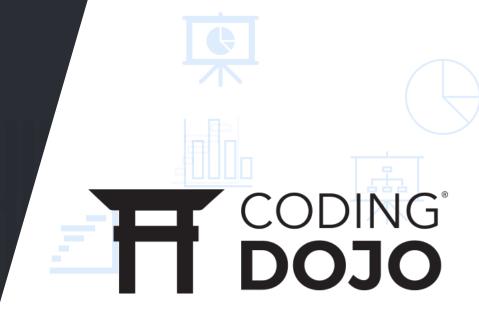
Data Science Bootcamp

Project 2
Transaction Fraud Detection

Marwa Salah









Project Overview



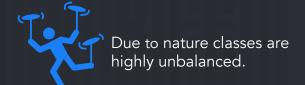
Identify potential fraudulant activity (transactions) using ML based on existing card transaction data



Anonymized credit card transactions labeled as fraudulent or genuine Transformed using PCA

<u>2 Features were not transformed using PCA</u> which are Time & Amount





0.17% vs 99.8% Class Balance



Transaction Faud Detection WHY?



\$2-3

Cost of each dollar lost to fraud



Impacts
Individuals | Businesses |

https://fortunly.com/statistics/cash-versus-credit-card-spending-statistics/#grefhttps://www.ncr.com/blogs/payments/credit-card-fraud-detection

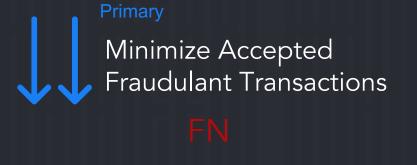
Financial Instituations



Objectives

Identify potentially fraudulent activity

Detecting fraudulent activity before those transactions are even completed





Secondary

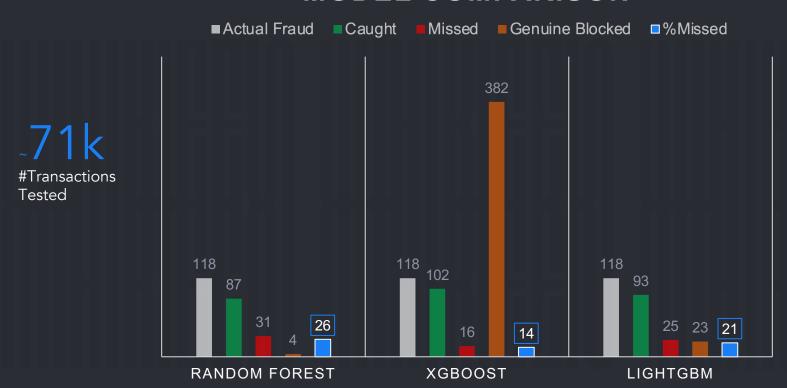
Avoid Declining
Genuine Transactions

FP



Outcome

MODEL COMPARISON





Conclusion & Next Step

Full Dataset

Fraud: \$ 60K Genuine: \$ 25M @86% Detection Rate

Fraud: \$ 9K Genuine: \$ 25M

Example based costsensitive learning & Cost-dependant Classification

	Actual Fraud y _{true} = 1	Actual Legitimate y _{true} = 0
Predicted Fraud y _{pred} = 1	True Positive cost _{TP} = Admin	False Positive cost _{FP} = Admin
Predicted Legitimate $y_{pred} = 0$	False Negative cost _{FN} = Transaction	True Negative cost _{TN} = \$0

ttps://towardsdatascience.com/fraud-detection-with-cost-sensitive-machine-learning-24b8760d35d9



Challenges

- Annomyzation of Data
- Diversity of Data
- Time
- Compute Resources



Technical Outcome Summary





Outcome Summary

	Model	precision	recall	f1-score					
	Random Forest	1	1	1					
O Genuine	XGBoost	1	0.99	1					
	LightGBM	1	1	1					
1 Fraud	Random Forest	0.96	<u>0.74</u>	0.83					
	XGBoost	0.21	0.86	0.33					
	LightGBM	0.83	0.80	0.81					

