UNIVERSITI TEKNOLOGI MALAYSIA

SECP1513 SECTION 08
TECHNOLOGY AND INFORMATION SYSTEM

FRAUD TRANSACTION DETECTION SYSTEM

Prepared for:

DR LAYLA RASHEED ABDALLAH HASAN

Prepared by:

FULL NAME	MATRIC NO.
NURUL ‘AFIFAH BINTI MOHAMAD YUSOF	A21EC0120
AIN SAFIAH BINTI MANAN	A21EC0155
LUQMAN HAKIM BIN MD SAID	A21EC0050
KAGINESWARAN A/L TAMIL VANAN	A21EC0035

Table of Content

1.0 Introduction

In our current world of ever-evolving technologies, people and companies are incorporating technology into almost everything, from the field of agriculture to the field of medicine such as delivering a baby, simulations of operation procedures, and many more. In addition to the COVID-19 pandemic, people have been more dependent on digital platforms and technology in daily activities. For example in making bill payments, purchasing daily necessities, etc.

The adoption of digital technology did help humans in a lot of aspects such as improvement in activities and innovation efficiency. However, despite all of the benefits it brings, these dependencies on technology can leave us vulnerable to much unwanted and malicious intent from third parties like cybercriminals. Cybercriminals are those who use technology to do malicious activities to gain benefit for personal or group purposes. Most of the cases are for financial gain.

Due to that, here are some common disadvantages of digitalization. Firstly, the lack of data security can cause the organization to lose information not only about the organization but also the personal information of the employers, employees, and even the customers which can be dangerous if it lands in the wrong hands. It may result in a financial loss at the enterprise or personal level. The worst thing is it may cause the organization to lose the trust of their clients and customers. Besides that, with the current growth of data received by organizations around the world, some cannot keep up with the increasing amount of data which makes them unable to set up proper cybersecurity procedures to protect all that data and detect any fraudulent data effectively.

These are just a few of the long list problems faced due to the lack of appropriate steps taken when using technology. Therefore, the step of prevention needs to be taken to prevent unwanted incidents.

It is known that technology is beneficial in many ways with the appropriate measures taken to prevent any malicious intention from the internal or external entities. The implementation of IR4.0 technologies plays a role in achieving those benefits. With this, our objective for this project is to create a system that can detect fraudulent activities in the financial aspect with the aid of IR4.0 technologies and appropriate Cloud Computing architecture.

2.0 Content of Report

2.1 Problem statement by a potential client

As Bank A started to grow, they received complaints from customers that they lost their money from their accounts, and some of them did receive unauthorized credit card charges. However, they did not make those transactions. Hence, Bank A wants to detect payment and identity fraud to prevent and stop fraudulent transactions and activities. The fraudulent activities here means the customers activities throughout the bank such as accessing the account information.

2.2 Selection of the Fourth Industrial Revolution, 4IR technology

Fourth Industrial Revolution technology is the latest revolution in the industry which is involved in the fusion of the digital, physical and biological aspects in the merging of new technologies. These technologies involved Artificial Intelligence (AI), Internet of Things (IoT), 3D printing, etcetera. The technologies involved in the 4IR did help in solving the problem faced by humans such as creating a model. With the invention of 3D printing, creating a 3D model would be an easy task and might not need direct human interaction with the model. It also has been an aid in the making of innovations and creations of new technology and systems.

For this project that we would like to create, we intend to create a system that can detect fraudulent activities or transactions throughout a bank. The creation of this system includes the 4IR technology such as Cybersecurity, Artificial Intelligence, Machine Learning, and Cloud Computing.

2.2.1 Cybersecurity

Cyber security involves the policies and application of technologies used in protecting digital components, systems, networks, programs, devices, and data from cyberattacks which can bring harm to the victims mostly on the economic and financial aspects. This is because most cyberattacks that occur are intended to gain profit.

A fraudulent transaction or activity may occur in various ways. It can happen with account takeover by the cybercriminal, getting access by stealing the bank account information, or many more ways. There are a lot of cyberattacks that may be involved in the process but here we would like to emphasize Phishing and Malware attacks.

Phishing is a cybercrime frequently used to gain sensitive information from users, such as usernames, passwords, and credit/debit card numbers. It occurs when a cybercriminal assumes the identity of a legitimate entity through an email, instant message, or text message and convinces a victim to open those email, instant messages, or text messages. The recipients are deceived into opening a malicious link, which results in the installation of malware, the freezing of the victim's machine as part of a ransomware attack, or the revealing of sensitive information. In other cases, it might occur when the victim is convinced to give out their personal information such as bank information through those platforms.

