TechE

TECHNOUTSAV 3.0

FRAUD DETECTION SYSTEM

Team : SPIT Mumbai_KarmaBytes

Team Members : Pulin Prabhu
Tania Rajabally
Anisha Gharat

College : Sardar Patel Institute of Technology

Theme : Data Science/Machine Learning

Deloitte.

Round 2 Phase 1 – Ideation

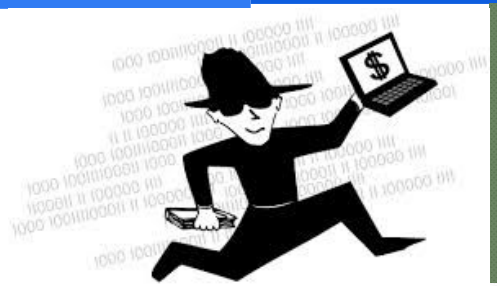
Deliverables	Use Case
Brief Business Problem	< Frauds in online transactions is seen in abundance with the increase in the use of online methods for making payments.Studies suggest 93% of the merchants perform manual review anywhere between 1% - 10% of orders for fraud detection and this is costly, time consuming and not only that but it also leads to higher false negatives.>
Proposed Solution	< A Machine Learning model that will be trained to detect frauds and then accept or reject the transactions based on the result. >
Proposed Tech Stack	< Django Server, Python, TensorFLow, Selenium Web Drivers, Matplotlib, Pandas, Seaborn, Numpy, Basic HTML, CSS for Frontend, Logistic Regression, Random Forest Algorithm>
Market Place/ Positioning	<Financial sector being used abundantly and has a higher possibility of fraud. Currently the merchants depend on manual ways of fraud detection which is slow, less accurate and costlier. The proposed method provides speed, scale and efficiency alongside accuracy. Businesses using the solution will in turn endup saving a lot of resources since they won't have to train employees in order to manually detect frauds. Also the frauds detected using the ML model will be way higher and accurate as compared to the ones done manually.>
Effort (Hours) and Cost (INR) of Implementation	< **Effort**: Person Hours - 3 person * 135(9 hours*5 days*3weeks) = 405 hours **Cost**: Software - Licensed Software's Cost >
Theme Alignment	Data Science / Machine Learning
Industry Alignment	Financial Services

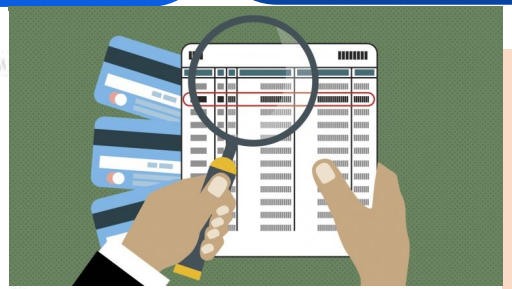
Brief Business Problem

Fraud Rates increasing alarmingly in the finance sector

Data Collected from various points that are not consolidated enable this

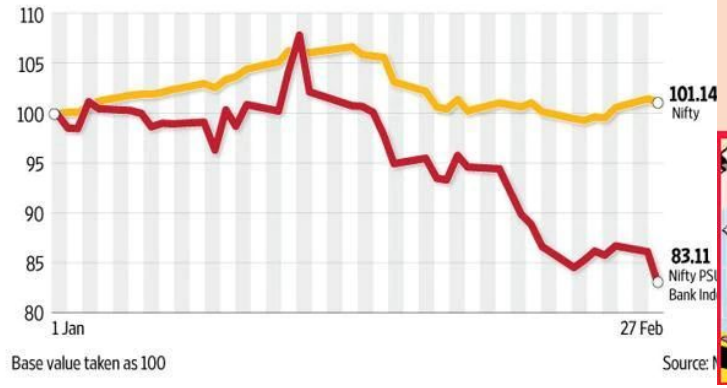
A Fraud Detection System to tackle this issue





ALL FALL DOWN

Public sector bank stocks have plummeted a massive 12% since the fraud at Punjab National Bank came to light.



