BLACKCOFFER

MANAGEMENT CONSULTING FIRM

STATISTICS & OPTIMIZATION

Statistics & Optimization with Big Data

Data driven, technology and data science company focused on helping enterprises to solve big data and analytics problems of any kind, from any source, at massive scale.

The Problem

How Fraud causing several problems in digital payments and online market place?

- ❑ It is hard to integrate, affordable and secure digital payments infrastructure for small and medium sized merchants engaged in online and mobile commerce in emerging markets.

- ❑ The rising threat of cyber security and credit card fraud is damaging the growth of cross-border online and mobile commerce

Solution

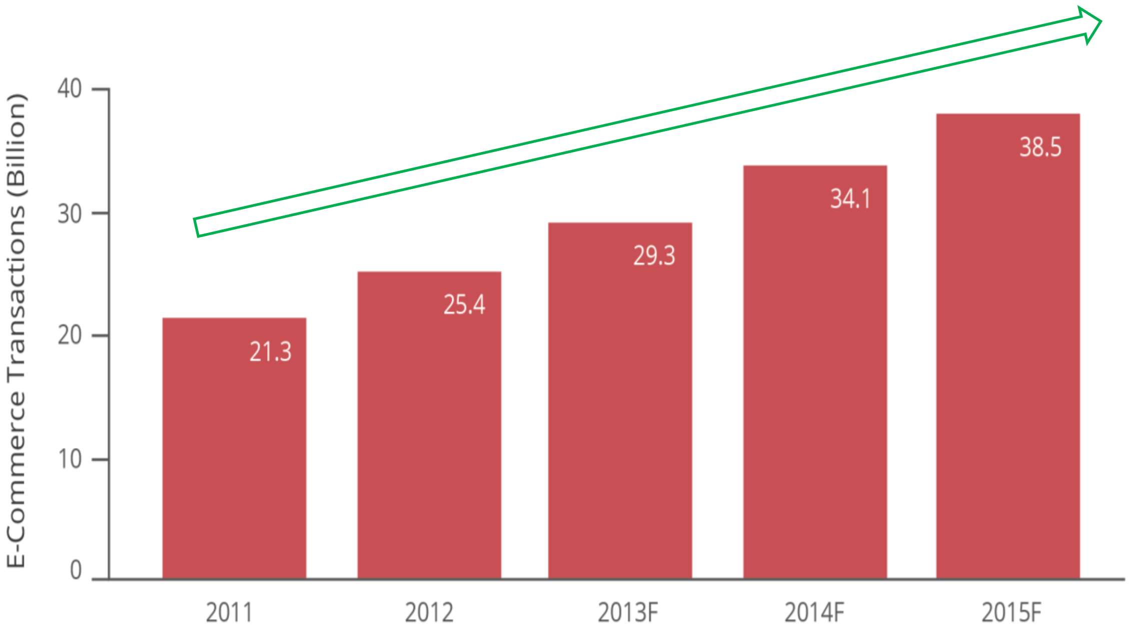
Fraud prevention and detection

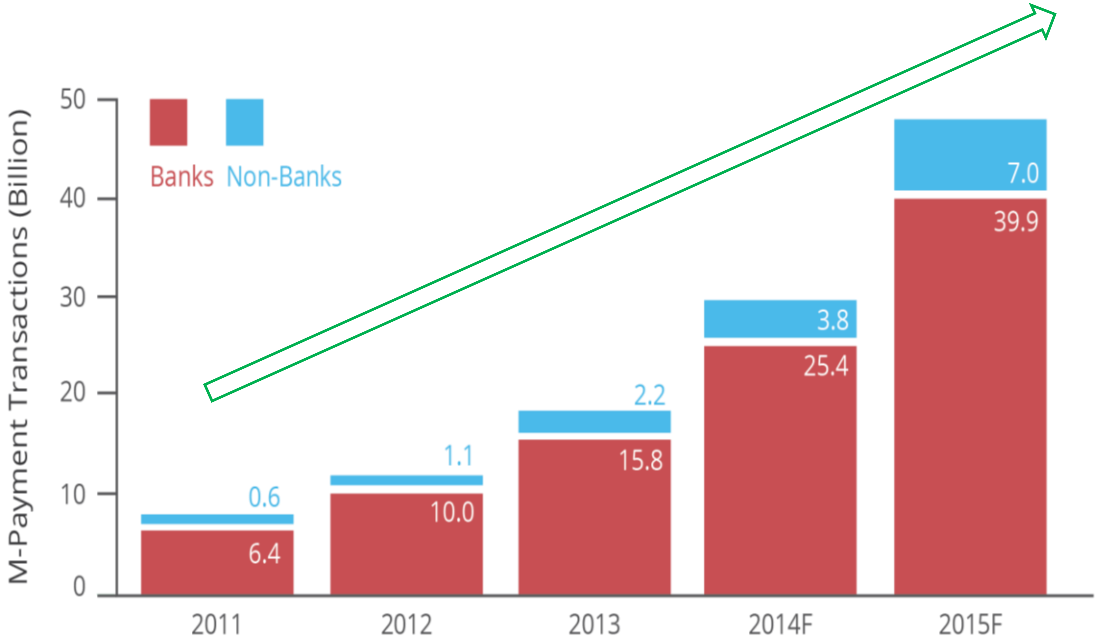
- Differentiated, guaranteed fraud prevention services by combining fraud detection technologies with machine learning capabilities

