

A Consideration Point of Fraud Detection in Bank Loans

Group members: Nguyễn Thị Thanh Thúy
Hồ Thị Quỳnh Trâm
Lê Xuân Lâm
Nguyễn Thành Long

John von Neumann Institute - Vietnam National University



JVN Institute
Vietnam National University HCM

Agenda

1. What are Fraud Detection for Bank Loans Application?
2. Current situations
3. Dataset
4. Model Comparison
5. Profit Comparison
6. Conclusion

What are Fraud Detection for Bank Loans Application?

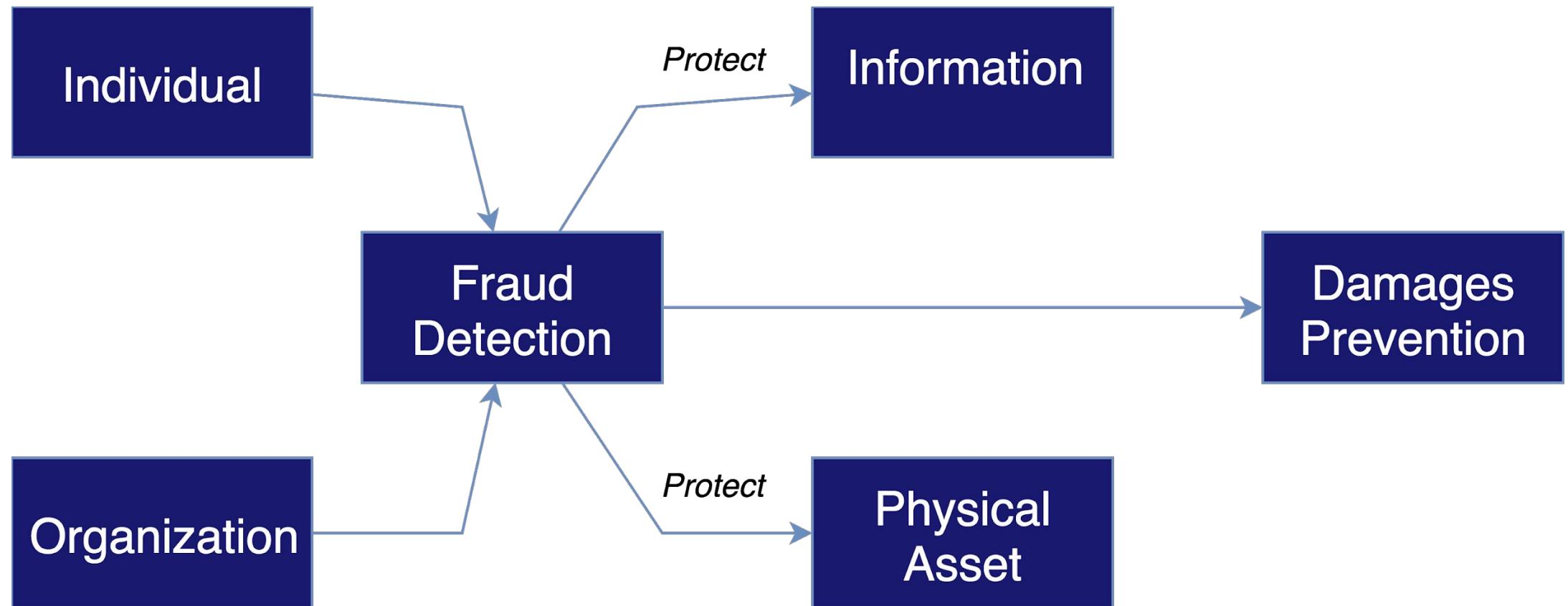


JVN Institute
Vietnam National University HCM

John von Neumann Institute - Viet Nam National University Ho Chi Minh City



Definition



Source: *Student's work*



Frauds Types - Internal Frauds

Transaction Reversal by Tellers

A problem in emerging markets, where bank staff reverse deposit transactions after the customer has left the branch and steal their money.



Account Manipulation

Where employees remove charges or change interest rates on loans or credit limits, usually to benefit their family or friends.



