BFS Capstone Project

Abstract

This report is aimed at addressing the increasing credit loss problem being faced by CredX, a leading credit card provider.

Following are the highlights of the report:

1. Assumptions
2. Exploratory Data Analysis - Takeaways
3. IV Analysis – Best Attributes
4. Model Building - Outcomes
5. Application Scorecard
6. Financial Benefit

Assumptions

Following are the set of assumptions made in the course of this analysis:

Policy and Data related-

- ✓ Minimum employability age is 18yrs, and hence the minimum age of holding a Credit card.
- ✓ Customers with *Education* as *Others* are below Graduation/Bachelor degree holders.

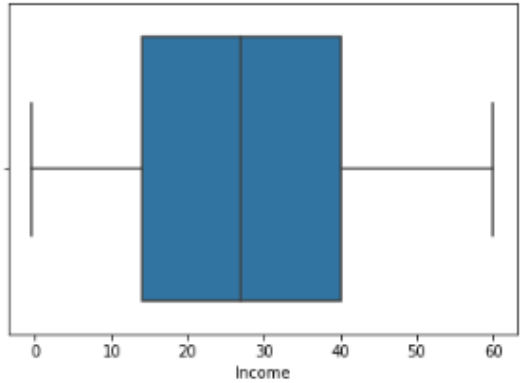
Finance related-

- ✓ On average 30% of Outstanding amount per defaulter is recovered by Collections and recovery department.
- ✓ On average 8% of Outstanding amount per customer making late payments is collected as Interest income by the company.

Exploratory Data Analysis - Takeaways

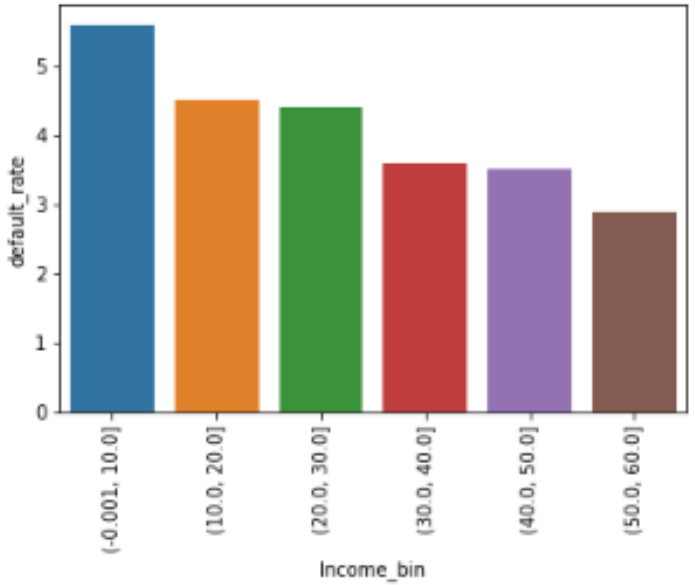
- Income

The Income of the customers majorly between 14 thousands and 40 thousands, and is averaging at about 28 thousands.



While, in terms of predictability of the default rate this variable is definitely intuitive and shows obvious results whereby demonstrating the inverse relationship with the default-rate.