- Using the power of data to predict future events, match patterns and map identities to build a 360˚ view of each transaction, scoring each transaction for risk of fraud from hundreds of data points resulting in accurate, real-time, and 100% automatic Accept or Decline decision

Digital Payment Overview

Digital payment over time from 2011 till 2015





Growth continues in the e- and m-payments markets, along with convergence between the two modes, as some e-payments transactions migrate towards m-payments due to increased use of tablets and smartphones.

Source: World Payments Report 2014, Authored by Capgemini and Royal Bank of Scotland

Fraud and Analytics

Type of Fraud

- ❖ **Merchant Identity Fraud**
- ❖ **Friendly Fraud**
- ❖ **Merchant Credit Risk**

- ❖ **Buyer Identity Fraud**
- ❖ **Chargebacks**
- ❖ **Phishing**

How Data Analysis Prevents Fraud

Statistical data analytics techniques for fraud prevention

- Data pre-processing techniques for detection, validation, error correction, and filling up of missing or incorrect data
- Calculation of various statistical parameters such as averages, quintiles, performance metrics, probability distributions, and so on
- Models and probability distributions of various business activities either in terms of various parameters or probability distributions
- Computing user profiles
- Time-series analysis of time-dependent data
- Techniques are also needed to eliminate false alarms, estimate risks, and predict future of current transactions or users.

How Data Analysis Prevents Fraud

Artificial Intelligence Data Analytics Techniques

- ❖ Data mining to classify, cluster, and segment the data and automatically find associations and rules in the data that may signify interesting patterns, including those related to fraud.
- ❖ Expert systems to encode expertise for detecting fraud in the form of rules.
- ❖ Pattern recognition to detect approximate classes, clusters, or patterns of suspicious behaviour either automatically (unsupervised) or to match given inputs.
- ❖ Machine learning techniques to automatically identify characteristics of fraud.
- ❖ Neural networks that can learn suspicious patterns from samples and used later to detect them.

Methodology

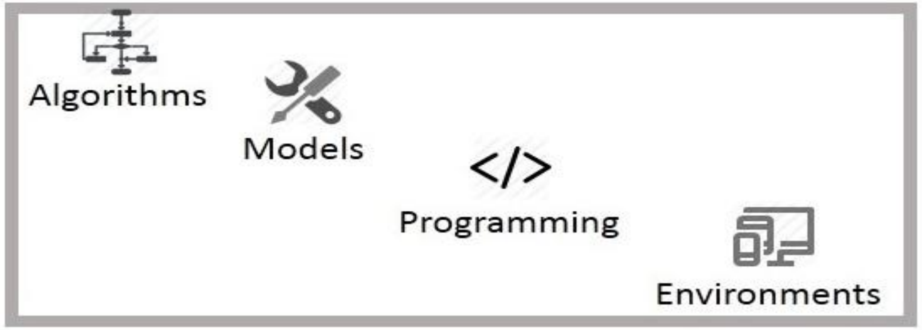
Mine Data ======➔ Analyse Data =====➔ Insights for fraud detection and prevention

❖ Demographic Data
- Name
- Address
- Age
- Gender

❖ Transactional Data
- What is being purchased
- Method of payments
- Frequency of past orders

❖ Device Data
- Device id
- Device used in past
- Mobile number
- IP Address

❖ Merchant Data
- Merchant Account
- Transaction Frequency
- IP Address
- Location Data



❖ Anomaly Detection
- Cluster Analysis
- K-nearest
- Deviations from Association rules
- Replicators neural Networks

