

CREDIT CARD FRAUD DETECTION

Team
Members
&
Supervisor

Anuj Srivastava 2015040 Amit Lal Ranjan 2015297 Vicky Kumar 2015273 Vishal Chaudhary 2015277 Deewakar Singh 2015312 Sumit Ramteke 2015256

> Supervisor: Dr. Sunil Agarwal

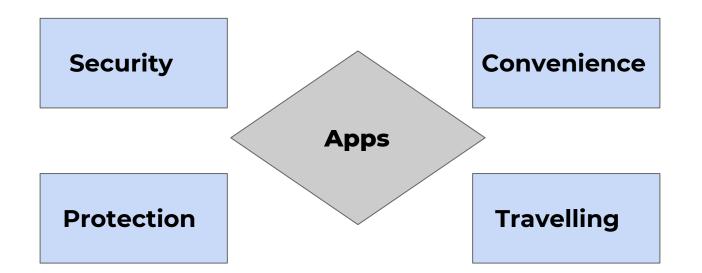


Credit Cards: A credit card is a payment card issued to users (cardholders) to enable the cardholder to pay a merchant for goods and services based on the cardholder's promise to the card issuer to pay them for the amounts plus the other agreed charges















Credit Card Fraud

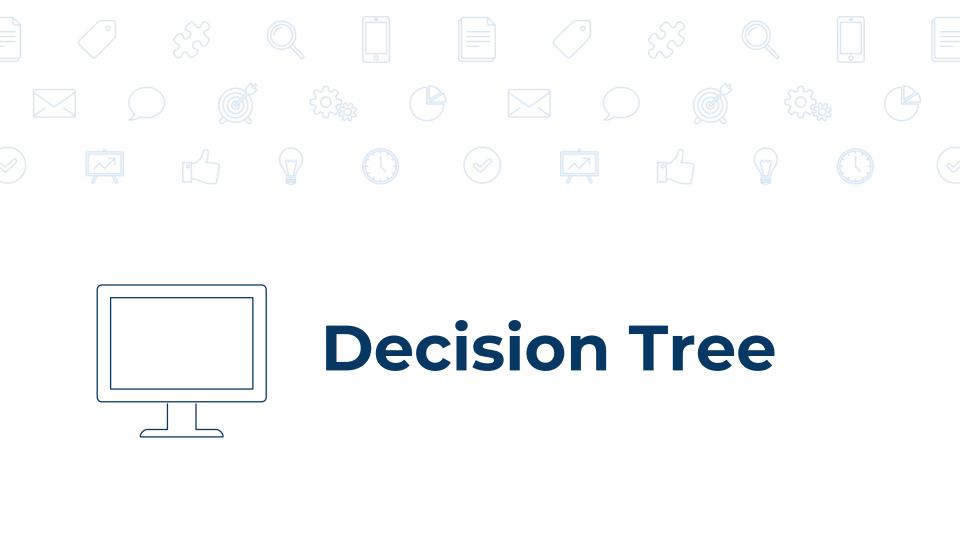
Frauds through phone calls

Computer Intrusion

Counterfeit card Fraud

CNP Fraud





Decision Tree

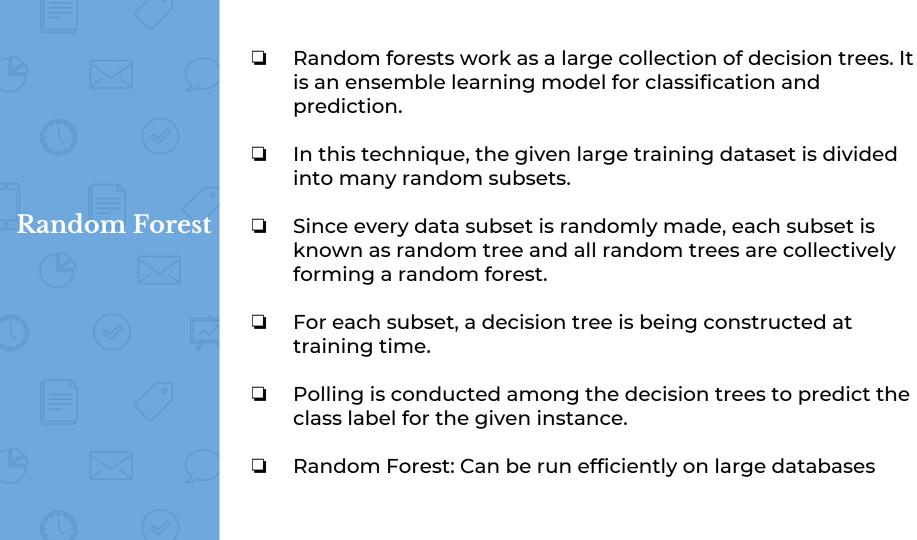
- Decision Tree is a supervised learning technique that produces a tree structure.
- This algorithm employs a top down recursive, greedy approach to build branches without backtracking.
- ☐ ID3 algorithm uses Entropy and information gain to build decision trees.
- Entropy E(A) is a measure of uncertainty. Entropy is calculated for every attribute and is used to measure the homogeneity of data.
- For the entire dataset, information gain I(p,n) is to be calculated where p is total number of class A labels and n is the total number of class B labels



Support Vector Machines

- SVMs are used in the task of classification which classifies and labels the data in feature space.
- SVMs work for the both linearly separable and linearly inseparable data.
- The aim of the algorithm is to design a hyperplane that classifies all data vectors in to classes.
- We can set different parameter for svm-kernel such as Radial, sigmoid, linear, polynomial for better classification and accuracy.
- For every kernel trick, misclassification error/ confusion matrix is to be calculated and whichever gives lesser error rate can be the best model.







Comparision among
Algorithms

Algorithm	Decision Tree	Support Vector Machine (polynomial)	Random Forest (200 Trees)
Precision	0.8217	0.7722	0.7920
Recall	0.7980	0.9176	0.9195
F1-Score	0.8097	0.8387	0.8510



Conclusion

- Decision Tree is having highest precision value.
- ☐ Random Forest with 200 trees having highest recall value.
- ☐ F1-score is highest of Random Forest with 200 trees.
- ☐ Training set ratio Test set is better in (80:20) than (70:30) among all the algorithms.
- SVM with polynomial kernel results best among 4 different kernels.
- F1 Score increases as we increase number of trees in random forest upto certain limit after that remains constant.





After trial it is noted that the K Nearest Neighbors(n = 3) results the best Fl Score compared to Random Forest with 200 Trees.

Algorithm	Random Forest	K Nearest Neighbors
Precision	0.7920	0.811
Recall	0.9195	0.9318
F1-Score	0.8510	0.86772



THANKS.