FRAUD DETECTION IN BANKING

ISM-6331 INFORMATION SECURITY SYSTEMS

Research Paper

TABLE OF CONTENT

FRAUD DETECTION

Fraud detection is a process that identifies and prevents fraudsters or cyber-hackers from obtaining money or property through false means. It is a series of practices undertaken to block these criminals to achieve their goals. Fraud detection can be found across multiple industries, some examples are insurance, government, banking, and healthcare. Generally, activities such as cyberattacks, money laundering, or identity theft, are all labeled as fraudulent. Therefore, is vitally important for corporations to implement modern techniques and risk management strategies to fight growing fraudulent transactions across diverse platforms. These cyber-techniques apply adaptive and predictive analytics to create a fraud risk score, along with real-time monitoring of events.

Very recently, the banking industry has prioritized cybersecurity highly and it makes sense to reinforce credibility and trust toward customers. After some research, I've identified several factors that demonstrate the growing necessity of cybersecurity in the field.

First, everyone seems using digital payment methods like credit or debit cards. Perhaps having strong security safeguards, might protect the privacy and confidential data.

Second, data breaches can have devastating consequences for any bank. If a breach has been caused by an inferior security solution, surely it will lead its customers to move their business base elsewhere.

Third, most of the time, when a bank's data is compromised, the customer loses time and money. Not to mention the unpleasant feelings the recovery would bring.

Forth, the inappropriate use of confidential data might be extremely harmful. Although cards are revoked and frauds dealt with, stolen sensitive data may still be exploited against customers.

Lastly, financial institutions need to be more cautious than most other firms. Since they retain valuable personal data, there is a high chance that they could be compromised if not safeguarded correctly.

THREATS

Cybercrimes have increased frequently over the past decades to the point where it is thought that they are one of the most significant hazards to the financial sector. Hackers have improved their technology and expertise, making it difficult for any bank to thwart the attack consistently. The following are just some threats to the bank's cybersecurity:

PHISHING ATTACK: perhaps one of the most frequent assaults in the banking sector. It can be used by a hacker to enter a financial institution's network to steal confidential data or conduct a more severe attack such as an APT "Advanced Persistent Threat, a sophisticated, sustained cyberattack in which an intruder establishes an undetected presence in a network to rob sensitive data over a prolonged time".

SPOOFING: hackers use a close site to lure honest customers and plunder their credentials. By posing as a financial website, hackers may architect a layout that resembles the original one in both appearance and functionality, plus establish a domain with a modest modification in spelling or extension. Then, the customer can access the duplicate via a third-party messaging service, text, or email. Once baited, the hacker obtains access to a customer's login information.

RANSOMWARE: The cyber threat encrypts confidential data and prevents institutions from having access to it until the company pays a ransom. It seems a lot of banks faces ransom attacks in the past. On top of that, such a cyberattack affects cryptocurrency as well. Due to their decentralized structure, fraudsters have easy access to breaking into trading systems and stealing money.

TROJAN: it is used to designate several dangerous tactics cybercriminals use to cheat their way into secured data. Until it is installed on a device, a trojan looks like trustworthy software. Unfortunately, it is a malicious computer application created to access private data, processed by online banking systems. Generally, such a program offers a backdoor that enables access to a targeted computer from the outside.

IDENTITY THEFT: it happens when a cyber hacker gets in possession of stolen personal information and uses it to commit fraud. Then, criminals can open bank accounts under someone else's name, take out credit loans without having an interest in repaying, and launder money without linking transactions to their identities. Generally, this type of threat comes from standard cybersecurity failures such as data breaches, phishing, or fake job advertisements.

ARTIFICIAL INTELLIGENCE APPLIED TO FRAUD DETECTION

Most financial institutions still use rules-based systems with manual evaluations to detect fraud. Until very recently, these systems were doing their job. However, with hackers increasing knowledge and sophistication, the results traditional systems give, are becoming inconsistent. Perhaps Machine Learning can overcome such shortcomings. An ML model analyzes huge sets of data using algorithm complexity to identify patterns. The intuition stems from seeing fraudulent transactions showing patterns differently from genuine ones. Thus, the algorithm detects illicit activity faster and more accurately than traditional rules-based systems, because programmed to deal with large datasets. While humans unknowingly overlook pieces of information, AI can be trained to analyze even the most seemingly unrelated information to find a pattern.

