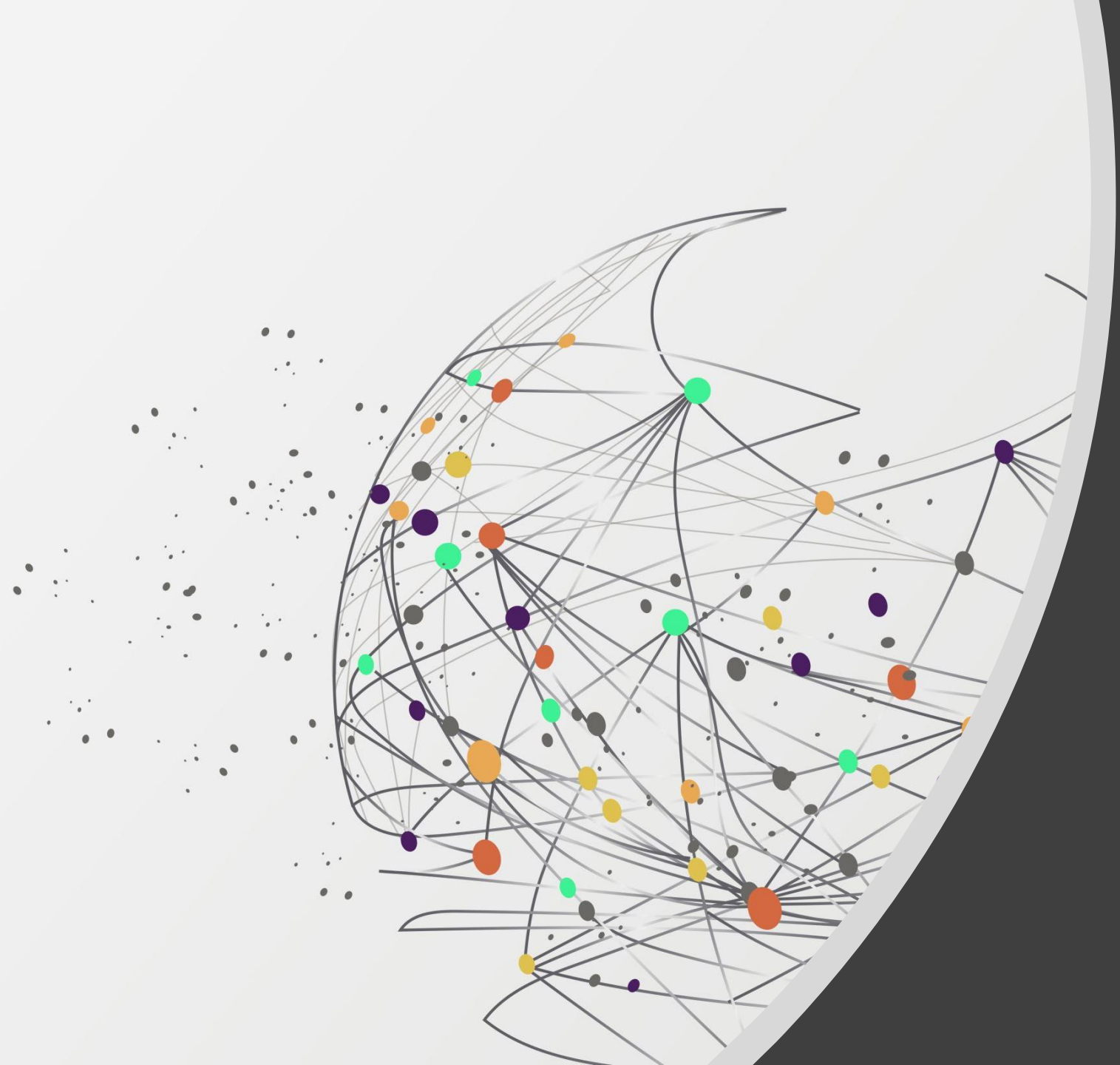


Predicting Credit Card Approval with Machine Learning Algorithms

Ruben Solis

Objective

- Use machine learning methods to develop a prediction model to determine whether a client is approved or disapproved.
- Explore to see which features play a role in determining this decision Financial institutions will find much use of this model.