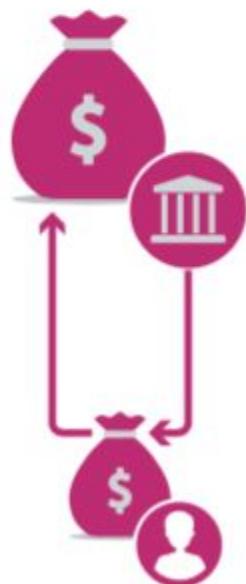
Source: *NetGuardians A-Z of Internal Banking Fraud*



Frauds Types - Internal Frauds

Loan Applications

Another common tactic is to steal a customer's personal details and use them to apply for loans or credit cards in the customer's name.



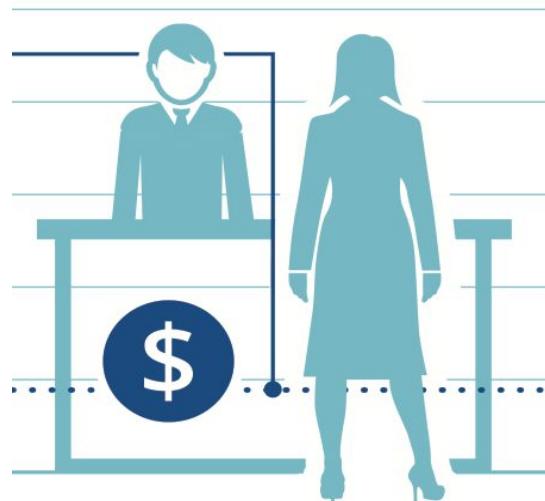
Hiding Losses

Losses on a private banking customer's investment portfolio can be hidden by temporarily taking money from the bank's suspense account – used to hold funds pending reconciliation and allocation to the final account – and transferring them into the customer's portfolio account to increase the balance. After the client meeting, the funds are transferred back to the suspense account. Anti-fraud technology flags unusual behavior around suspense accounts.

Source: *NetGuardians A-Z of Internal Banking Fraud*



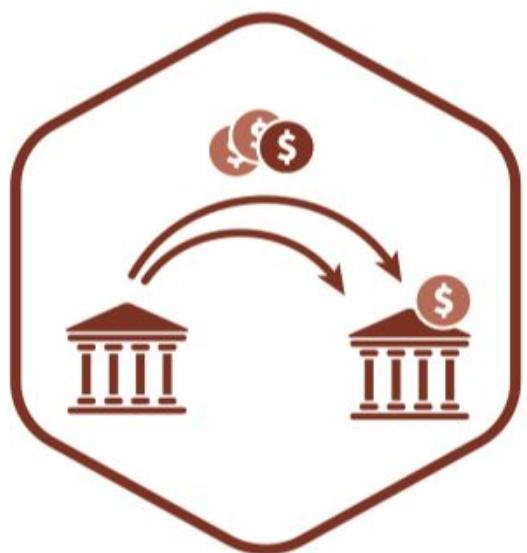
Frauds Types - External Frauds



Loan Frauds



ID Frauds



Money Laundry



Credit Card Frauds

Source: *NetGuardians Banking Fraud*



Other Frauds and Crimes

- Telecommunications Fraud
- Securities Fraud
- Insurance Fraud
- Mortgage Fraud
- Corporate Crime
- Governmental Crime
- Occupational Crime
- Income Tax Evasion
- Money Laundering
- Computer Fraud
- Friendly Fraud

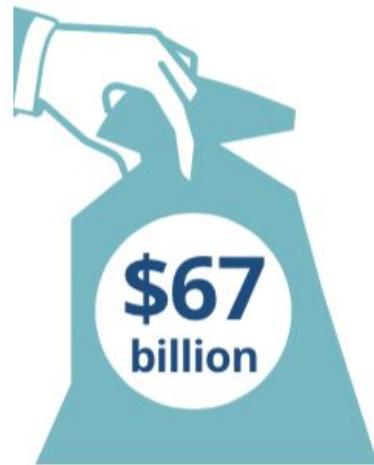
Current situations



JVN Institute
Vietnam National University HCM

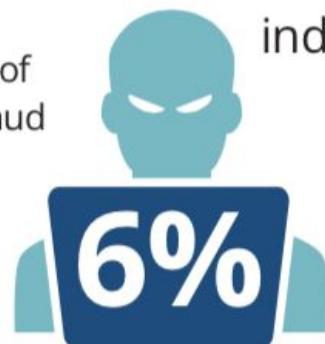


Damages on Financial Institutions



Total value of banking fraud in 2014

(Source:
Association of Certified Fraud Examiners)