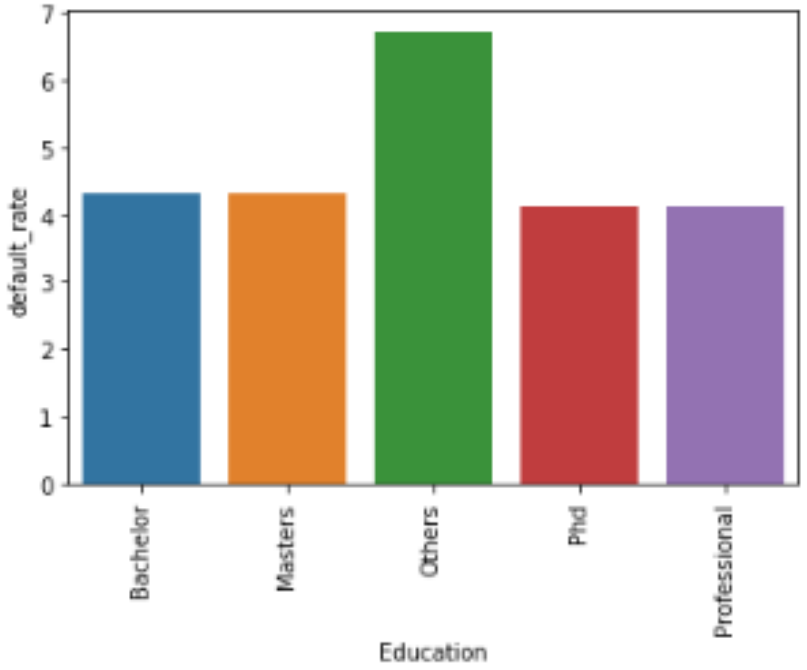
In simple words, *More the income of a candidate, less likely he is to default*.



Exploratory Data Analysis - Takeaways

- Education

Education variable having a great variance in the different classes, also looks intuitive in terms of predictability of the default rate, the key insight here is that *people with below Graduation degree tend to default at a much higher rate than the Graduates and above*.



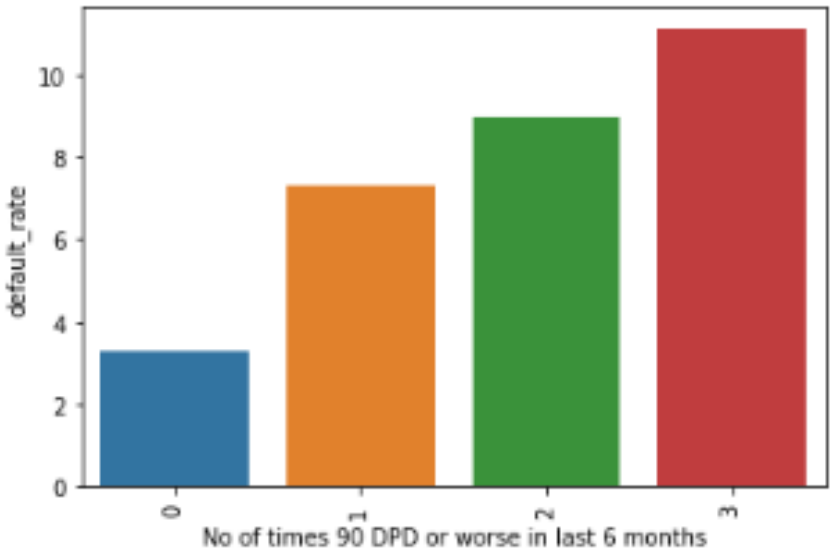
Professional	24839
Masters	23970
Bachelor	17697
Phd	4549
Others	121

Exploratory Data Analysis - Takeaways

- No of times 90 DPD or worse in last 6 months

This attribute have some really intuitive stats to show in terms of predictability of the default rate, the frequency of 90 days past due or worse in last 6 months is directly proportional to the default rate.

So, *higher the frequency of 90 days past due, more the chances of default*.