Malware is a generic phrase that refers to any malicious software designed to infect or exploit any programmable device, service, or network. Malware is classified into numerous categories, including computer viruses, worms, Trojan horses, ransomware, spyware, adware, rogue software, wiper, and scareware. It is used to extract data or information from victims. This data may include sensitive data such as financial information, healthcare records, personal emails, and passwords. Else, it can be used to gain access to the victim's side such as the victim's accounts or devices.

With the related data, information, or access gained by those cybercriminals. Then, they might be able to make a fraudulent transaction. As the final protection for the customers, the bank can implement cybersecurity on the transaction process by detecting and stopping fraudulent transactions.

2.2.2 Artificial Intelligence and Machine Learning

In detecting fraudulent transactions or activities, Artificial Intelligence and Machine Learning technologies can be used. Artificial Intelligence, AI is the simulation of human intelligence processes by machines, particularly computer systems, and Machine learning is a subset of artificial intelligence that has the ability to approach data analysis that involves building and adapting modes, which help the programs to improve their performance over time. Machine learning also contains the algorithm that will help the models to adapt and improve the ability to make any predictions.

AI and Machine learning can be used in transaction or activities verification because it is possible to swiftly evaluate many types of events and identify a variety of fraudulent data, ranging from false transactions to identity fraud or theft that could lead to a financial loss to the victims as an individual or as an organisation. These systems may improve with time since Machine learning is used. Machine learning learns from the data received from the previous cases. This machine learning will then segment and extract the required features from the collective data. From this, machine

learning can learn and train itself to detect a similar type of fraud or fraud that are not the same but have the similarity with the previous cases.

In the case of fraud detection, a machine learning model and example dataset from the previous transaction are being used. Perhaps, the transaction amount, the IP address of the device used for online transactions, location, or others. The fraud detection deploys a machine learning model and the example dataset to train the model to recognize fraud patterns. By comparing the data transaction with the fraud pattern, the model is able to recognise or predict whether the data is fraud or not. The model is dynamic since it is able to predict the new fraud pattern. It is possible with the ability of the model to self-learn with the new data. While detecting fraud data, the data received by the model is used to improve the fraud pattern.

2.2.3 Cloud Computing

For implementing the fraud transaction detection system, we would like to use Cloud computing technology to support the architecture of the system. Cloud computing is the practice of using a network of remote servers hosted on the internet to store, manage, and process data, rather than a local server or a personal computer. Examples of cloud computing providers are Amazon Web Services, Microsoft Azure, and Google Cloud. Most cloud computing providers apply pay-per-use services.

The use of cloud computing technology in implementing the system can bring a lot of benefits such as cost optimization since we don't need to build and maintain the hardware infrastructure of the system. The implementation of the on-premises infrastructure might result in a capitalized cost where cost needs to be covered or paid even if the service is not in use. In managing the hardware infrastructure, a lot of things need to be considered, the security aspect, maintenance, the storage of the server, the service needs to be used, and much more.

By using cloud computing, we didn’t need to worry about the hardware infrastructure since it is being taken care of by the cloud service provider. In contrast, if we decided to use the on-premises infrastructure, the risk for the hardware infrastructure to be attacked are high, especially without a high-security precaution which may result in the system malfunction and can’t operate properly. Even if there are no attacks on the hardware infrastructure, constant maintenance needs to be done. The utility cost such as the bills also needs to be considered.

The transaction process in and out of the bank involved a huge amount of data and an inconsistent data flow of numbers. Therefore, in the deployment of the system, we need to have a scalable infrastructure to prevent the system malfunction or when the system is down since it can't support any sudden increase in the amount of data. Hence, cloud computing can be used since it is scalable.

In this project, the fraud detection system, AI and machine learning are being used. Therefore, cloud computing is suitable for system deployment since the cloud service provider AWS did provide services for AI and machine learning implementation.

3.0 Architecture planning & design(luq & kagi)

3.1 Design of the current system

3.1.1 Current system infrastructure

The bank uses the reputation list technique to prevent fraud activities from happening. The method works by matching up the data in the system like email, IP address, credit card number and phone number with the reputation database for any known blacklist identity. If the data matches the database, it will detect as fraudulent activities. The limitation is that the reputation database can only capture a static identity and not a dynamic one. For example, nowadays, people can create multiple one-time email addresses that can be used to commit fraud activities. Still, the system can only block the person if it uses the same email address, not for a different email address.