70%
of fraud is committed by industry insiders ie employees

of banks' global pre-tax profits were lost as a result of criminal activity*

Three-quarters of financial services companies experienced at least one incidence of fraud in 2012-13



On average, these business incurred losses equivalent to **1.5%** of their revenues
(Source: Economist Intelligence Unit)

141%

increase in the number of financial firms reporting losses of between **\$10m** and **\$19.9m**
(Source: 2014 US State of Cybercrime Survey)

By the Temenos and NetGuardians A-Z of Banking Fraud, NetGuardians



Damages on Financial Institutions



30% of financial services companies have been affected by data theft – individually the most common form of fraud within the industry



IT complexity cited as the top risk factor that organisations face



Fear of **bad publicity** is the most frequently cited reason why cases of fraud are not referred to criminal prosecutors. Plus the cost and time necessary to carry out an investigation

By the Temenos and NetGuardians A-Z of Banking Fraud, NetGuardians



Traditional Fraud Detection Methods

1.

Endpoint-Centric

Analysis of users and their end-points

2.

Navigation Centric

Analysis of navigation behavior and suspect patterns

3.

Account-Centric

Analysis of anomaly behavior by channel

DISCRETE ANALYSIS



Weaknesses

Unable to detect

- Fraud rings
- Fake IP-addresses
- Hijacked devices
- Synthetic Identities
- Stolen Identities
- And more...



Particular vulnerabilities of banks include:



Some banks do not have dedicated fraud teams



No real-time controls



Banks have invested heavily in IT security to counteract threats from malware but far less in behavioral analytics

TRANSACTION SIZE



Controls used to analyze transactions are frequently too narrow and can lead to false positives



Employees failing to observe correct processes create gaps in bank defenses

*KPMG's report Global Profiles of the Fraudster (May 2016)



NetGuardians



Eight reasons why technology wins

Timeliness



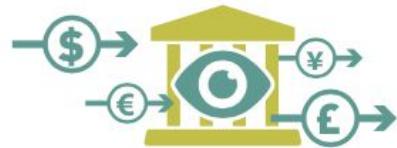
Automated anti-fraud systems can detect possible instances of fraud in real time and block them before they happen

Risk sensitivity



Banks avoid blocking legitimate transactions and identify others that seem genuine but have suspect characteristics

Comprehensiveness



A technology-based approach allows a bank to monitor every transaction in its system – an impossible feat for humans

Focus on individual customer



Banks must understand each customer's behavior patterns so every transaction makes sense when compared to their profile

360-degree surveillance



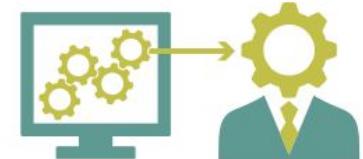
A tech-based approach enables banks to monitor both customers and their own staff through a single system

Record-keeping



Automated fraud-detection systems facilitate record-keeping, helping banks comply with regulatory requirements

Efficiency



Expert staff are freed up to focus on the investigation and verification of suspect cases flagged by the system

Ability to learn



Intelligent systems make it possible to identify new risks before they lead to losses and anticipate new types of fraud



JVN Institute
Vietnam National University HCM

Tech-led solution

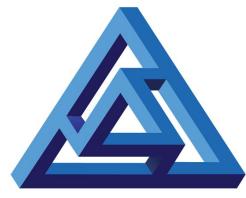


Application



JVN Institute
Vietnam National University HCM

John von Neumann Institute - Viet Nam National University Ho Chi Minh City



Data overview

Title: German Credit Data

Source: Prof. Dr. Hans Hofmann - University at Hamburg

Number of Instances: 1000 (80% for training and 20% for testing)