It all starts by gathering and categorizing as much historical data as possible. All-inclusive information about legitimate and fraudulent transactions that are labeled as good or bad. Then, the training data is used to teach the model how to predict whether a certain customer or transaction is fraudulent or not. For any ML program to be successful, it requires to have as much frauded data as possible, so to feed the algorithm with a lot of references to learn from. Once the training is completed, the model becomes specific to the business and can be considered ready to use in a bank's fraud management framework. Sadly, the model must be updated occasionally, since it is not impeccable. But it offers a good solution for fraud detection.

Why should an institution implement an ML model? Because it offers a plethora of benefits; **Speed** – can evaluate vast amounts of data in a few minutes. Plus, it can perpetually collect and analyze new data in real time. **Efficiency** – It can perform repetitive tasks and detect subtle changes in patterns across huge datasets. It can analyze hundreds of thousands of payments per second, way more than several human analysts can do. **Scalability** – As the number of transactions increases over time, the pressure increases as well. Thus, additional costs and time are spent on analysis. With a machine learning algorithm, the more data the better. It improves as more data comes in, enabling it to detect fraud faster and with more accuracy. **Accuracy** – a model can be trained to detect patterns across seemingly insignificant data or identify non-intuitive trends which would be hard, or perhaps impossible, for analysts to snatch. As a result, there will be fewer false positives and frauds that go undetected.

ADDITIONAL SOLUTIONS

Threats are constantly evolving; hence, the bank must take action to protect itself. Normally, hackers tend to adapt to new defenses, by developing more sophisticated tools and strategies to compromise security. With all that in mind, it is vital to have a collection of security tools – in addition to machine learning models – to protect at best sensitive data and systems. More defense lines are listed below:

Network security surveillance: it is a perpetual network scanning to detect signs of dangerous or intrusive behavior. Frequently employed with firewalls, antiviruses, or IDS (intrusion detection systems). Generally, it allows for either manual or automatic scanning.

Risk Management: data integrity, data security, security awareness training, and risk analysis, are all part of the package.

Software security: Safeguard apps are essential to business operations. They possess features such as helping synchronize security policies with file-sharing permissions and multi-factor authentication.

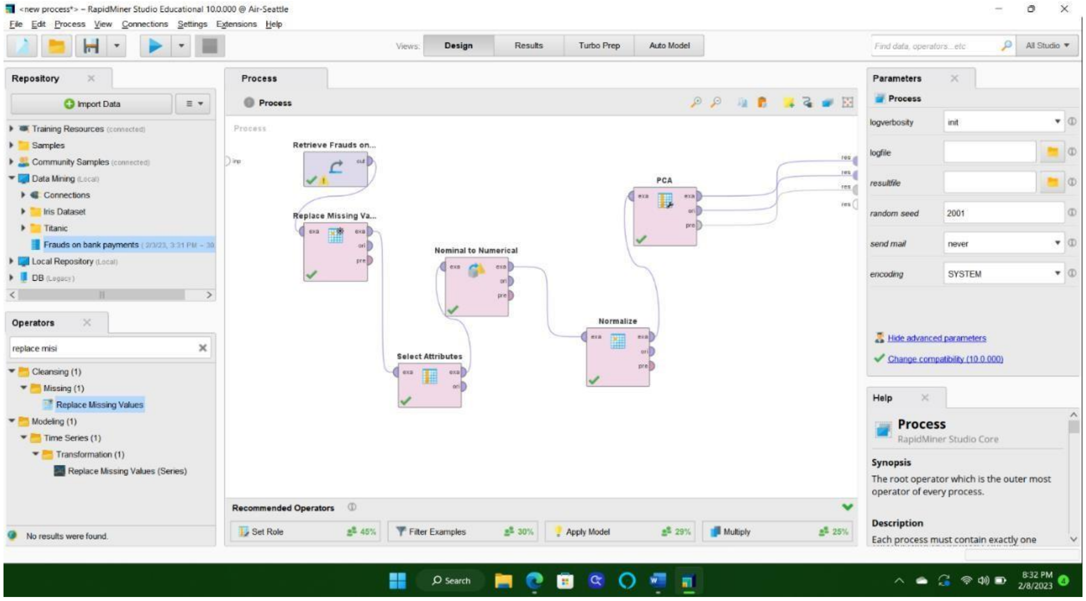
Protecting Critical systems: wide-area network connections help avoid attacks on massive systems. Plus, it monitors all programs and performs security checks on users, servers, and the VPN.