Base value taken as 100

Source: N

63% of businesses have experienced the same or more fraud losses in the past 12 months

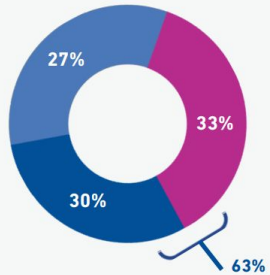


Figure 5

Reeling Under the Pressure

Bank Group/ Institution	2017-18 Number of Frauds (₹ cr)	2017-18 Amount Involved (₹ cr)	2018-19 Number of Frauds (₹ cr)	2018-19 Amount Involved (₹ cr)
Public Sector Banks	2,885	38,260.87	3,766	64,509.90
	-48.8	-92.9	-55.4	-90.2
Private Sector Banks	1,975	2,478.52	2,090	5,515.10
	-33.4	-6	-30.7	-7.7
Foreign Banks	974	256.00	762	955.30
	-16.5	-0.6	-11.2	-1.3
Financial Institutions	12	164.70	28	553.40
	-0.2	-0.4	-0.4	-0.8
Total	5,916	41,167.00	6,801	71,542.90

Source: RBI

TOP 5 BANK FRAUDS OVER THE LAST FEW YEARS

1. Punjab National Bank : ₹13,000 cr
2. PMC Bank: ₹4,355 cr
3. [illegible]
4. Cosmos Bank: ₹90 cr
5. Bhushan Power: ₹2,348 cr

Proposed Solution

Extract the Data + Provide Training Sets

- ➔ The data collected from various data points in a single platform will be split into three different segments – training, testing and cross - validation.
- ➔ The **train** data will be trained for prediction.
- ➔ The **test** data will be used for predicting the frauds.
- ➔ The high performing models will be cross - validated to ensure consistency in results.

Training & Building Models

- ➔ The data is trained for certain input parameters to predict boolean outputs.
- ➔ Models will be built for prediction based on previous examples of data. We can further divide the prediction problem into two types of tasks:
  1. Classification
  2. Regression