Number of attributes: 20

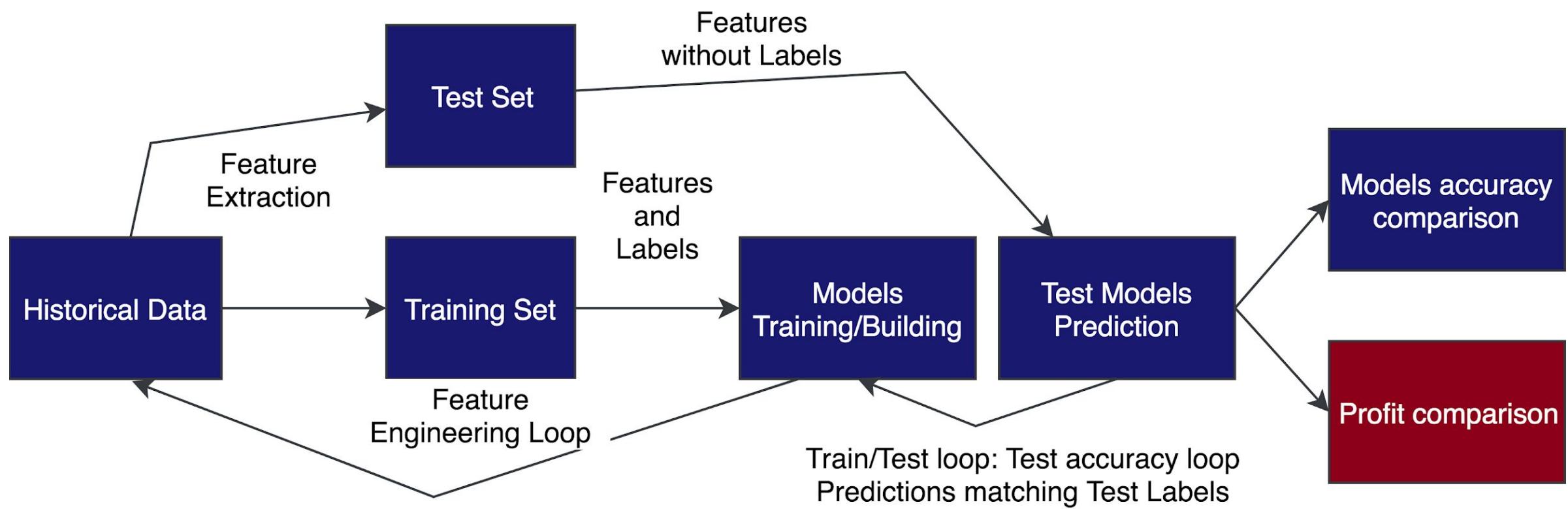
Attribute description:



Order	Type	Name of attribute
1	qualitative	Status of existing checking account
2	numerical	Duration in month
3	qualitative	Credit history
4	qualitative	Purpose
5	numerical	Credit amount
6	qualitative	Savings account/bonds
7	qualitative	Present employment since
8	numerical	Installment rate in percentage of disposable income
9	qualitative	Personal status and sex
10	qualitative	Other debtors / guarantors
11	numerical	Present residence since
12	qualitative	Property
13	numerical	Age in years
14	qualitative	Other installment plans
15	qualitative	Housing
16	numerical	Number of existing credits at this bank
17	qualitative	Job
18	numerical	Number of people being liable to provide maintenance for
19	qualitative	Telephone
20	qualitative	foreign worker
21	Category	Class - Determined BAD or GOOD



Methodology



Source: *Student's work*



	Pros.	Cons.
XGB Tree	<ul style="list-style-type: none">• Highly accurate• Easy to use• Easy to interpret and control	<ul style="list-style-type: none">• Overfitting• Scale-up to larger dataset.
SVM	<ul style="list-style-type: none">• Performs similarly to logistic regression• Performs well with non-linear boundary• Handle high dimensional data well	<ul style="list-style-type: none">• Susceptible to overfitting/training issues