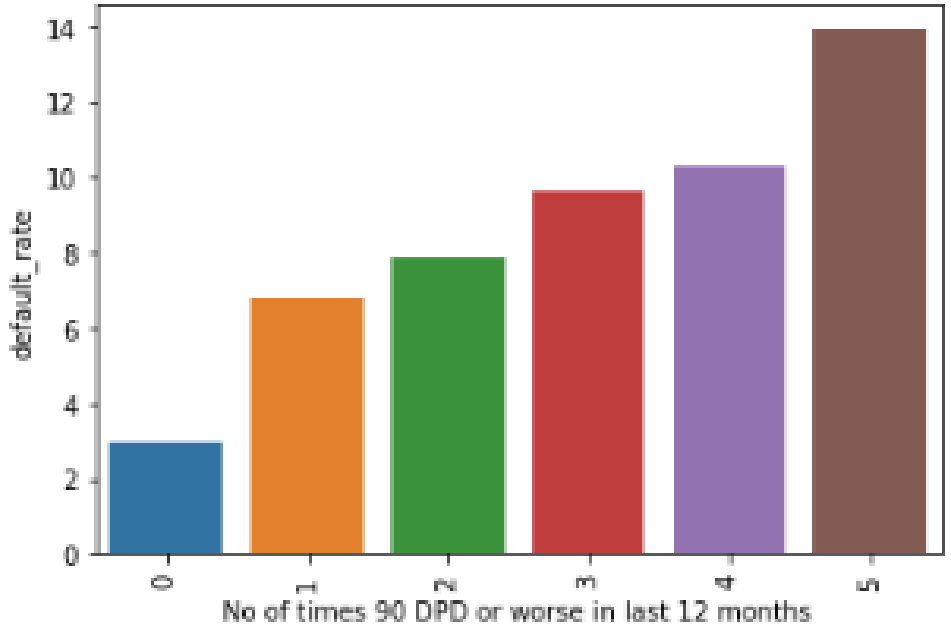


Exploratory Data Analysis - Takeaways

- No of times 90 DPD or worse in last 12 months

Again the variable shows the same signs like the previous one, the frequency of 90 days past due or worse in last 12 months is directly proportional to the default rate.

Higher the frequency of 90 days past due, higher the probability of default.

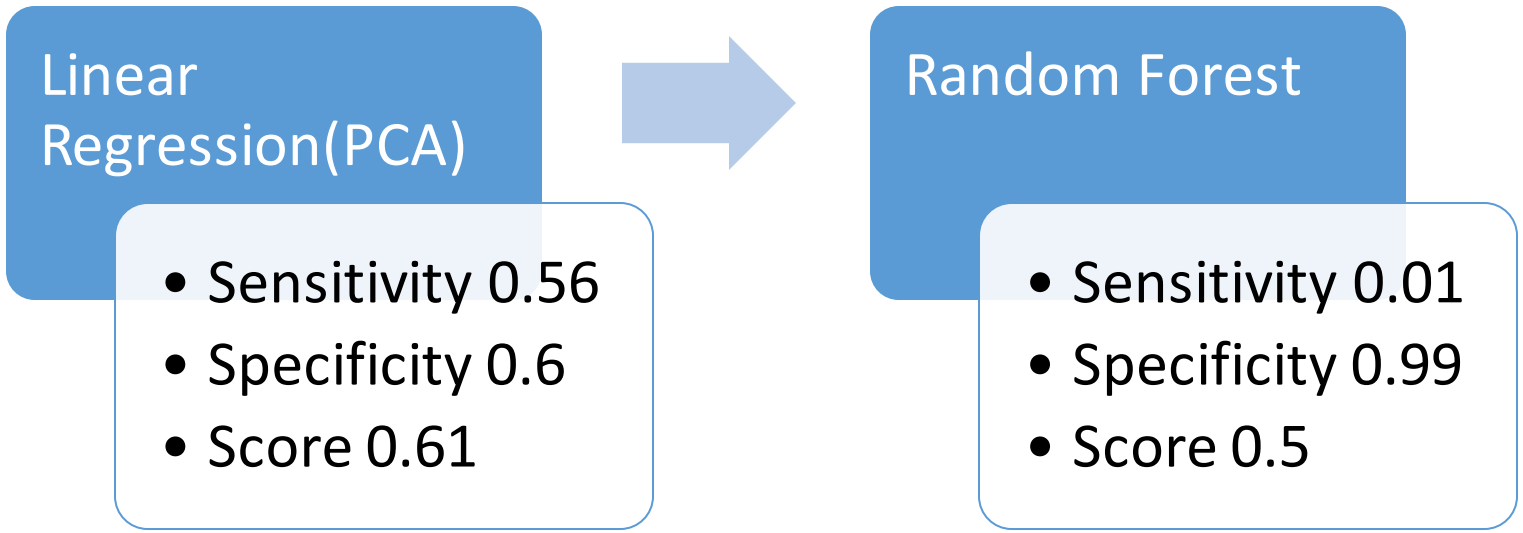


IV Analysis – Best Attributes

According to our IV analysis, following is the categorization of the attributes based on their Predictability.

