Credit Card Fraud Detection-Classification

Abdulrahman
Tahani
Mutab

Introduction

- Our business problem is to accurately identify fraudulent credit card transactions.
- This is an important issue as:
  - The global credit card fraud volume was $24 billion in 2018.
  - It's growing by over 10% a year.
  - Fraud is becoming more sophisticated.
  - It impacts how much customers trust financial institutions and merchants.

Problem statement

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be fraud.

This model is then used to identify whether a new transaction is fraudulent or not.

Our aim here is to detect 100% of the fraudulent transactions while minimizing the incorrect fraud classifications

Why credit card fraud matters

The Federal Trade **Commission estimates that 10 million people are victimized by credit card theft each year**

Credit card companies lose close to $50 billion dollars per year because of fraud

Dataset Description

There are 492 frauds out of 284,807 transactions.

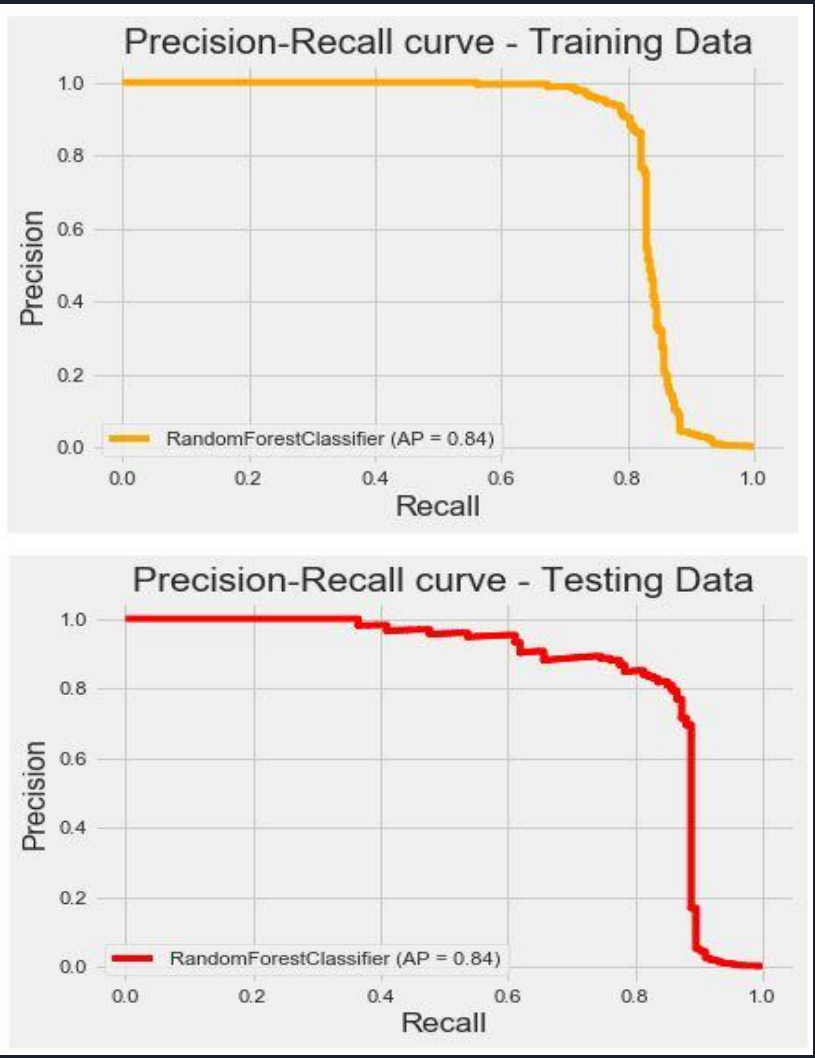
There are 31 variables.