	Pros.	Cons.
RF	<ul style="list-style-type: none">• Decorrelates trees (relative to bagged trees)• Reduced variance (relative to regular trees)	<ul style="list-style-type: none">• Not as easy to visually interpret
Logistic	<ul style="list-style-type: none">• Low variance• Provides probabilities for outcomes• Works well with diagonal (feature) decision boundaries	<ul style="list-style-type: none">• High bias

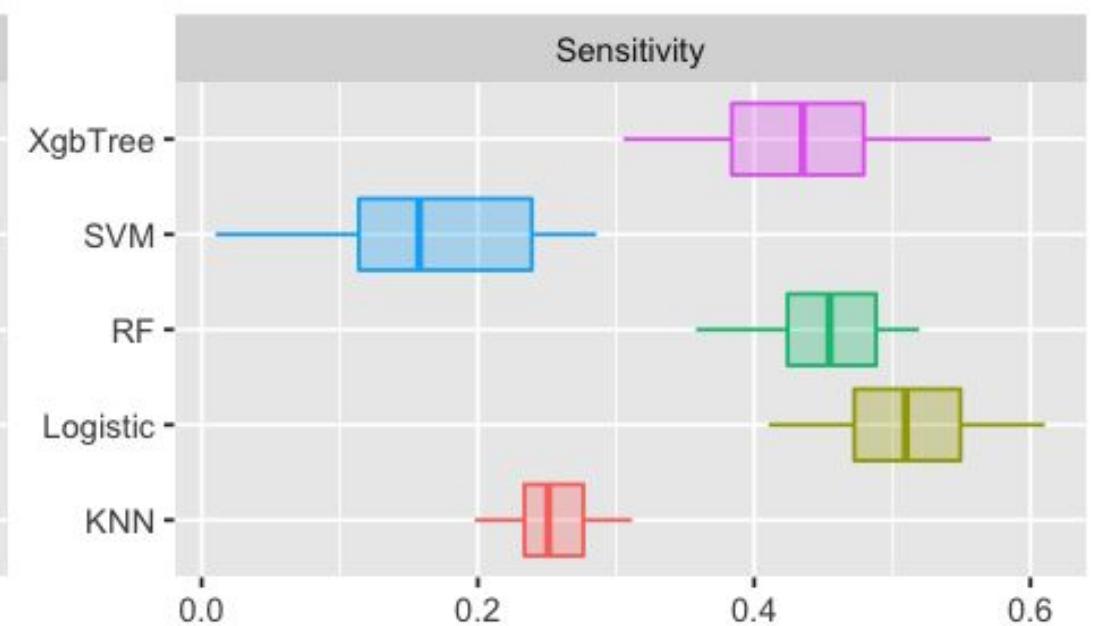
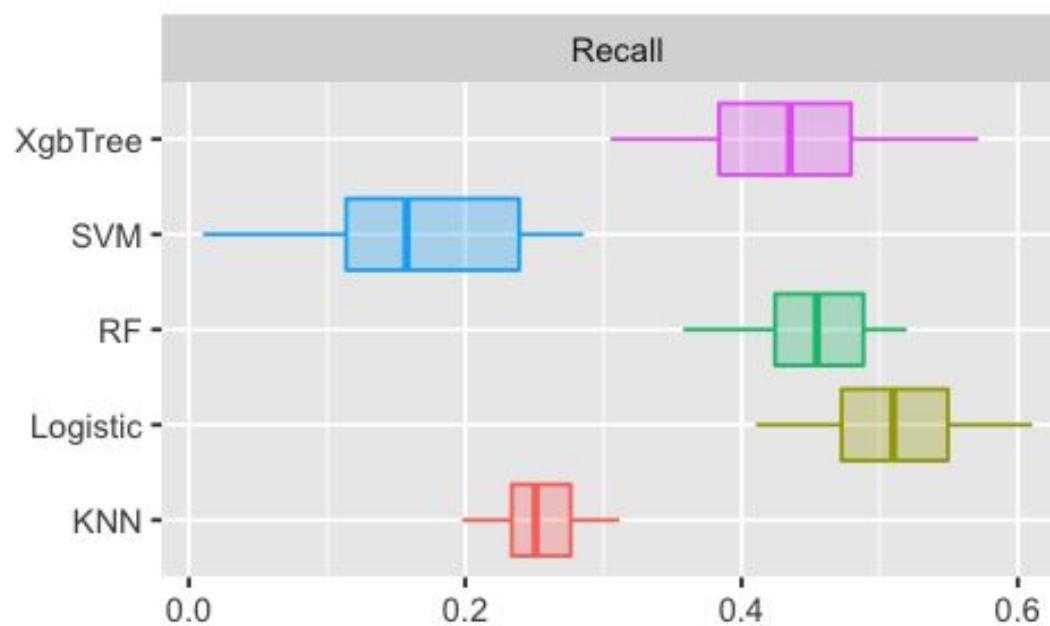
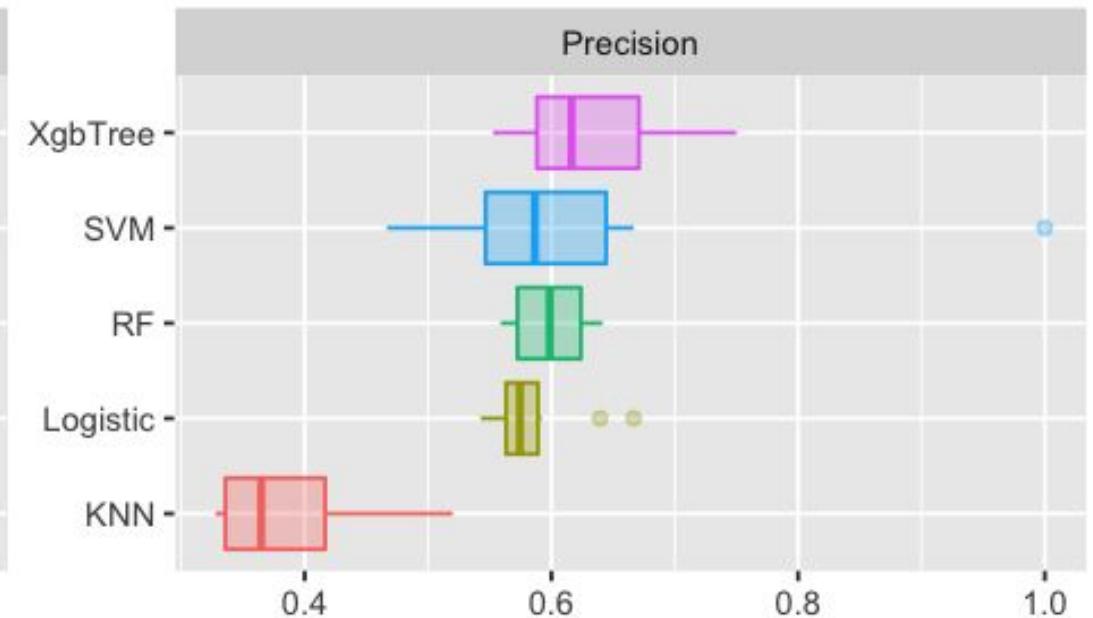
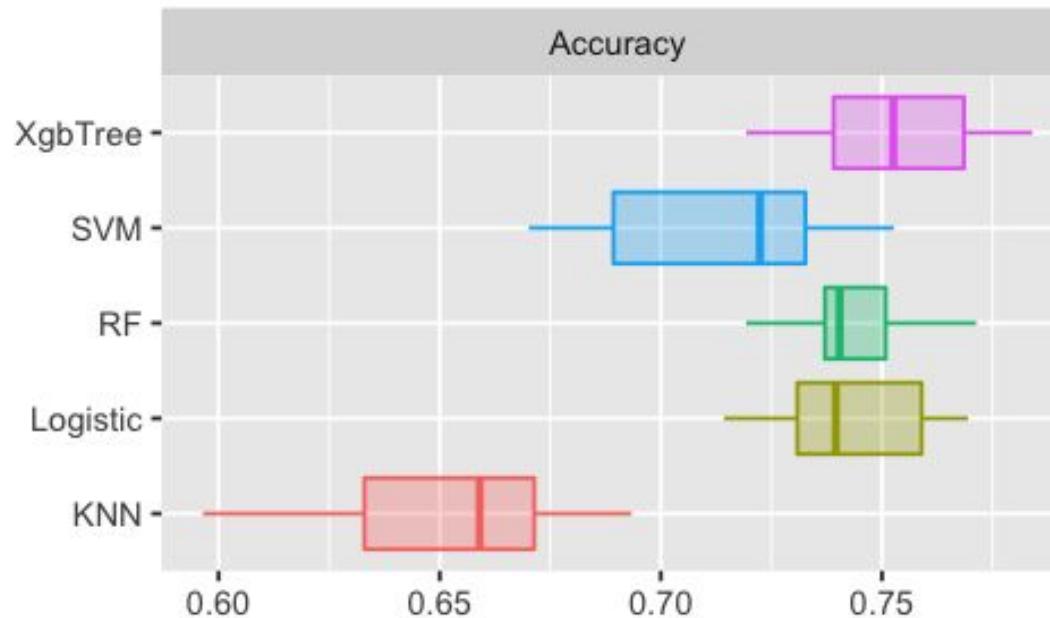