The Data

Column Name	Data Type
male	object
age	float64
debt	float64
married	object
bank_customer	object
education_level	object
ethnicity	object
years_employed	float64
prior_default	object
employed	object
credit_score	int64
drivers_license	object
citizen	object
zip_code	object
income	float64
approval_status	int64

Approval Status and Application by Gender

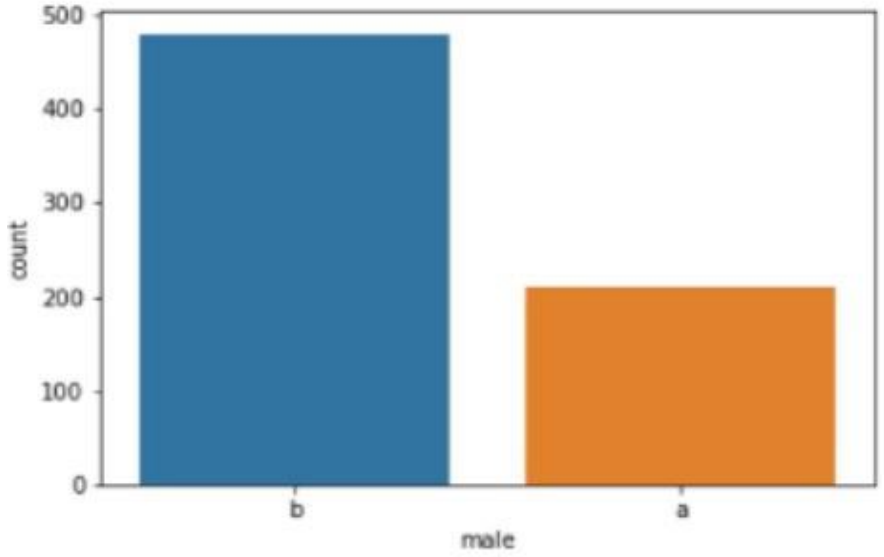