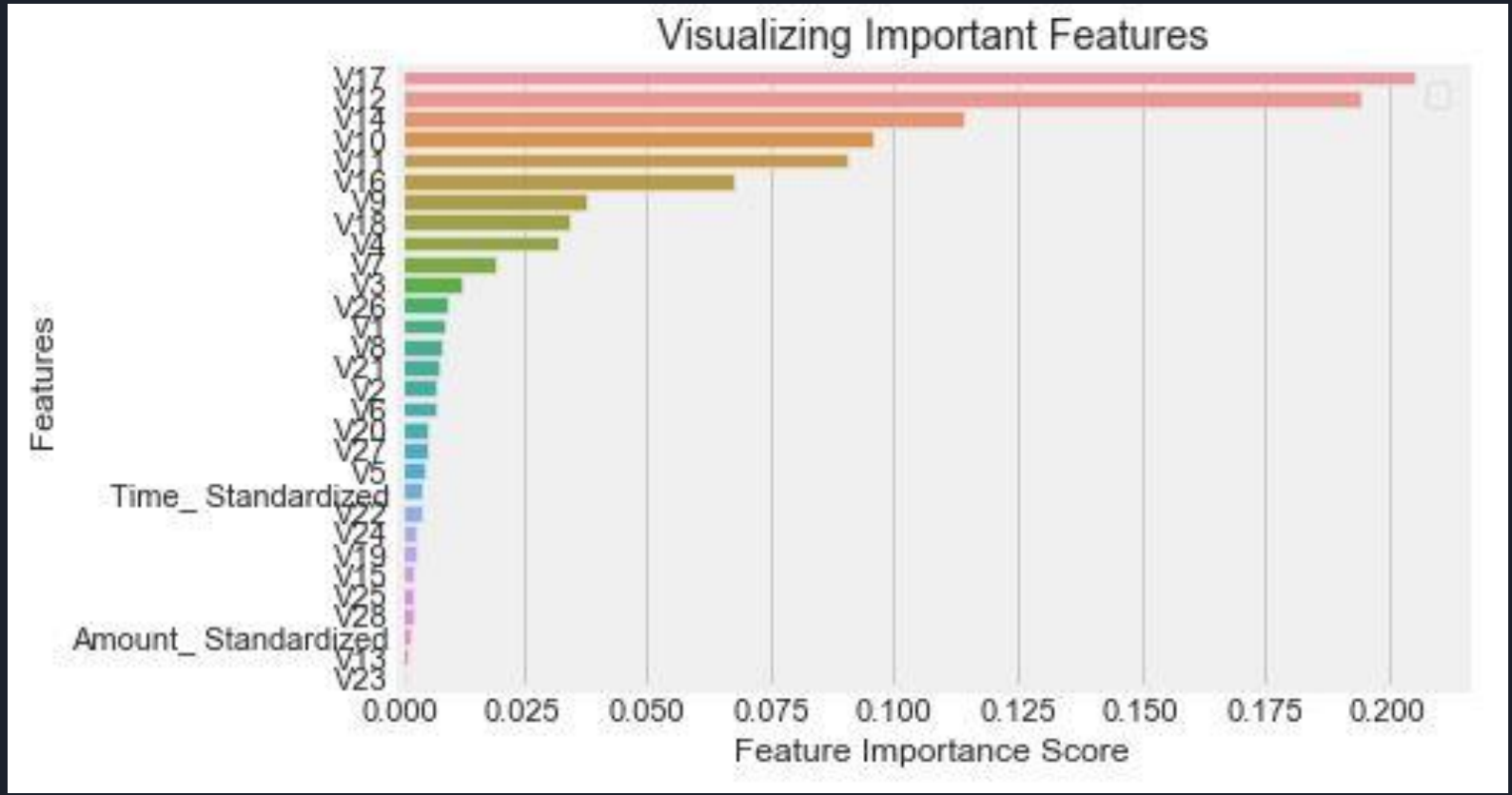
	Time	V1	V2	V3	V4	V5	...	V25	V26	V27	V28	Amount	Class
178698	123737.0	1.876980	-0.696291	0.450719	1.508030	-1.020571	...	-0.153727	-0.475249	0.113934	-0.023137	23.45	0
76266	56484.0	-2.905057	-2.553843	1.842890	-0.212129	0.480680	...	0.162048	0.701272	-0.250248	0.127260	427.64	0
51127	44807.0	1.095063	-0.302038	0.561085	-0.038436	-0.673402	...	0.008530	0.782091	-0.078727	0.015462	81.75	0
201147	133746.0	-2.739547	1.594012	-1.905468	0.576733	0.886780	...	-0.294596	-1.103543	-0.922379	-0.229078	79.00	0
73617	55206.0	0.990980	-0.365387	-0.414428	0.237689	-0.404407	...	0.473201	1.054140	-0.148832	0.019787	178.80	0
273023	165382.0	1.978600	0.680265	-1.001308	3.520812	1.053766	...	-0.110191	-0.339700	-0.036152	-0.040014	8.64	0
82712	59517.0	1.198267	0.265913	0.399505	0.631122	-0.465584	...	0.138864	0.066304	-0.030512	0.021475	1.98	0
112805	72813.0	-0.510041	0.801484	1.955988	1.010750	0.430620	...	0.381213	-0.195755	-0.342545	-0.403932	57.00	0
30261	35844.0	-3.222500	-3.641709	2.577789	-0.929482	2.402566	...	0.685703	-0.106645	0.173078	0.184205	12.48	0
79333	57972.0	-2.558141	-2.292401	1.450960	0.108837	-2.089814	...	0.022859	0.123339	0.944169	-0.198398	445.00	0
123561	76978.0	1.050040	-1.777259	1.310847	-0.410554	-1.618252	...	0.513019	0.173519	0.089214	0.022859	123.00	0
14975	26291.0	-0.451848	0.927968	0.816701	1.963048	-0.920190	...	-1.007315	-0.324020	-0.216883	0.132860	44.03	0
255807	157410.0	-0.213966	0.990962	-1.084647	-0.310215	0.605261	...	-0.458031	0.511308	-0.202088	-0.015042	10.55	0
46909	42985.0	-4.075975	0.963031	-5.076070	4.955963	-0.161437	...	-0.304987	-0.106089	1.899714	0.511462	1.00	1
235167	148278.0	2.046638	-0.099353	-1.203877	0.207330	0.119362	...	-0.270674	0.202065	-0.073970	-0.073668	1.98	0

Result

- Precision-Recall curve for the training and test data.
- Model performance is similar for both the training and test data.
- The average precision is 1% lower for the test data versus the training data.
- There is no indication that the model is overfit the training data.

RF	Training	Test
Precision	0.96	0.89
Recall	0.73	0.75
F-Measure	0.83	0.81




Visualizing Important Features
Features
Time_ Standardized
Amount_ Standardized
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
Feature Importance Score

Conclusion

Overall the main project goal involved determining the optimal algorithm for analysis as well as the best-performing combination of factors to detect credit card fraud.

- **Reduction in number of fraud transactions.**
- **User can safely use his credit card for online transaction.**
- **Added layer of security**

(Recommendation)
How to prevent credit card fraud

Keep Your Credit Cards Safe

Avoid Giving out Your Credit Card Information

Report Lost or Stolen Credit Cards Immediately

Review Your Billing Statements Each Month

Make Strong Passwords and Keep Them Safe