	Pros.	Cons.
KNN	<ul style="list-style-type: none">• Simple to understand and easy to implement• Works easily with multiclass data sets• Does not assume any probability distributions on the input data• Can quickly respond to changes in input.	<ul style="list-style-type: none">• Sensitive to localized data• Be able to suffer from skewed class distributions• High Computation time• Many dimensions.



Models Comparison

Evaluation and Comparison of Model Performance





Models Comparison

To be more precise

	Accuracy	Precision	Recall	Sensitivity
XgbTree	75.20	62.95	43.03	43.03
Logistic	74.36	58.44	50.91	50.91
RF	74.36	59.68	45.30	45.30
SVM	71.41	61.85	16.52	16.52
KNN	65.34	38.59	25.39	25.39



Profit Calculation Mechanism

- Example of 200 sample of XgbTree Model

Confusion Matrix		Actual	
		Bad	Good
Prediction	Bad	24	16
	Good	36	124

- The model resulted in approved 160 applications
- However, there are 36 fraud cases. Created Losses to Bank.

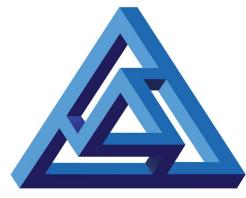


Profit Calculation Mechanism

- Create N output result (in this case 100) of each model based on Training Data Set
- Assumed the margin of each loan - in this case 30%
- Simulated the profit from model output results generated by each other from Loan Amount - in this case simulated 1,000 times on Test Data Set
- Profit = Rate*(True Positive - False Positive)

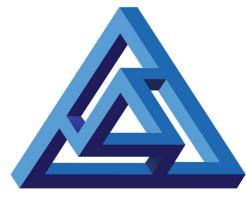
In short, beside the Accuracy Rate:

- When model got high False Positives bank will loss money**
- When model got high False Negatives bank will loss the expected profit**