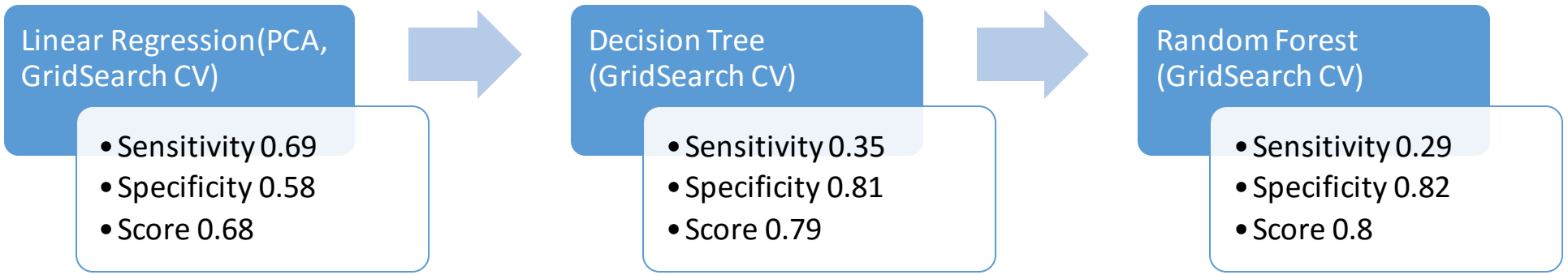
Predictive Power	Attributes
Useless	*Marital Status, Gender, Education, Residence Type, Open Autoloan, Profession, Dependents, Age, Open Homeloan*
Weak	*Time in Current Company, Income, Time in Current Residence*
Medium	*Inquiries in 6 months, 90DPD in 6months, Trades opened in 6months, 60DPD in 12months, 60DPD in 6months, 90DPD in 12months, 30DPD in 12months, PL trades opened in 6months, Total trades, Outstanding balance, 30DPD in 6months, Inquiries in 12months, PL trades opened in 12months, Trades opened in 12months*
Strong	Average card utilization in 12months

Model Building - Outcomes (Demographic Data Only)



As we can see, the demographic data alone is not good enough to drive predictions.

Model Building - Outcomes
(Demographic & Credit Bureau Data combined)



So the best model, considering the best combination of performance stats is Random Forest.

Additionally when checked for the rejected candidates, model showed an accuracy of 95%.

Application Scorecard

Parameters	Accepted Population	Rejected Population
Overall Score Range	312 - 385	325 - 340
Defaulters' Score Range	312 - 334	325 - 334
Non-defaulters' Score Range	333 - 385	333 - 340

From the above table of scores distribution based on our final model, below are two major conclusions:

✓ The cutoff score should be 335

✓ Candidates scoring between 332 to 335 should be manually reviewed by Approvals' department.

Financial Benefit

Based on the assumptions mentioned before, following are the findings and key benefits out of our model,

Current situation

- ✓ Average Credit loss per default is Rs.880,044
- ✓ Total interest income is Rs.2,57,74,04,854
- ✓ Total Credit loss to the company is Rs.2,59,34,89,668

Model benefit

- ✓ Total savings by rejecting defaulters is Rs. 79,64,39,820
- ✓ Total loss of interest by rejecting non-defaulters is Rs. 46,75,84,240
- ✓ If we use model for auto approval/rejection we save Rs. 32,88,55,580

Thank You