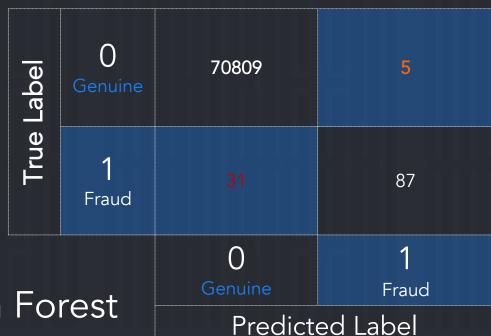


Confusion Matrix



96% Precision

Random Forest



70814

118

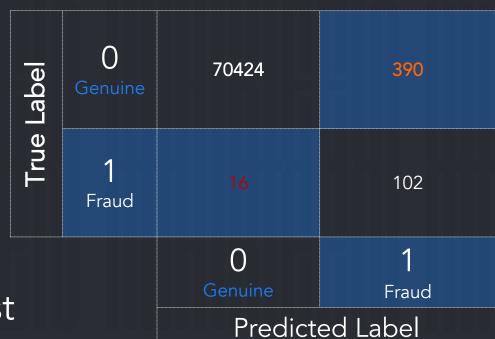


Confusion Matrix



21% Precision

XGBoost



70814

118



Confusion Matrix



80%
Precision

LightGBM



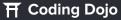
70814

118



Project Description

- In this project we were initially tasked with selecting a couple of datasets for analysis that meet the following rules:
 - Available to use
 - Suitable for use in a professional environment
 - O Does not contain any personal information
 - O Not used for any assignment, lecture or task from this environment
- My Proposed Sets: Credit Card Transaction Data & Financial Sales Data. Former was approved.
- So far, the project covers the following progress:
 - Data Cleaning & EDA
 - Evaluating & Tuning 3 Different Models (RandomForest, XGBoost, LightGBM)
 - Testing Impact of example-based cost sensitive learning
 - Model performance comparison



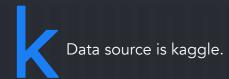


Data Description



Anonymized credit card transactions labeled as fraudulent or genuine Transformed using PCA

Contains transactions made by European credit card holders in Sep 2013 2 Features were not transformed using PCA which are Time & Amount





Due to nature classes are highly unbalanced.



Data Shape is:

- 28 Components, 2 Features & 1 Label 284807 Rows (Observations)



Data Types: All features are Integers of type int64





Classification Target:

Feature 'Class' is the response variable, and it takes value 1 in case of fraud and 0 otherwise.



Data Cleaning

- Due to PCA preprocessing Data Cleaning needed was very limited.
- No Nulls found
- No Data Type Inconcistencies
- No Categorical Inconsistencies would exist
- No Illogical Values
- A Number of Duplicates were there (~1000 rows) which were dropped