Figure 2: Application Count by Gender

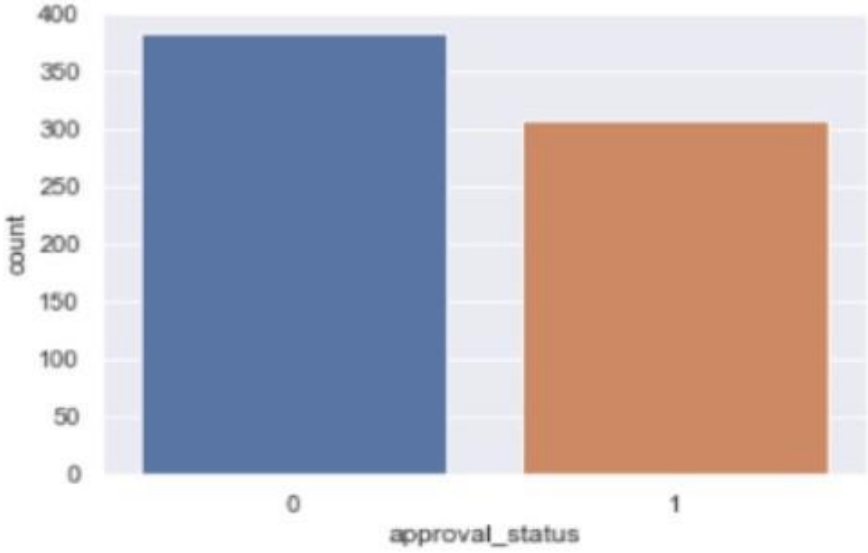


Figure 1: Approval and Disapproval Counts for Applicants

Education and Prior Default for Credit Card Candidates

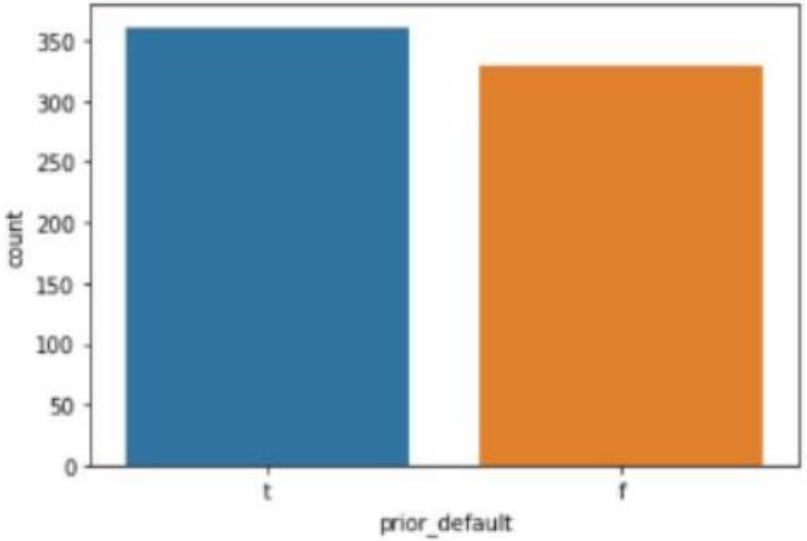


Figure 4: Prior Default of Credit Card Applicants

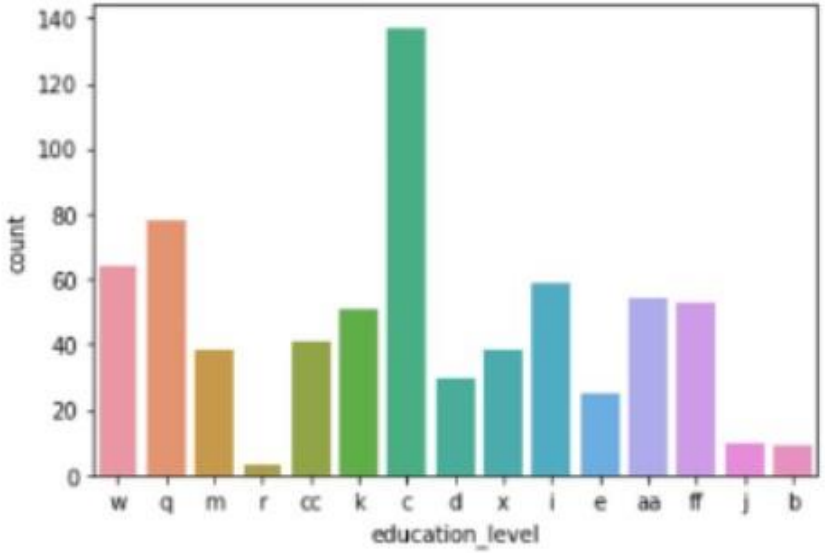


Figure 3: Education Level for Credit Card Applicants

