

FRAUD RISK MANAGEMENT SYSTEM

*Project report submitted in partial fulfillment of the requirement for the degree
of*

BACHELOR OF TECHNOLOGY

IN

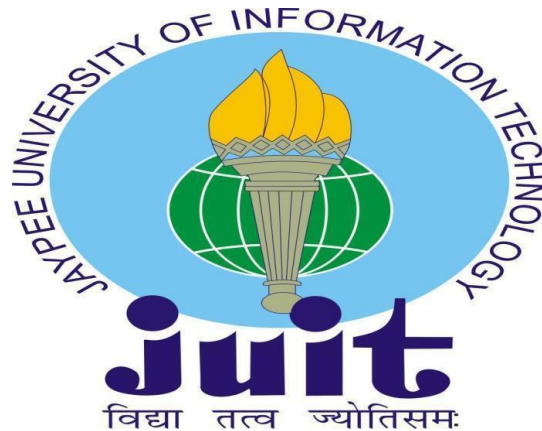
ELECTRONICS & COMMUNICATION ENGINEERING

By

Pragya Tiwari (191004)

UNDER THE GUIDANCE OF

Dr. Pardeep Garg



**Jaypee University of Information Technology Waknaghat, Solan-
173234, Himachal Pradesh**

Table of Contents

DECLARATION	3
ACKNOWLEDGEMENT	4
LIST OF FIGURES	5
ABSTRACT	6
1. INTRODUCTION	8
1.1. INTRODUCTION	8
1.2. MOTIVATION FOR THE WORK	12
1.3. OBJECTIVES	15
1.4. METHODOLOGY	16
1.5. SOFTWARE REQUIREMENTS	21
2. LITERATURE REVIEW	27
2.1. LITERATURE REVIEW	27
3. SYSTEM DEVELOPMENT	30
3.1. ALGORITHMS	30
4. CONCLUSION	35
5.1. CONCLUSION	35
5.2. FUTURE SCOPE	37
5. REFERENCES	39
PLAGIARISM REPORT	

CANDIDATE'S DECLARATION

I hereby declare that the work presented in this report entitled “**Fraud Risk Management System**” in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Electronics & Communication Engineering** submitted in the department of Electronics and communication engineering, Jaypee University of Information Technology Wagnaghat is an authentic record of my own work carried out over a period from August 2022 to May 2023 under the supervision of **Dr. Pardeep Garg**, Assistant Professor in Electronics and Communication Engineering Department.

I also authenticate that I have carried out the above-mentioned project work under the proficiency stream **Embedded System**.

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Pragya Tiwari, 191004.

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Dr. Pardeep Garg,
Assistant Professor,
Electronics & Communication Engineering
Dated: 9th May 2023

ACKNOWLEDGEMENT

Firstly, I express my heartiest thanks and gratefulness to almighty God for His divine blessing to make it possible to complete the project work successfully.

I am grateful to my supervisor Dr. Pardeep Garg, Assistant Professor, Department of ECE Jaypee University of Information Technology, Waknaghat. The deep knowledge & keen interest of my supervisor in the field of “Genomic Signal Processing, Internet of Things(IoT), Image Processing, and Applications of Machine Learning” helped me to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to Dr. Pardeep Garg, Department of ECE, for his kind help to finish my project.

I would also generously welcome each one of those individuals who have helped us in making this project a win. In this unique situation, I might want to thank the various staff individuals, both educating and non-instructing, which have developed their convenient help and facilitated my undertaking.

Finally, we must acknowledge with due respect the constant support and patients of my parents.

Pragya Tiwari, 191004

LIST OF FIGURES

Figure 1 Flowchart	22
Figure 2 Uploading the excel sheet	33
Figure 3 Training the model	34
Figure 4 Accuracy of algorithm	36
Figure 5 Trained model result	36
Figure 6 Integrating MySQL with python code	37

ABSTRACT

The number of UPI fraud transaction cases is increasing in the modern era of the digital payment system. Since a few years ago, the number of fraud cases has almost doubled[1]. A real-time fraud management system that can identify and prevent fraudulent UPI transactions is the focus of my project, Fraud Risk Management System. It utilizes API calls to operate instantly.

The MYSQL server houses the customer's transactional information. The algorithms retrieve transaction information from the MySQL server, analyze the customer's prior transaction patterns, and save the required information in the pickle library, a Python library. Real-time transaction information is compared to historical data, and a fraud score is generated based on a number of factors. The bank server checks this score, and if it exceeds a predetermined threshold, the specific UPI transaction is immediately blocked.

UPI fraud transactions have become more frequent as a result of the quick adoption of digital payment technologies, raising serious concerns among both consumers and financial institutions. My project, the Fraud Risk Management System, provides a trustworthy and practical solution for the timely detection and avoidance of fraudulent transactions in order to address this issue.

This technology is an effective tool for preventing fraudulent transactions since it uses cutting-edge algorithms to analyze transaction data in real-time and spot suspicious activity through API calls. The system's machine learning algorithms examine customers' transaction histories, which are kept in the MYSQL server and detailed information is kept in a pickle-named python library. The system then computes a fraud score using a set of criteria, and if the score exceeds a specified threshold, it immediately blocks the transaction.

