Best credit card fraudulent classifier.

the dataset

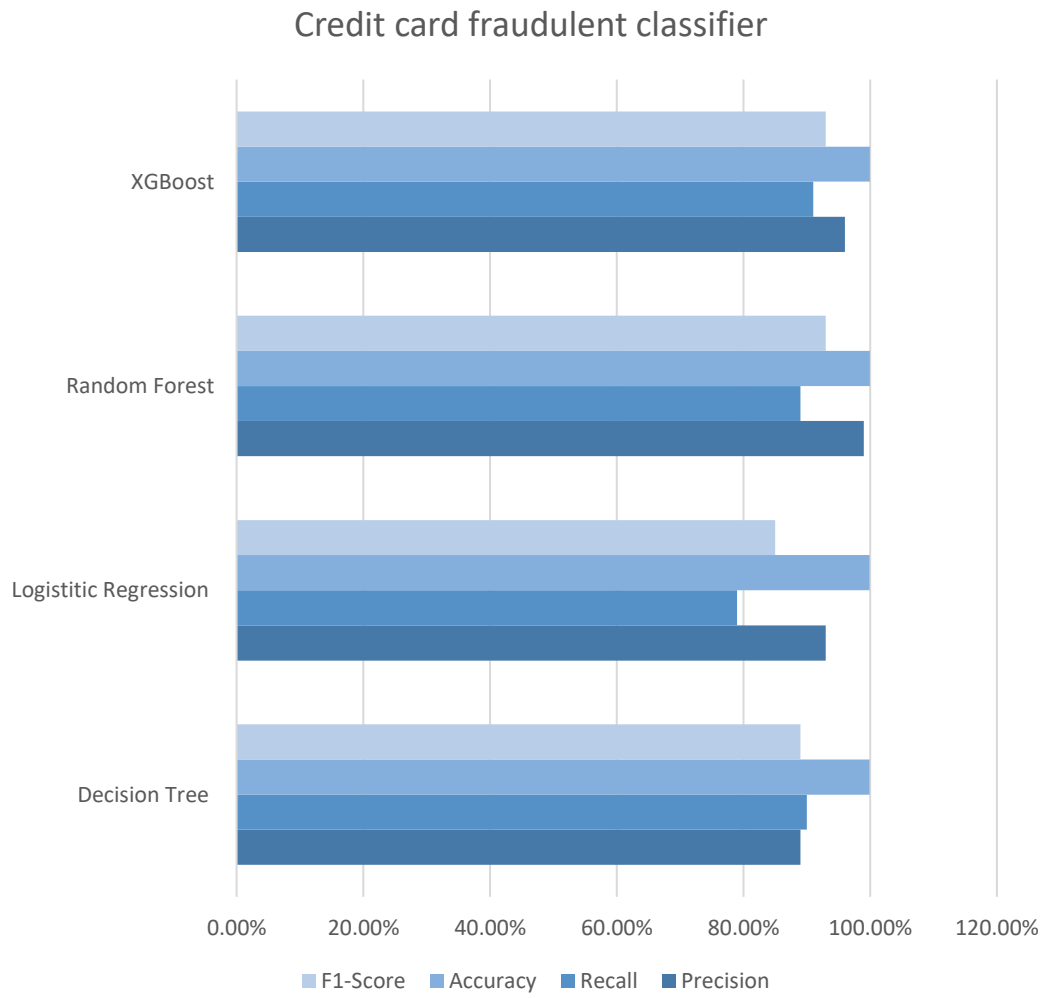
- the dataset https://www.kaggle.com/mlg-ulb/creditcardfraud
- features are obscured (maybe a result to protect user identities and sensitive data(v1-v28))
- time and class heading not obscured
- the majority of the features are constant.
- 31 columns > 284,000 rows @144MB

considered models: logistic regression, decision tree, random forest, xgboost

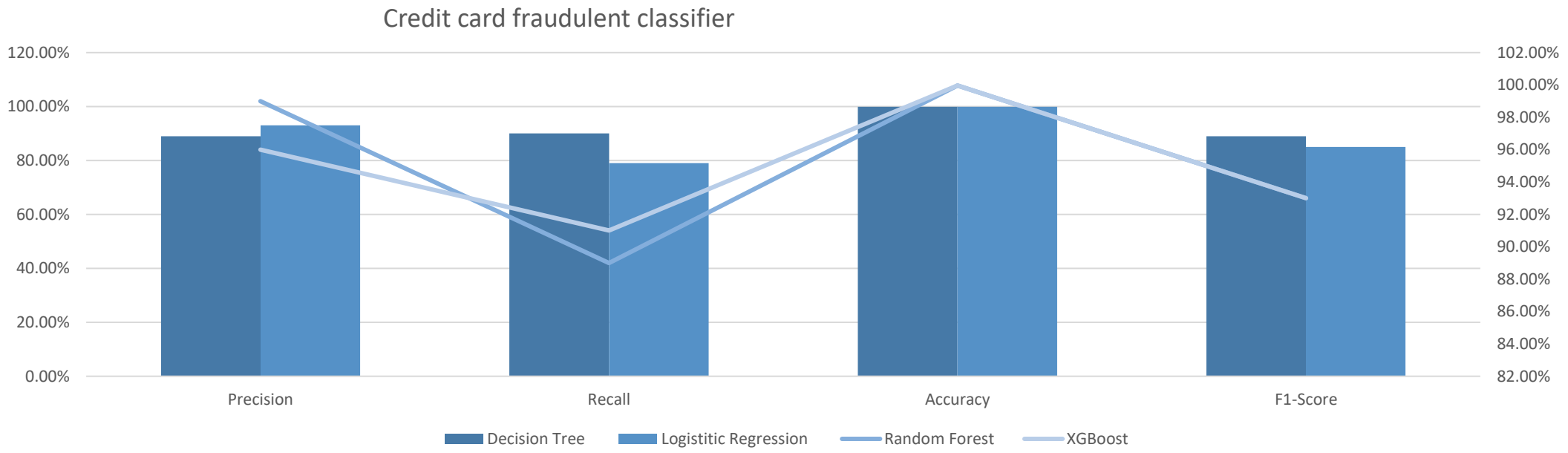
metrics: precision, recall, accuracy, and F1-score

STOPPING FRAUDULENT TRANSACTIONS AND NOT LOSING MONEY!

random forest classifier: number 1 in precision and accuracy at 99%!



xgboost: number 1 in recall at 91% with 96% in precision.



classification final results

Model	Precision	Recall	Accuracy	F1-Score
Decision Tree	89.00%	90.00%	0.99926	89.00%
Logistitic Regression	93.00%	79.00%	0.99912	85.00%
Random Forest	**99.00%**	89.00%	**0.99960**	**93.00%**
XGBoost	96.00%	**91.00%**	0.99958	**93.00%**

- The business objective of this process identified that *precision* would be the metric used to determine the best classification model. The highest precision (followed by accuracy) would alleviate the banks obligations in any future fraudulent charges.

Thank You!