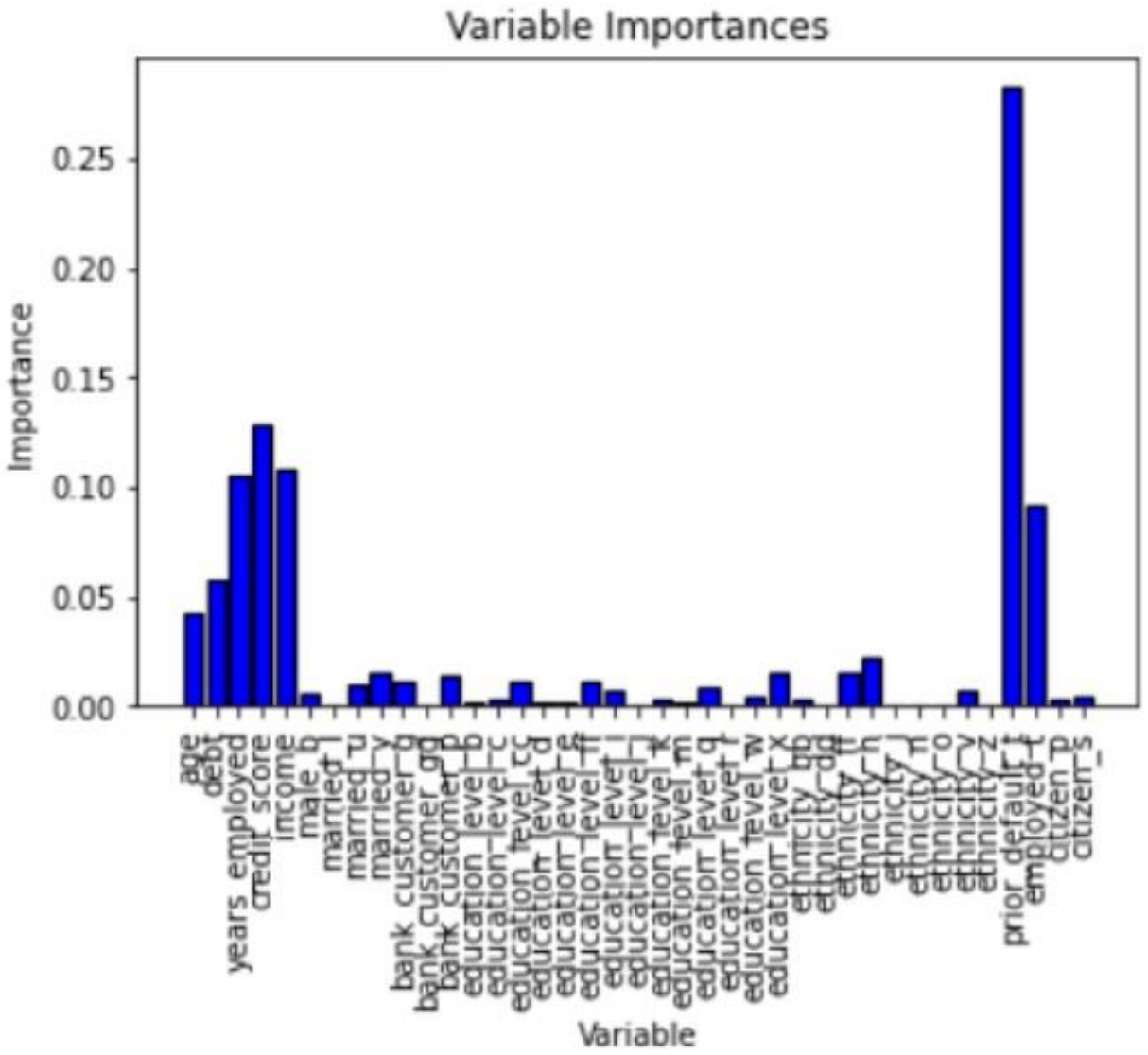
Results

Model	Accuracy	F1	Precision	Recall	Cross-Validation Score
Logistic Regression	0.826087	0.833333	0.810811	0.857143	0.8625
Random Forest	0.847826	0.844444	0.876923	0.814286	0.8659
XGBoost	0.840580	0.840580	0.852941	0.828571	0.8497
KNN	0.811594	0.803030	0.854839	0.757143	0.8660
Decision Tree	0.782609	0.776119	0.812500	0.742857	0.8188