❖ Predictive Analytics
- Regression Techniques
- Decision Trees

Pattern Recognition
- Naïve Bayes
- Support Vector Machines
- K-Means Clustering

❖ Detecting Impersonations

❖ Authentication of Merchants

❖ Live Fraud Detection

❖ Minimize Chargebacks

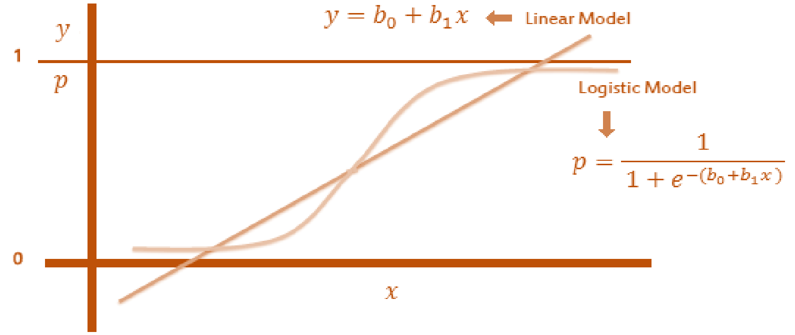
❖ Authentication of Payment Methods

❖ Real time reports and monitoring

Machine Learning Techniques

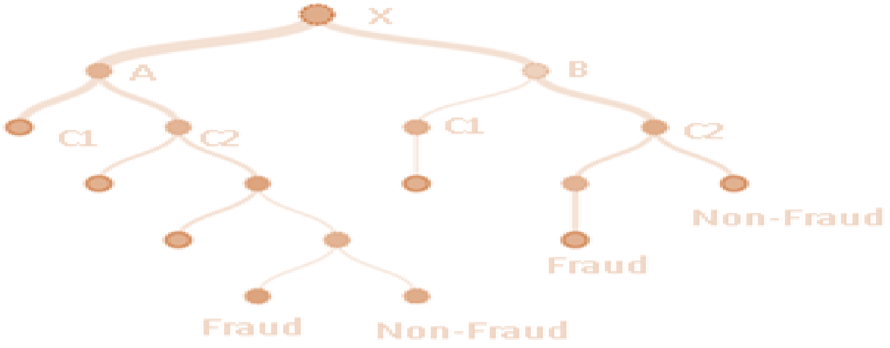
Cutting edge technologies can help to detect fraud

1. Regression Techniques



- ❖ Techniques like Logistic, Probit, Zero-inflated Poisson regressions can used to predict the probability of a company turning out to be a fraud or not.
- ❖ Similarly, various factors of fraudulent companies can be analyzed using multivariate regression to identify the main indicators of fraudulent characters in a firm.

2. Decision Trees

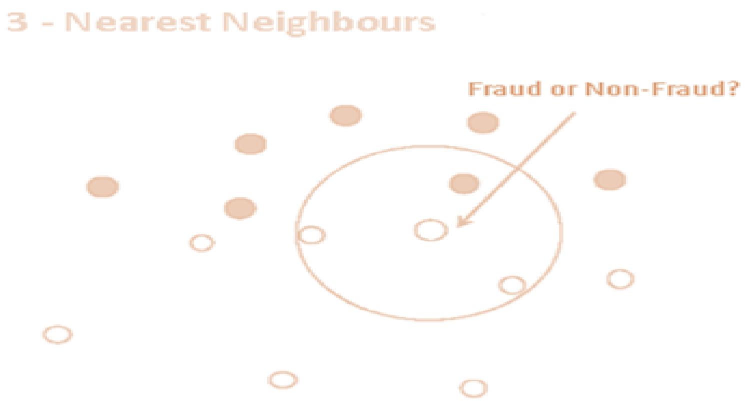


- ❖ A decision tree is a predictive machine-learning model that decides the target value (dependent variable) of a new sample based on various attribute values of the available data.
- ❖ The internal nodes of a decision tree denote the different attributes, the branches between the nodes tell us the possible values that these attributes can have in the observed samples, while the terminal nodes tell us the final value (classification) of the dependent variable.