Next, the bank lacks identity verification. Any transaction will proceed as long as the user has logged in to the system using only a username and password. Hackers nowadays can get information regarding customers' usernames and passwords using various methods like phishing. Therefore, it is easy for cybercriminals to manipulate the customer's data because they have full access to the account.

The bank also uses a traditional data centre that requires physical facilities to store information. Therefore, if the bank wants to scale up its data storage because of the non-stop incoming data, the bank will need to pay expensive costs to buy a new server that matches the data storage. Moreover, the system will lack storage objects required to accumulate data for processing, which will limit the system's work.

3.2 Planning and design of the new system

3.2.1 New system infrastructure

The customer will be asked to provide facial recognition for pre-payment authentication. The scanned facial recognition will then be used to verify the identity of the user before beginning the process of transferring, paying bills, checking bank balances and many more. After that, the user will be able to key in their account information which will then trigger the lambda function and will invoke the machine learning inference endpoint through Amazon Sagemaker. Amazon Sagemaker then predicts the data to be fraud or not fraud and sends the processed data to Amazon Simple Storage Service (Amazon S3) to be stored for future uses. If the fraudulent data is being detected, then the system will give the alarm to the bank to block the activities.

3.2.2 Cloud services being used

- Storage - Amazon S3

Amazon Simple Storage Service (Amazon S3) is the main storage object used for the system. It will hold all the data before and after being processed such as the scanned facial recognition, the data about account information before being processed and the data after the account has been predicted to be fraud or not fraud. The Amazon S3 can also store any amount of data so that it will not cost any additional cost to scale up the storage.

- Compute - AWS Lambda

AWS Lambda is a serverless and event-driven compute service that can run code without provisioning or managing servers. AWS Lambda will only function when it is triggered. The data will be received by AWS Lambda and upon receiving the data it will trigger the lambda function which then invokes the machine learning inference endpoint through Amazon Sagemaker. Moreover, it will run the code without thinking about the servers thus not worrying about the computing speed of the system.

- Machine learning services - Amazon Sagemaker (Amazon Fraud Detector)

By using Amazon Sagemaker, the system will use Amazon Fraud Detector to build a fraud detection machine learning model that is customized to your data in a few clicks using a fully automated process. The detection logic works by combining your model with decision rules to turn model scores into actionable outcomes. For real-time fraud detection, call the Amazon Fraud Detector API with online event data to receive fraud predictions. Our system will be primarily using this service to identify suspicious online payments, It detects potential fraud to swiftly and correctly find abnormalities and patterns. This will then predict the data to be fraud or not fraud.

- Identity verification - Amazon Rekognition

By using Amazon Rekognition, the customers will be asked to provide a selfie picture and identity document picture. The system uses facial biometrics powered by machine learning to do

identity verification. It also uses Amazon Rekognition Face Comparison, which helps measure the similarity of two faces to help you determine if they are the same person. Amazon Rekognition Face Index and Search enables the system to create a face collection of existing users and search for new user selfie pictures against all faces in the collection to detect duplicate or fraudulent account creation attempts.

4.0 Conclusion

4.1 Achievement

Throughout the making of this project part 1, we learn about the problem-solving techniques that can be applied in our daily life as well. These problem-solving techniques help us to determine the main problems and focus on what we are targeting to be done. By providing these services, we can get full partnership because the company trusts our services in improving their data security system. Plus, we can get to know more about AWS services that provide strong security in data information such as Amazon SageMaker. Machine learning and artificial intelligence also help in recognizing the pattern of the fraud and learning from the past data to protect the system from being attacked by hackers again. Cybersecurity systems can work together with machine learning to ensure that the data information will be more secure than before by notifying the IT administrator of that company about those unusual activities.

4.2 Contribution

Next, from this system Bank A's data information security will increase. This is due to the services and 4th IR technologies that help in protecting the data from any fraud so data stolen from unauthorised identity will be decreased. By having cloud computing, there will be no limitations in accessing data compared to IT traditional architecture. With machine learning and artificial intelligence, the security of data will be flexible where the machine learning system can react and make good decisions when repeated fraud happens again. This cloud computing also will be easy to maintain with low cost compared to traditional IT architecture where cloud computing will collaborate with machine learning and artificial intelligence where they do not need human intervention.

4.3 System Limitation

Last but not least, this system has limitations where hackers or cybercriminals can hack the machine learning algorithm and they can manipulate the experience that machine learning has been involved in. From this, machine learning cannot protect the system from being hacked because machine learning did not have the algorithm for the usual pattern that those hackers use.