | |
|  |  | Fully paid | Charged off |
| predict | Fully paid | 3596 | 105 |
| Charged off | 796 | 693 |

Table 6: Contagious Table of Test Set

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Weight | Training MSE | Test MSE | lillie.test p-val | bptest p-val |
|  | 596.512 | 557.0247 | 6.14e-05 | < 2.2e-16 |
|  | 628.7692 | 607.5463 | 0.03527 | < 2.2e-16 |
|  | 640.0233 | 620.44 | 7.692e-07 | < 2.2e-16 |
|  | 628.8545 | 607.5413 | 0.02514 | < 2.2e-16 |
|  | 635.0558 | 610.283 | 7.692e-07 | < 2.2e-16 |
|  | 635.0303 | 610.2441 | 5.782e-05 | < 2.2e-16 |
| notes |  | | | |

Table 7: Weighted Least Squares

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Train MSE | Test MSE | lillie.test p-val | | | bptest p-val | | |
| robust regression psi=psi.bisquare | 630.8849 | 608.7783 | 0.0003873 | | < 2.2e-16 | | |
| robust regression psi=psi.hampel | 628.3365 | 607.4484 | 0.008273 | < 2.2e-16 | | |
| robust regression psi=psi.huber | 629.5456 | 607.9952 | 0.001817 | < 2.2e-16 | | |

Table 8: Robust Regression