Visuals



1	0.12	-0.011	-0.42	-0.11	0.17	-0.063	0.085	-0.038	-0.0079	0.031	-0.25	0.13	-0.066	-0.1	-0.18	0.011	-0.074	0.09	0.03	-0.051	0.046	0.14	0.051	-0.016	-0.23	-0.042	-0.0052	-0.0093	-0.011	-0.012
0.12	1	0.0069	-0.0081	0.0023	-0.007	0.00041	-0.0092	-0.0012		0.00082	0.001		-0.00057		-0.0006	-0.0033	-0.0035		0.00092	-0.0014	0.0028	-0.0014			-0.00022		-0.016	-0.0049	-0.23	-0.094
-0.011	0.0069	1	0.0053	-0.0015	0.0052	-0.00059	0.0074	0.0029	-0.00027	0.00062	-0.00063	0.0023	0.00068	0.0027	0.0015	0.004	0.0032	0.0025	-0.00036	-0.0013	-0.0049	0.0012	-0.0039	0.0007	-0.0016	0.00025	0.0076	0.0016	-0.53	0.085
-0.42	-0.0081	0.0053	1	0.0028	-0.0069	-0.0015	-0.012	-0.0018	-0.0036	-0.0096	0.0023	-0.0059	0.00011	-0.003	-0.0012	-0.0044	-0.0082	-0.0035	-1.6e-05	-0.0023	0.0035	-0.00027	0.00045	-7.2e-05	0.00043	-9.4e-05	-0.0071	-0.00013	-0.21	-0.18
-0.11	0.0023	-0.0015	0.0028	1	0.0017	-0.00088	0.0047	0.00089	0.0022	0.0028	-0.0012	0.0034	0.00018	0.0028	0.00057	0.0033	0.0037	0.0023	-0.00056	0.00032	-0.001	0.00012	0.00073	-0.00012	0.00016	0.00078	0.0013	0.00023	0.1	0.13
0.17	-0.007	0.0052	-0.0069	0.0017	1	-0.00094	-0.0087	0.0014	-0.0012	-0.006	0.00041	-0.0023	1.9e-05	-0.001	-0.0012	-0.0024	-0.0045	-0.0027	0.00044	-0.0012	0.0016	-0.00056	0.0012	0.0002	6.9e-05	0.00039	-0.0058	-0.00082	-0.39	-0.088
-0.063	0.00041	-0.00059	-0.0015	-0.00088	-0.00094	1	0.00044	0.003	-0.00073	-0.0022	-0.00021	-0.0012	0.0004	0.00018	-0.00047	0.00012	-0.0017	0.00054	0.00011	-0.00018	-0.0021	0.0011	-0.00076	0.0012	0.0007	-2.8e-05	0.00029	0.00092	0.22	-0.044
0.085	-0.0092	0.0074	-0.012	0.0047	-0.0087	0.00044	1	-0.0064	-0.0049	-0.014	0.0025	-0.0062	-0.00017	-0.0038	-0.0014	-0.0059	-0.0088	-0.0043	0.00085	-0.0012	0.009	-0.0023	0.0033	-0.00038	-7.2e-05	0.00062	-0.0045	0.0017	0.4	-0.17
-0.038	-0.0012	0.0029	-0.0018	0.00089	0.0014	0.003	-0.0064	1	0.001	0.00048	0.0047	-0.0044	-0.0014	-0.0084	0.001	-0.0044	-0.0056	-0.0013	-0.00063	0.00027	0.019	-0.0062	0.005	0.00011	1.1e-05	-0.0014	0.00061	-9.9e-05	-0.1	0.033
-0.0079	0.0018	-0.00027	-0.0036	0.0022	-0.0012	-0.00073	-0.0049	0.001	1	-0.013	-0.00022	-0.0024	0.00075	0.002	-0.00028	-8.6e-05	-0.0023	-0.00037	0.00025	-0.0018	0.00068	0.00078	0.00068	-0.0001	-0.00028	0.0013	0.0082	0.0056	-0.044	-0.094
0.031	0.00082	0.00062	-0.0096	0.0028	-0.006	-0.0022	-0.014	0.00048	-0.013		0.00084	-0.0069	0.0014	0.00017	-0.0023	-0.0037	-0.0079	-0.0025	0.0011	-0.0044	0.0038	-0.00048	0.0019	0.00015	-0.00056	0.0011	0.011	0.0092	-0.1	-0.21
-0.25	0.001	-0.00063	0.0023	-0.0012	0.00041	-0.00021	0.0025	0.0047	-0.00022	0.00084	1	0.0056	0.00046	0.0077	-0.00087	0.0048	0.0074	0.0021	-0.00049	-0.00099	-0.0028	-0.00015	-3.7e-05	8e-05	4.7e-05	-0.0002	0.002	0.0026	-1.5e-05	0.15
0.13	-0.0015	0.0023	-0.0059	0.0034	-0.0023	-0.0012	-0.0062	-0.0044	-0.0024	-0.0069	0.0056	1	-0.00055	-0.01	0.00088	-0.0074	-0.013	-0.0035	0.00059	0.0012	0.0033	0.00015	0.00049	0.00059	-0.00018	-0.00014	-0.00093	-0.00061	-0.0093	-0.25
-0.066	-0.00057	0.00068	0.00011	0.00018	1.9e-05	0.0004	-0.00017	-0.0014	0.00075	0.0014	0.00046	-0.00055	1	-0.0011	0.00023	-0.00081	-0.00017	-0.00016	8.6e-05	0.00038	0.00052	1.6e-05	0.00025	-4.9e-05	0.00025	-0.0001	-0.0016	-0.0006	0.0052	-0.0039
-0.1	-0.0027	0.0027	-0.003	0.0028	-0.001	0.00018	-0.0038	-0.0084	0.002	0.00017	0.0077	-0.01	-0.0011	1	0.00087	-0.0091	-0.014	-0.0045	0.0016	0.0027	0.0056	-0.0019	0.00067	-2.6e-05	0.00016	-0.0007	-0.0046	-0.0047	0.034	-0.29
