

CAR INSURANCE FRAUD DETECTIONS

Presented by Feras Alyahya ,Ahmed Alonaizi ,Hussain
Alhadab and Mohammed Alhamoud

OUTLINE

01

introduction

02

Data

03

Result

04

Best model

05

Conclusion

INTRODUCTION

PROBLEM :

With the large number of traffic accidents insurance companies receive claims for financial compensation to the beneficiaries, and with the many claims appear multiple fraud cases and undeserved financial claims, so insurance companies face difficulty in identifying and detecting fraud in car accidents, insurance companies need a solution to help them detect fraud and identify the factors and causes of fraud, based on all factors related to the accident .

SOLUTION :

IN THIS PROJECT, WE ARE GOING TO DEVELOP A CLASSIFICATION MODEL TO DETECT FRAUD.

The background features a dark teal or black abstract graphic. It consists of several overlapping circles of varying sizes and thin, intersecting lines that create a sense of depth and motion. The overall aesthetic is minimalist and modern.

DATA

DATA DESCRIPTION

Dataset contains 34 columns and more than 11000 record

The background of the slide features a dark teal or black color. Overlaid on this are several abstract shapes: a large circle in the center-left, a smaller circle above it, and a larger circle below it. These shapes are defined by thin, light-colored lines that create a sense of depth and overlap. The overall aesthetic is minimalist and modern.

RESULT

Result

	test	F1	AUC	Log-Loss
Logistic Regression	using all features	0.000	0.717	0.208
	using some of the features	0.000	0.502	0.228
	using some of the features with dimensionality reduction	0.028	0.782	0.195
	using some of the features with Over-sampling using SMOTE	0.924	0.975	0.203
KNN	using some of the features	0.000	0.540	1.276
	After resampling & GridSearchCV	0.924	0.975	0.203
	using some of the features	0.122	0.799	0.253
Random Forest	using some of the features with Over-sampling using SMOTE	0.967	0.993	0.120
	using some of the features with RandomizedSearchCV & Over-sampling using SMOTE	0.964	0.992	0.152

Result

	test	F1	AUC	Log-Loss
xgboost	some of the features with Over-sampling using SMOTE	0.969	0.992	0.106
Support Vector Machine	Some of the features with Over-sampling using SMOTE	0.984	0.996	0.053

Result

model	accuracy	f1-score	prediction	precision	recall	f1-score
Logistic Regression	0.94	0.92	Not Fraud	0.92	0.93	0.92
			Fraud	0.93	0.92	0.92
KNN	0.98	0.98	Not Fraud	0.99	0.98	0.98
			Fraud	0.98	0.99	0.99
Random Forest	0.96	0.96	Not Fraud	0.94	0.98	0.96
			Fraud	0.98	0.95	0.96
xgboost	0.97	0.96	Not Fraud	0.94	0.98	0.96
			Fraud	0.98	0.96	0.97
Support Vector Machine	0.98	0.98	Not Fraud	1.00	0.96	0.98
			Fraud	0.97	1.00	0.98

The background features a dark teal or black gradient with several large, semi-transparent white circles of varying sizes. Some circles overlap, creating a sense of depth. Thin black lines form a grid-like pattern across the circles.