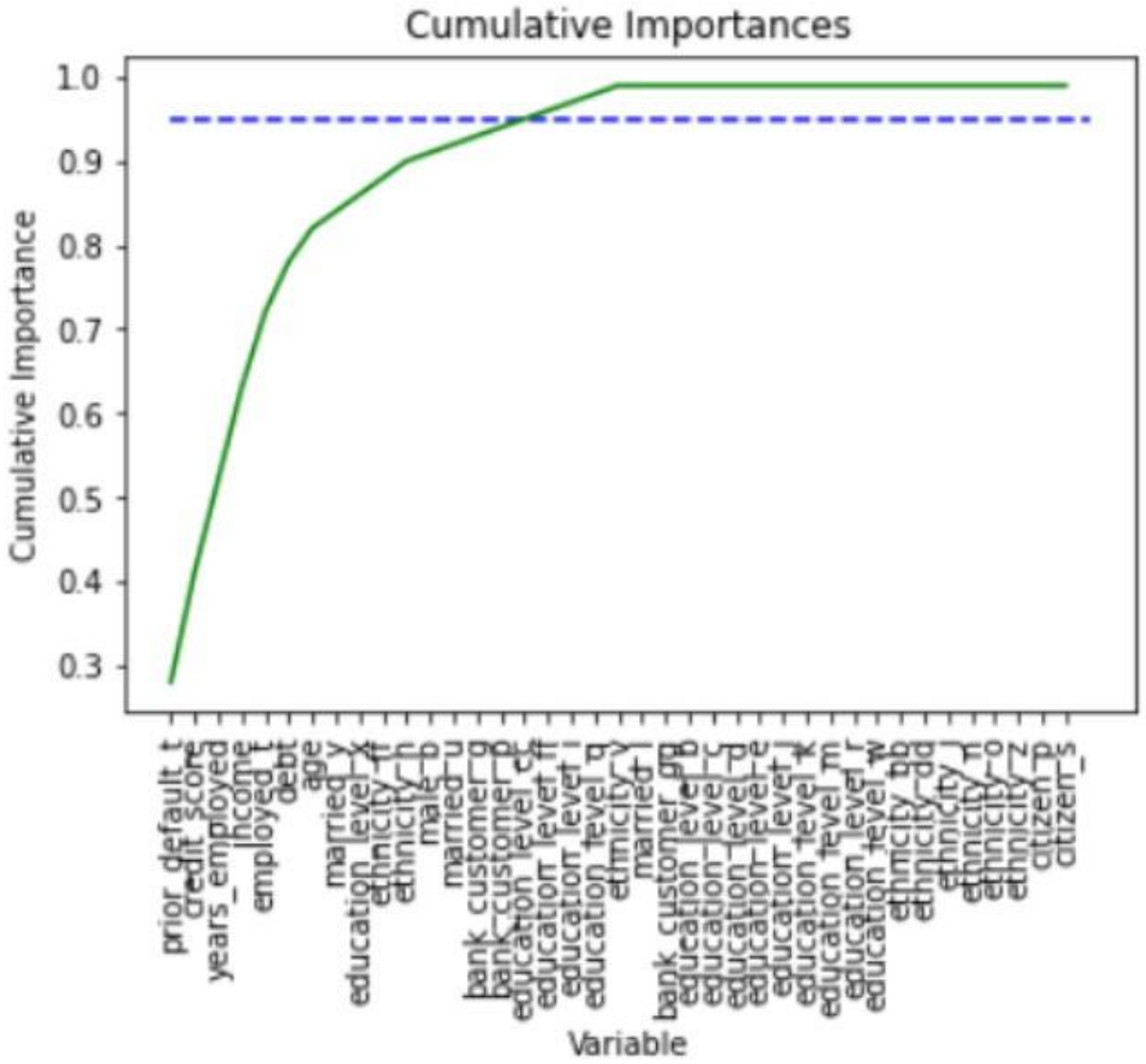
Final Model Results

Model	Accuracy	F1	Precision	Recall
Final Logistic Regression	0.8623	0.8527	0.9322	0.7857

Variable Importance Random Forest



Cumulative Importance



Recommendations

- As a result, it is not recommended that clients who have defaulted in the past be issued credit-cards, unless there are good reasons to do so.
- Credit score, income, and years employed are important when making this decision. These factors impacted the model more than debt, it does not mean that they are necessarily bad candidates for credit cards. The other factors should play a factor as well