-0.18	-0.0006	0.0015	-0.0012	0.00057	-0.0012	-0.00047	-0.0014	0.001	-0.00028	-0.0023	-0.00087	0.00088	0.00023	0.00087	1	-0.00028	-0.00017	-0.00043	0.00055	-0.00075	-0.00027	-0.0012	0.00097	0.00011	0.00044	-0.002	-0.00064	0.00086	-0.0033	-0.0033
0.011	-0.0033	0.004	-0.0044	0.0033	-0.0024	0.00012	-0.0059	-0.0044	-8.6e-05	-0.0037	0.0048	-0.0074	-0.00081	-0.0091	-0.00028	1	-0.0091	-0.0054	0.0024	0.0011	0.0043	-0.00082	0.0012	-0.00048	0.00021	-0.0012	-0.004	-0.0016	-0.0045	-0.19
- 0.074	-0.0035	0.0032	-0.0082	0.0037	-0.0045	-0.0017	-0.0088	-0.0056	-0.0023	-0.0079	0.0074	-0.013	-0.00017	-0.014	-0.00017	-0.0091	1	-0.0053	0.00099	0.0015	0.0036	-0.00016	0.00067	0.001	-0.00068	0.00016	-0.0034	-0.0027	0.0077	-0.31
0.09	-0.0035	0.0025	-0.0035	0.0023	-0.0027	0.00054	-0.0043	-0.0013		-0.0025	0.0021		-0.00016	-0.0045	-0.00043		-0.0053	1		-0.00024		-0.00053		-0.00071		-0.0006	-0.0042	-0.0013	0.036	-0.11
		-0.00036	-1.6e-05					-0.00063	0.00025		-0.00049	0.00059	8.6e-05	0.0016	0.00055	0.0024	0.00099	-0.0001	1		0.00024	0.0013	0.00038		-8.4e-05		-0.00054	0.00035	-0.056	0.034
-0.051	-0.0014	-0.0013	-0.0023	0.00032	-0.0012	-0.00018	-0.0012	0.00027	-0.0018	-0.0044	-0.00099	0.0012	0.00038	0.0027	-0.00075	0.0011	0.0015	-0.00024		1	0.0054	-0.0016	-0.0011	-0.0003	-0.00064	-0.00031	-4.9e-05	0.0027	0.34	0.021
0.046	0.0028	-0.0049	0.0035	-0.001	0.0016	-0.0021	0.009	0.019	0.00068	0.0038	-0.0028		0.00052	0.0056	-0.00027	0.0043	0.0036	0.0016		0.0054	1	0.0096			-0.00087	-0.00087	-0.0052	-0.0044	0.11	0.026
0.14	-0.0014	0.0012	0.00027	0.00012	-0.00056	0.0011	0.0023	0.0062	0.00078	-0.00048	-0.00015 -3.7e-05	0.00015	0.00025	0.0019	0.0012	0.00082	0.00016	0.00053	0.0013	-0.0016	0.0096	0.0019	0.0019	-3.1e-05	-0.0002	-0.0015	-0.003	-0.0014	-0.065	0.0049
	-0.00072	0.0007	-7.2e-05		0.0012	0.00076	-0.00038	0.00011		0.0019	-3.7e-05	0.00049	4.9e-05	-2.6e-05	0.00097		0.001	-0.00071		-0.0011	0.0012		0.00027		-0.00019		-0.00088		0.0051	-0.0063
	.0.00072	-0.0016		0.00012	6.9e-05	0.0012	-7.2e-05	1.1e-05	-0.00028		4.7e-05	-0.00018		0.00016	0.00011	0.00048	-0.00068	-0.00071	-8.4e-05	-0.00064	-0.00087	0.0002		-0.00019	1	4.8e-05	-0.0003		-0.048	0.0072
	.0.00022	0.00025		0.00078	0.00039		0.00062	-0.0014	0.00028	0.0011	-0.0002	-0.00014	-0.00023	-0.0007	-0.002	-0.0012	0.00016	-0.0006	0.00086	-0.00031	-0.00087	-0.0015		0.00019		4.86-03	-0.0033	-0.001	-0.0034	0.0032
-0.0052	-0.016	0.0076	-0.0071		-0.0058	0.00029	-0.0045	0.00061	0.0082	0.011	0.002	-0.00093		-0.0046	-0.00064	-0.004	-0.0034		-0.00054	-4.9e-05	-0.0052	0.003	-0.002	-0.00088	-0.0013		1	-0.014	0.028	0.022
-0.0093	-0.0049	0.0016	-0.00013		-0.00082	0.00092		-9.9e-05	0.0056	0.0092	0.0026	-0.00061	-0.0006	-0.0047	0.00086	-0.0016	-0.0027	-0.0013		0.0027	-0.0044	0.0014	-0.0032	0.00032		-0.001	-0.014	1	0.01	0.0097
-0.011	-0.23	-0.53	-0.21	0.1	-0.39	0.22	0.4	-0.1	-0.044	-0.1	-1.5e-05	-0.0093	0.0052	0.034	-0.0033	-0.0045	0.0077	0.036	-0.056	0.34	0.11	-0.065	-0.11	0.0051	-0.048	-0.0034	0.028	0.01	1	0.0058
-0.012	-0.094	0.085	-0.18	0.13	-0.088	-0.044	-0.17	0.033	-0.094	-0.21	0.15	-0.25	-0.0039	-0.29	-0.0033	-0.19	-0.31	-0.11	0.034	0.021	0.026	0.0049	-0.0063	-0.0072	0.0032	0.0043	0.022	0.0097	0.0058	1



Outcome

MODEL COMPARISON