TRAINING A MACHINE LEARNING MODEL WITH DATA MINING

I've used an intuitive dataset from Kaggle.com large enough to train the model successfully. The dataset represents the bank transactions of a European bank and contains the following information:

- 4412 Customers.
- 650000 Transactions collected in the past 6 months.
- 50 Merchants
- 15 Product categories

RapidMiner has been applied to run the experiment, with the CRISP-DM framework approach to complete the project. The first one is a powerful data mining tool that enables everything from data mining to model deployment, and model operations. Whereas the second one, CRISP-DM stands for a cross-industry standard process for data mining. In essence, the most common framework adopted throughout the industry, and it is employed as a guideline for DM projects.

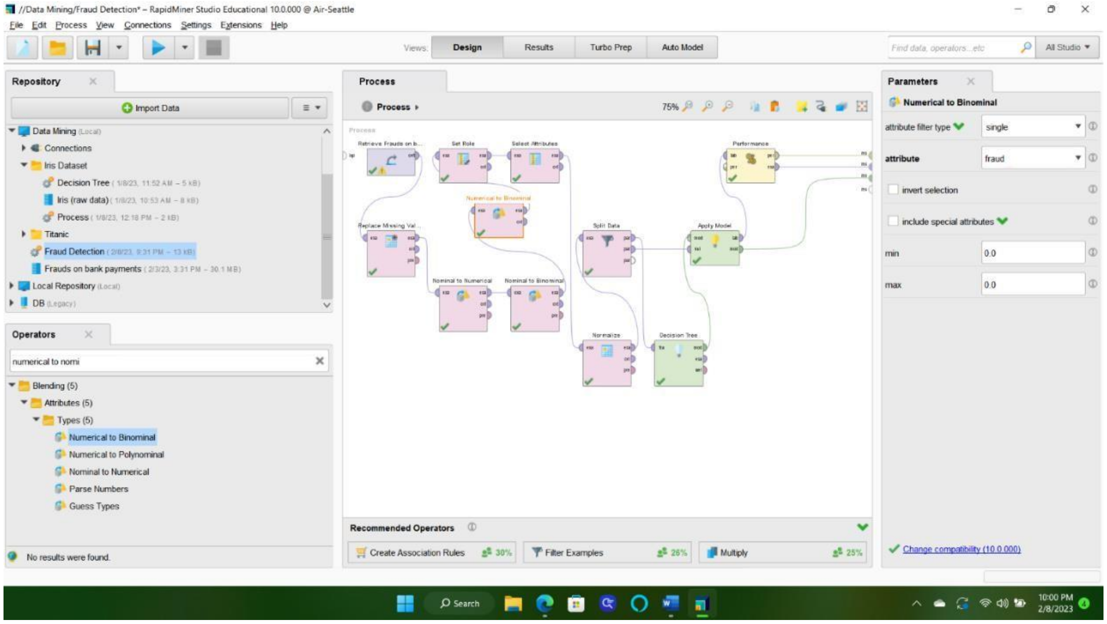


logverbosity: init
logfile:
resultfile:
random seed: 2001
send mail: never
encoding: SYSTEM

Process
RapidMiner Studio Core

Synopsis
The root operator which is the outer most operator of every process.