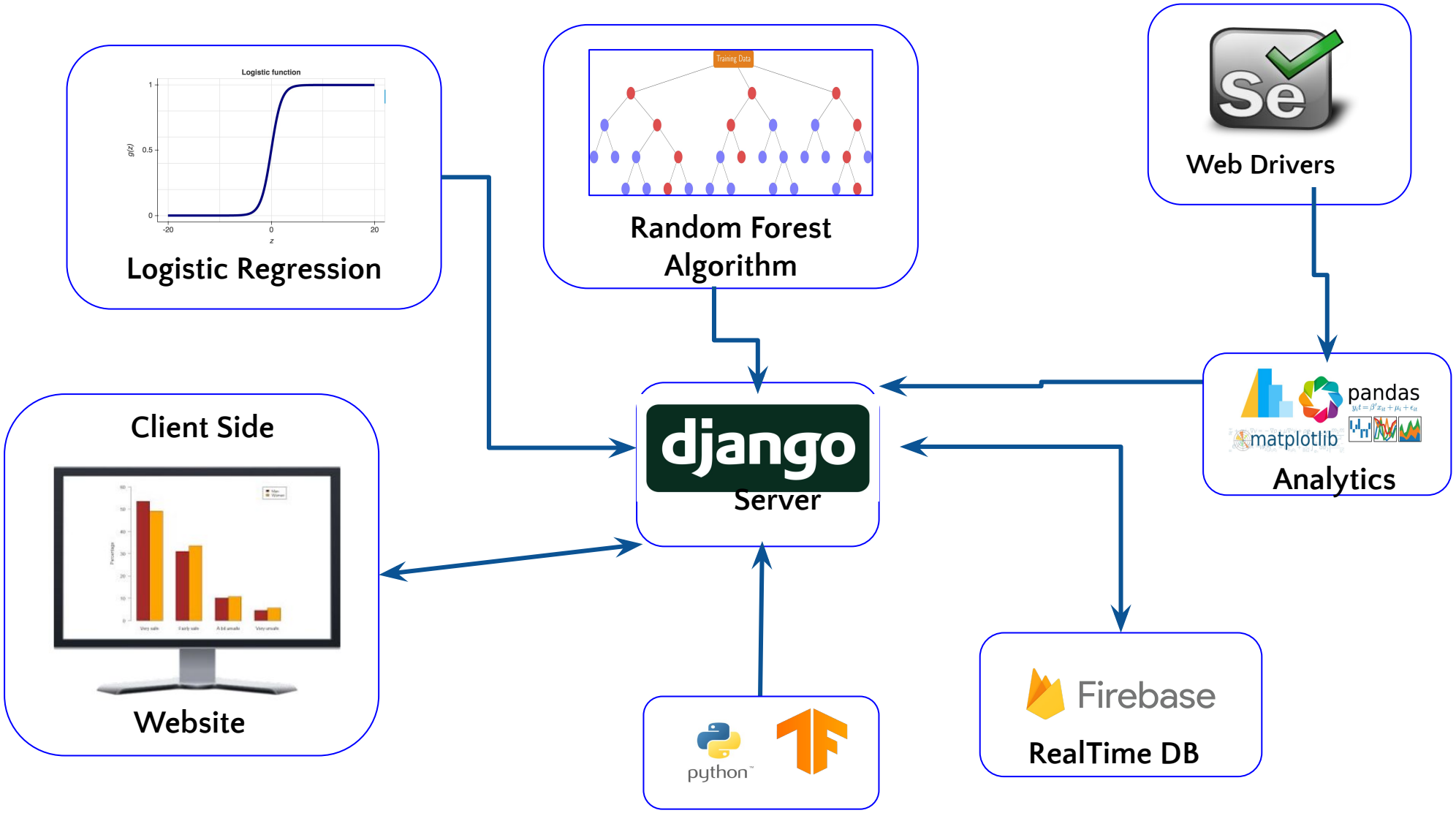
Detect/Predict Frauds

- ➔ This model will predict whether a new transaction is **fraudulent** or not.
- ➔ The model will give a probability score of fraud based on earlier scenarios.
- ➔ Any fraud by individual or companies will be detected on a single platform.

Notify/Take actions

- ➔ The risk of frauds will be estimated.
- ➔ Based on the predictive analysis and probabilities of frauds the company can either **Accept** or **Reject** transactions.

Proposed Tech Stack



Marketplace / Positioning

- ➔ Financial Services are everywhere and are function with high complexity.
- ➔ The solution mainly focuses on 3 main features namely speed, scale and efficiency.
- ➔ With betterment in these areas alongside accuracy it will reduce time which in turn shall save a lot of money which is one of the key saving in this specific industry.
- ➔ This is beneficial for all sectors that are even mildly aligned to finance or make use of it.
- ➔ The solution will help them analyse based on various factors if the ongoing transaction can be harmful and will prevent many unfortunate incidents.

Theme Alignment

- Data Science/Machine Learning
  - Indeed, even with present day investigation devices, it sets aside a great deal of effort for people to peruse, gather, sort and break down the information. ML instructs machines to distinguish and check the significance of examples instead of people.
  - Especially for use situations where information must be investigated and followed up on in a short measure of time, having the help of machines permits people to be progressively productive and act with certainty

Industry Alignment

- Financial Services
  - India has a diversified financial sector undergoing rapid expansion, both in terms of strong growth of existing financial services firms and new entities entering the market.
  - The sector comprises commercial banks, insurance companies, non-banking financial companies, co-operatives, pension funds, mutual funds and other smaller financial entities.