Profit Comparison

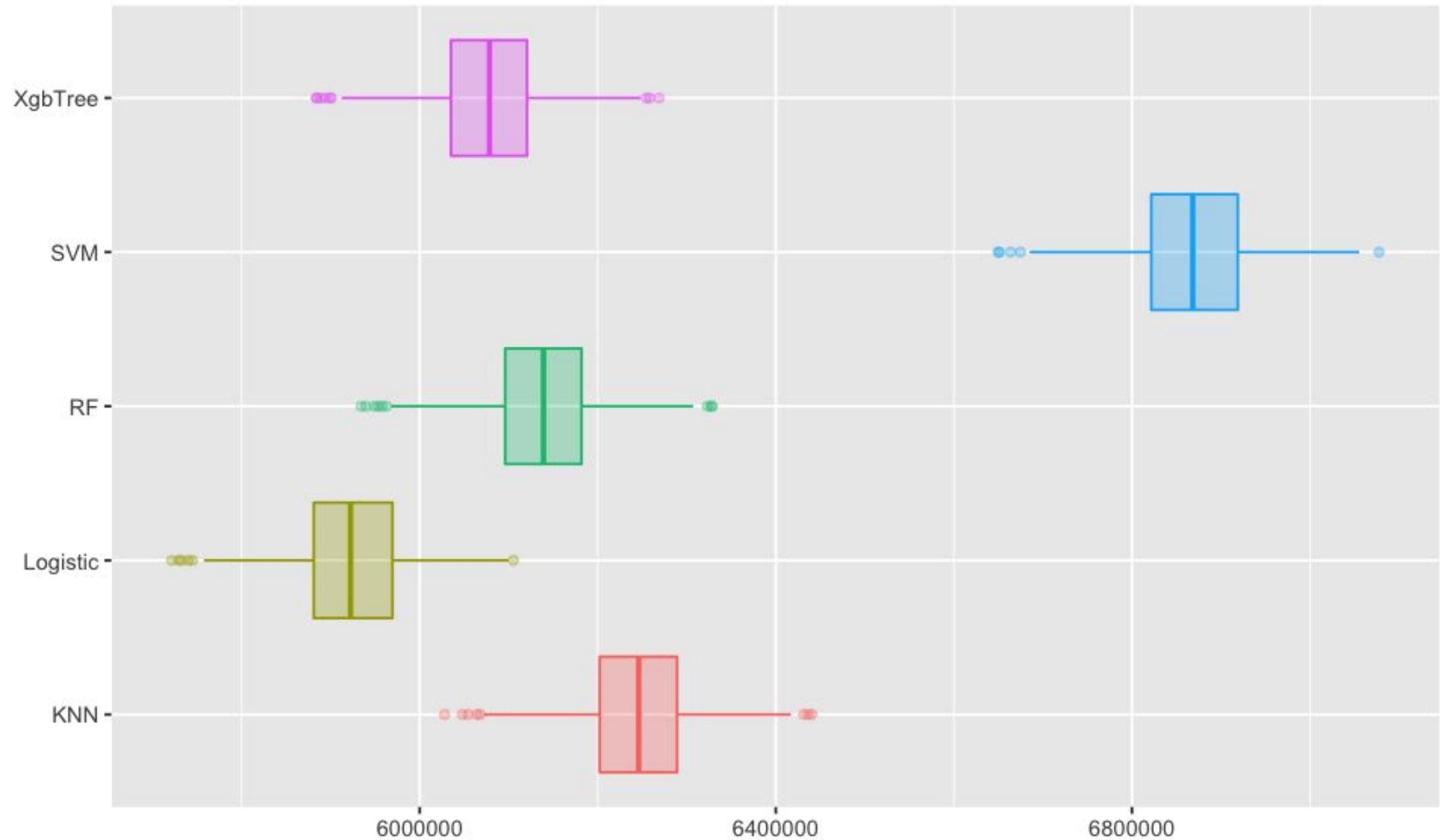
	mean	median	min	max
SVM	6,870,087	6,868,312	6,649,531	7,077,496
KNN	6,246,711	6,246,003	6,027,972	6,440,565
RF	6,139,046	6,139,071	5,934,271	6,328,565
XgbTree	6,079,105	6,078,347	5,883,934	6,268,982
Logistic	5,925,101	5,922,345	5,721,494	6,105,366



JVN Institute
Vietnam National University HCM

Profit Comparison

Simulated Profit Based Monte Carlo Method with Interest Rate is 30% for Test Data Sets





Conclusions

- Fraud Detection Methods have been played a important role in the modern Financial World.
- Beside the normal concentration on Accuracy Ratios in predictive modelling field, we also need to consider other specific critical criteria in different industries
- In this case, Financial Institutions focus on Profit so a good model is not the highest accuracy one, but the most profitable one.
- This presentation is the very early investigation on the Fraud Detection Models vs. Profit Model of Banks. In order to go further, we need to dig deeper into the Accuracy Improvement of the Models and the precisely of Bank Profit Caculation

Thank you for your listening



JVN Institute
Vietnam National University HCM

John von Neumann Institute - Viet Nam National University Ho Chi Minh City