Model Card - Credit Prediction

Model Details:

- Developed by Shashank Vankadari, a graduate student at Lehigh University Data Science program.
- This project designs various models and chooses the best binary model that can predict if a person will be able to repay the loan or not based on the historical data on which the model is trained.
- This is a revised version which is developed based on fairness assessments and value audits done on the previously developed model. It now ensures the data is cleaned properly as per various privacy and ethics principles, makes sure the model's fairness is assessed across different groups and thus gives suitable recommendations. Also, the model selection is done considering multiple metrics rather than considering one single metric.
- The Random Forest model is found out to be the best for predicting loan repayment.
- Developer can be contacted at shv222@lehigh.edu

Intended Use:

- Intended to be used by financial institutions to determine person's worthiness of receiving a loan.
- Intended for any person with proper identity (Native to United States or international resident with proper immigration status).
- Not intended to make judgments on any individual.

Metrics:

 The performance metrics such as accuracy, P-value, sensitivity, specificity, balanced accuracy was considered relevant factors in choosing the best performing model and Random Forest model is found to be the best

- and most effective. All metrics were evaluated from the confusion matrix obtained after testing the model. You can find the results of model evaluation under Quantitative analysis section.
- A P-value less than 0.05, sensitivity and specificity close to 1 and balanced accuracy close to 100% is considered.
- For evaluating the fairness of the Random Forest model, we considered Precision, Recall, False positives and made sure the fairness is improved on Race and Gender compared to earlier version.

Ethical considerations:

- The sensitive attributes like race, sex, gender is not considered while model building following privacy and ethics principles. This is done in accordance with Equal Credit Opportunity act which considers sex, race, or country as protected attributes.
- The dataset used has class bias on sensitive attributes and since they were not considered for model training, the model can be less biased.

Training data:

- The credit dataset contains occupation, personal and financial information necessary to predict loan repayment and thus making it more effective in building the models.
- The individual's native country (attribute) is removed from the data set before the data split, as we have improper distribution of records for some countries. Due to this, the test data may have some records which were not present in train data and vice versa. This creates a problem while evaluating the model.
- The historical credit dataset is split into 70% of training data and 30% of testing data as

- this distribution ensures sufficient data points for model training and testing.
- The data split is done after necessary preprocessing (like converting attributes to suitable data types, ordering of categorical data, removal of certain unwanted attributes etc.)

Testing Data:

• 30% of the data in the data split was used to evaluate the performance of the models.

Quantitative analysis:

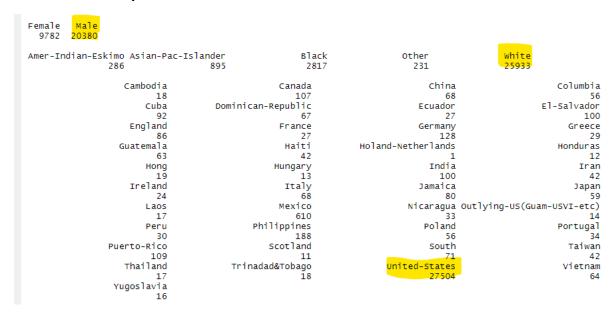


Fig-1: Class bias on sensitive attributes - Sex, Race and Country

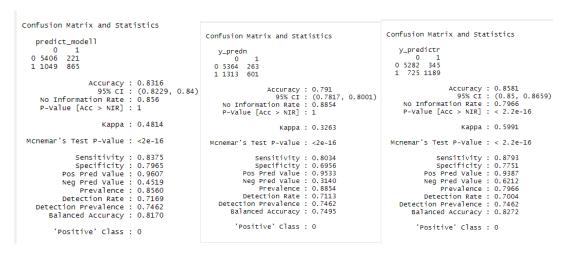


Fig-2: L-R Model

Fig-3: Naïve Bayes model

Fig-4: Random Forest



Fig-5: Fairness assessment based on gender

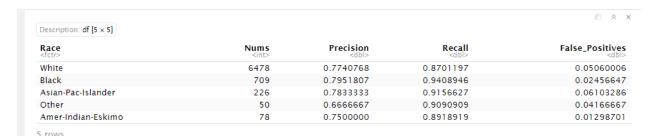


Fig-6: Fairness assessment based on race

Caveats and Recommendations:

- Since there is bias in the data, it cannot be possible to achieve complete fairness. Bias can be reduced at data collection stage by adjusting the sampling process. Once bias is removed, fairness can be improved as per legal/ethical principles of the place.
- In the fairness metrics, the false positives vary across various groups, so there is risk of improperly predicting the outcome based on race / sex. This can raise concerns about fairness of the model.
- We can collect data points in such a way that class bias is low or use advanced techniques such as up sampling / down sampling to reduce bias and improve fairness.