BEST MODEL

BEST MODELE

test F1 AUC Log-Loss

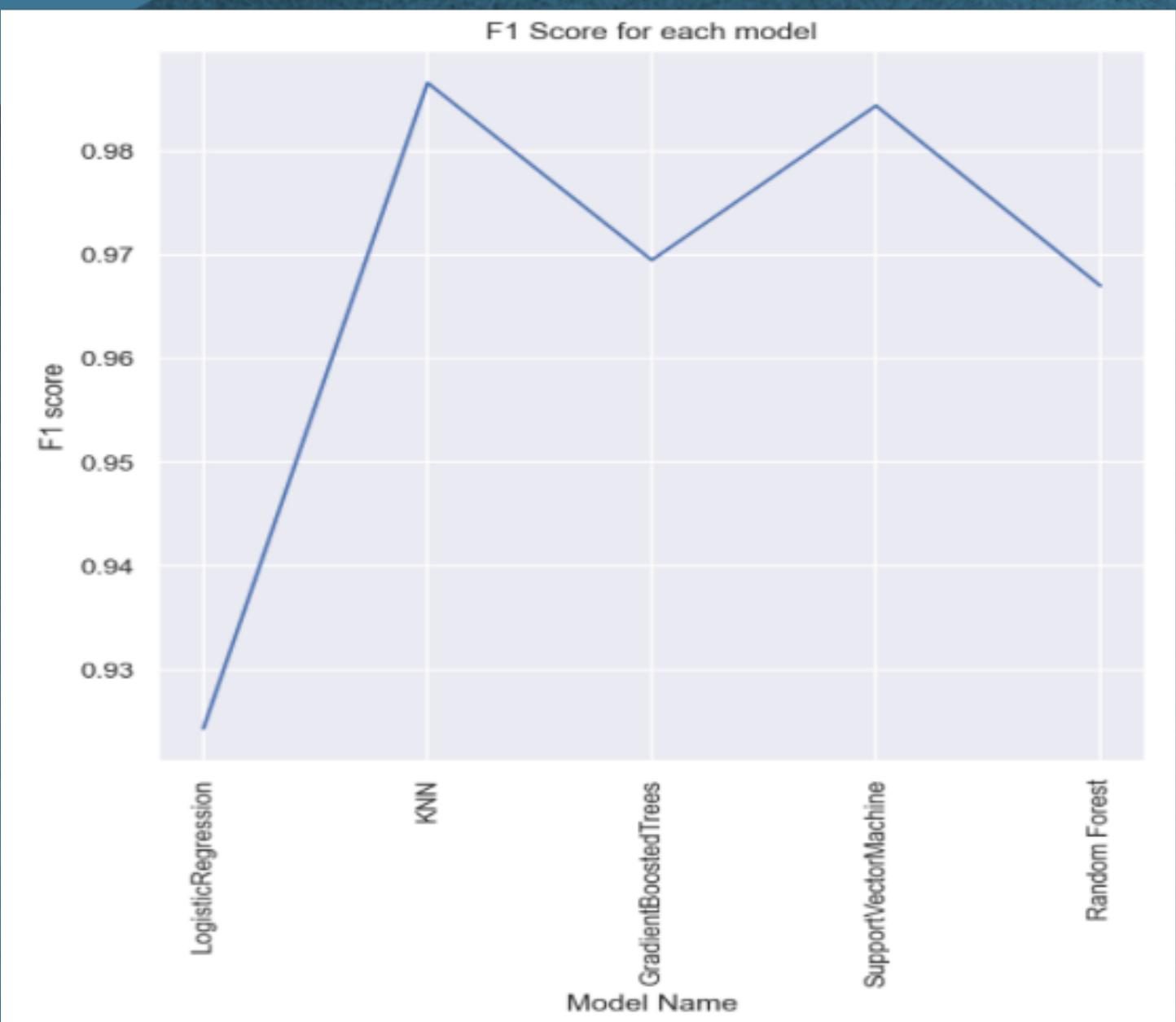
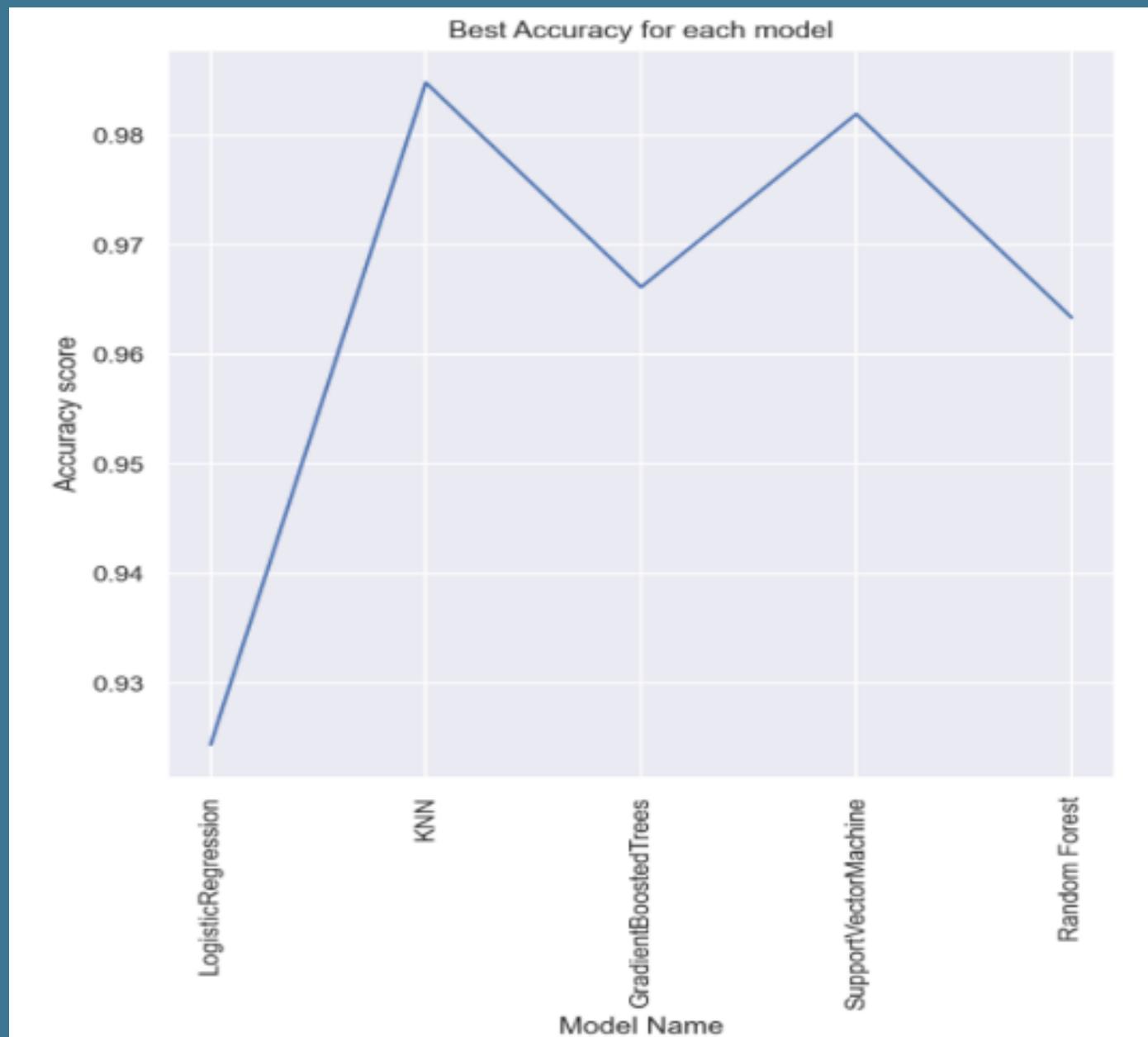
KNN