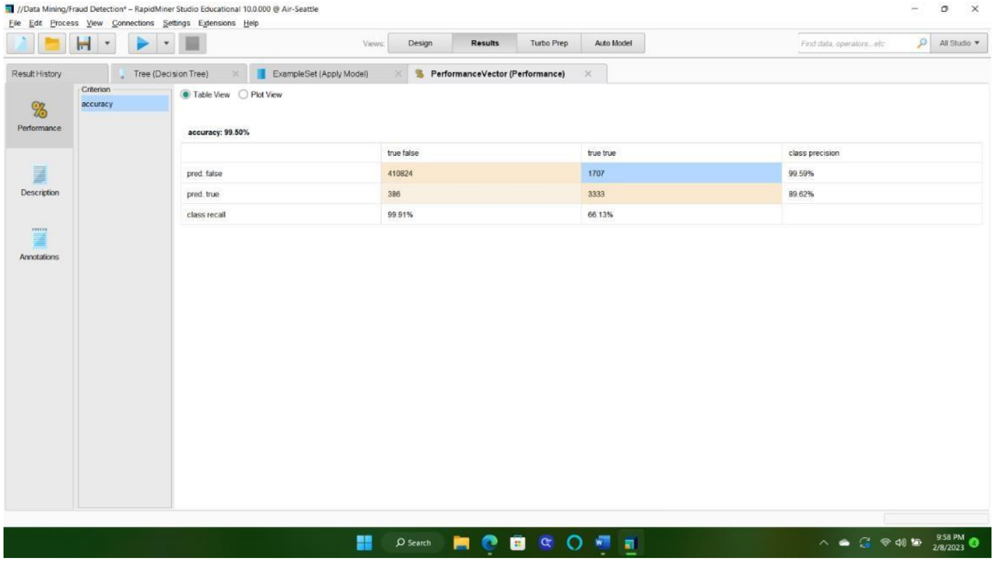
Description
Each process must contain exactly one



Parameters: Numerical to Binominal

attribute filter type	single
attribute	fraud
□ invert selection	
□ include special attributes	
min	0.0
max	0.0

A decision tree model brought these results. RapidMiner represents the outcome of a training session in a confusion matrix, where the scientistic can check false positives and false negatives. The focus is the blue rectangle. It tells exactly how many frauds have been passed undetected while inspected by the machine learning model. Sadly, the model shows a precision of just 60%. In essence, only 6 out of 10 bank frauds are blocked.


//Data Mining/Fraud Detection* – RapidMiner Studio Educational 10.0.000 @ Air-Seattle
File Edit Process View Connections Settings Extensions Help
Views: Design Results Turbo Prep Auto Model
Result History
Tree (Decision Tree)
ExampleSet (Apply Model)
PerformanceVector (Performance)
Performance
Description
Annotations
Criterion
accuracy
● Table View ○ Plot View

accuracy: 99.50%

	true false	true true	class precision
pred. false	410824	1707	99.59%
pred. true	386	3333	89.62%
class recall	99.91%	66.13%	

P.S: THE TRIAL AND PRESENTATION REQUIRE MORE WORK; A BETTER MODEL MIGHT BE TRIED TO IMPROVE ACCURACY IN FRAUD DETECTION.

CHALLENGES

Some contributing elements have presented a significant challenge to applying digital security in the financial industry. Some are mentioned as follows:

Lack of Knowledge – the general public's understanding of cybersecurity is relatively low, and very few institutions have invested in raising awareness among stakeholders.

Tight budget & Poor Mgmt. - cybersecurity tends to receive short budgetary shrift due to low priority from top management.

Weak cyber protections – the core component of cybersecurity is identity and access management; a company should avoid that scenario where a hacker accesses a business network with just one compromised login.

Ransomware – criminals have started to employ various techniques to mitigate being identified by endpoint protection code that concentrates on executable files.

Smartphones & Apps – almost all financial institutions conduct business through mobile devices. The base grows over time, making it the best target for hackers.

Social Media – criminals have increased their exploitation because of social media adoption. Uninformed customers are more likely to experience their data getting abused by a criminal.

REFERENCES

- Fraud Detection on Bank Payments | Kaggle
- Financial crime and fraud in the age of cybersecurity (mckinsey.com)
- 24 .pdf (armgpublishing.com)
- 5 Biggest Threats To Cyber Security In The Banking Industry In 2022 | DeskAlerts (alert-software.com)
- RapidMiner | Amplify the Impact of Your People, Expertise & Data