Machine Learning Techniques

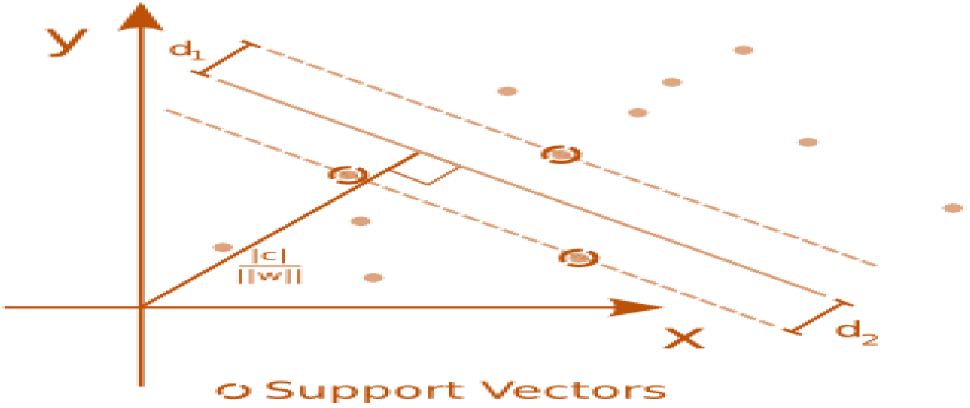
Cutting edge technologies can help to detect fraud

3. K-Nearest Neighbours



- Neighbours-based classification is a type of instance-based learning or non-generalizing learning: it does not attempt to construct a general internal model, but simply stores instances of the training data.
- Classification is computed from a simple majority vote of the nearest neighbours of each point: a query point is assigned the data class which has the most representatives within the nearest neighbours of the point.
- The k-neighbours classification is the most commonly used technique. The optimal choice of the value k is highly data-dependent: in general a larger k suppresses the effects of noise, but makes the classification boundaries less distinct.

4. Support Vector Machine

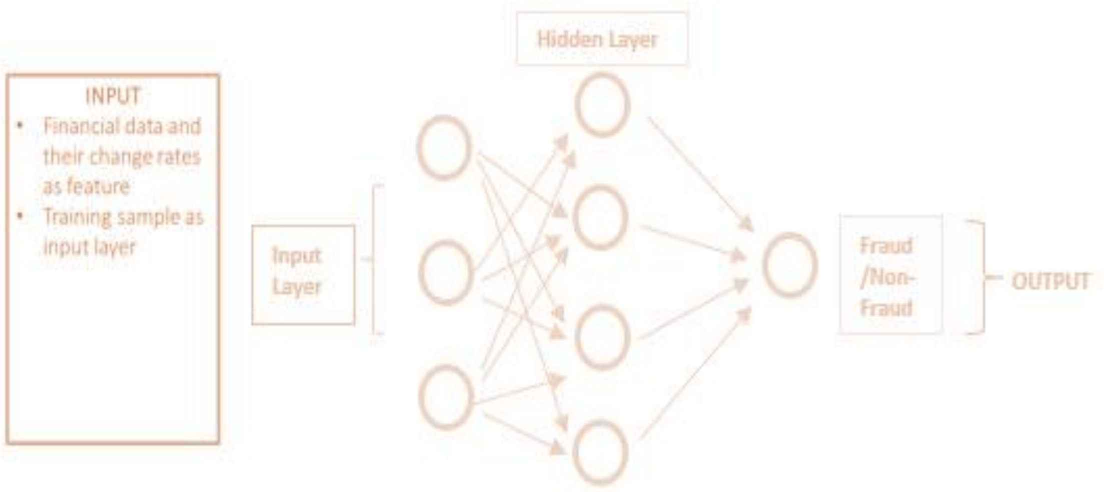


- Support vector machine is a supervised machine learning method for classification and regression.
- Given a set of training samples from two classes, SVM algorithm can find the best decision hyper plane that separate the two classes.
- For those datasets that are non-linearly separable, SVM algorithm can implicitly map the training data into a higher dimensional feature space by a kernel function.

Machine Learning Techniques

Cutting edge technologies can help to detect fraud

5. Artificial Neural Networks


Hidden Layer
INPUT
Financial data and their change rates as feature
Training sample as input layer
Input Layer
Fraud /Non-Fraud
OUTPUT

- Artificial neural networks can be viewed as functions that convert vector of input variables to another vector of output variables.
- A typical method for training artificial neural networks is the back-propagation (BP) algorithm.
- Figure below shows the architecture of the neural network for financial fraud prediction.
- The nodes are called artificial neurons and arranged in three layers: input layer, hidden layer, and output layer. The neurons in adjacent layers are connected with different weights.

- The BP learning algorithm has two phases: a feed-forward stage and a back-propagation stage.
- In the feed-forward stage, each neuron calculates weighted sum of input neurons' values and then applies an activation function to the sum as the output of this node.
- The values flow until they reach the output layer. In the back-propagation stage, the weights between neurons are updated by the learning rule for reducing the discrepancies between actual output and the target value.

- Many other techniques like Genetic Algorithm, Bayesian Belief Network etc. can also be considered for the modeling (predictive modeling) for robustness check. We select the model whichever best suits the data.