After resampling &
GridSearchCV

0.987

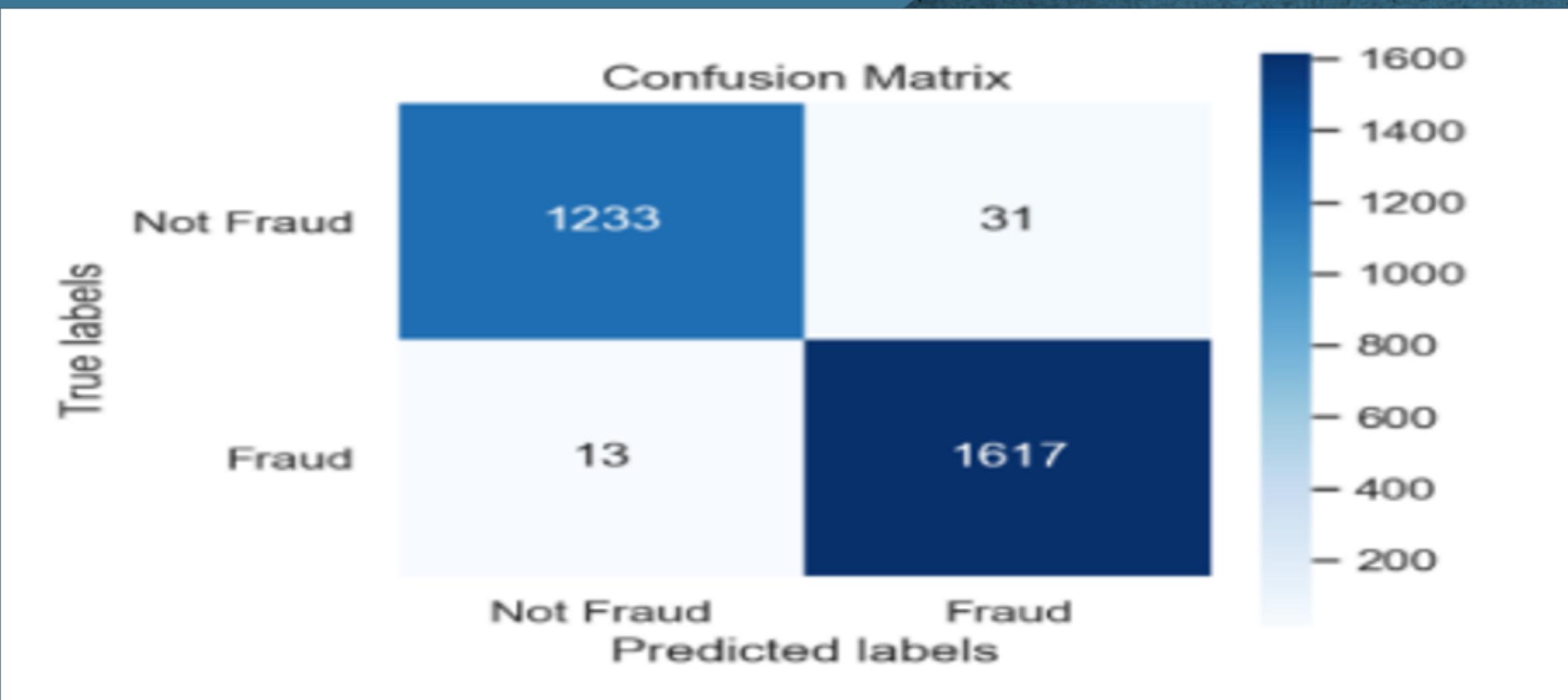
0.984

0.525



BEST MODELE

	test	F1	AUC	Log-Loss
KNN	After resampling & GridSearchCV	0.987	0.984	0.525



CONCLUSION

We made 5 models, We came up with the best model to do fraud detection
After receiving the characteristics for each claim, the model will help insurance companies to help them detect fraud clime. We achieved very good accuracy **98%** In the best model

The background features a dark teal color with abstract white shapes. On the left, there are two large, semi-transparent circles: one is light blue and the other is a darker shade of teal. Overlaid on these circles are several thin, black, wavy lines that intersect and curve across the frame.

THANK YOU!