The Fraud Risk Management System is adaptable to the requirements of financial institutions and may be smoothly integrated with the current banking infrastructure. The system can

be customized to meet the needs of various organizations and produce the best outcomes thanks to this capability.

Financial institutions must put effective safeguards in place to stop fraudulent activity given the growing potential of UPI fraud transactions. The Fraud Risk Management System can support the development of a cashless economy in the nation by spotting and stopping fraudulent transactions in real-time.

1. INTRODUCTION

1.1. INTRODUCTION

Internet payments have paved the way for cashless transactions in the modern world, but they have also given criminals access to a new internet platform. Nowadays, frauds involving Unified Payment Interface (UPI) payments are a prevalent problem due to the growth of these transactions. But if we exercise caution when making any UPI payments and use a secure method, we can simply defeat these con artists and protect our money from their wrongdoing[2].

It's critical to know that payment fraud is dynamic and always evolving. Fraudsters will rapidly intensify those kinds of attacks as they develop new, effective techniques. Your entire perspective on payment management may alter as a result of new payment processing systems, strategies, channels, suppliers, and integrators offering new opportunities and solutions.

I have developed a "fraud risk management system" to handle the issue of UPI fraud. This project's goal is to identify every UPI scam transaction. An individual transaction can be classified as either authentic or fraudulent based on the score that this model generates. I used a SQL dataset to analyze the person's transactional patterns and fetched the essential information in order to build the model. The fraud cases were then identified using Python, and a score was created to help identify the type of transaction.

The capacity of fraudsters to constantly alter their tactics and strategies presents one of the largest obstacles in the detection of UPI fraud. To make sure the fraud risk management system is still capable of identifying fresh forms of fraud, this calls for ongoing review. Furthermore, tight security controls like two-factor authentication, encryption, and frequent security updates can help stop fraud before it even starts.

It is feasible to utilize machine learning algorithms, which can learn from prior fraud cases and adapt to new fraud types, to increase the effectiveness of the fraud risk management system. This can aid in spotting transactional patterns and abnormalities that might point to fraud. But

even if technology is crucial for detecting and stopping fraud, human monitoring and intervention are equally crucial. This can be useful when a transaction is flagged as potentially fraudulent by the fraud risk management system but is legitimate.

The growing complexity of the payment ecosystem, with numerous players involved in the process, presents another difficulty in mitigating fraud risk. For everyone to be aware of the potential dangers and take action to mitigate them, there needs to be collaboration and coordination between various stakeholders. The goal is to maximize consideration of client journey touchpoints. The use of automated processes, cloud-based technology, and artificial intelligence analysis have all made it possible to better optimize risk controls for the situation.

People should be aware of the typical UPI scam kinds so they may take preventative measures to avoid them. Phishing is a typical form of UPI fraud, in which con artists impersonate real messages or phone calls to convince victims to divulge their UPI credentials. Another sort of fraud involves con artists requesting money over UPI while posing as someone the victim knows, like a friend or relative. Malware is another tool that fraudsters may use to access people's phones and UPI credentials.

There are several actions that folks can take to stop these frauds. It is crucial to maintain your UPI credentials secure and to never divulge them to anyone. Prior to replying to any messages or calls, always confirm their legitimacy. For UPI transactions, it is also advised to setup two-factor authentication and use secure passwords. The software on your mobile device should constantly be updated, and you should refrain from downloading apps from untrusted sites.

It is crucial for people and companies to stay up to date on the most recent advancements in payment fraud and take precautions as technology develops. This entails utilizing dependable payment methods and frequently checking transactions for any irregularities. We may guard ourselves against falling prey to UPI fraud and other types of payment fraud by remaining cautious and adopting the required steps.

In order to protect their customers' money and stop any fraud, businesses must also invest in reliable fraud management solutions. This entails putting strategies in place to identify and stop fraudulent transactions, such as real-time transaction monitoring, customer verification, and

machine learning algorithms. Businesses may defend their reputations and guarantee the security of their customers' payments by taking a proactive approach to fraud management.

In conclusion, UPI payments have transformed how we conduct business, but they have also given criminals access to new distribution channels. But with the correct knowledge and safeguards, we may safeguard ourselves and avoid becoming a victim of UPI scam. To protect the security and safety of their customers' payments, firms must adopt a proactive approach to fraud management. Together, we can fight payment fraud and develop a secure payment ecosystem that benefits all users.

It is important to note that payment fraud can have negative effects on a company's finances and reputation. Businesses that do not sufficiently protect the payments made by their consumers could face legal and regulatory repercussions in addition to losing money and customer trust.

Businesses should periodically undertake internal audits to find any potential weaknesses in their payment systems and educate their staff members on best practices for payment security in order to reduce the risks connected with payment fraud. Businesses should also think about forming alliances with payment processors and fraud management firms that provide cutting-edge security measures and fraud detection services.

Payment fraud is a widespread issue in the contemporary digital era that has an impact on people and businesses everywhere. Governments and regulatory organizations must collaborate in order to create and implement thorough frameworks and standards for payment security. This entails the creation of global standards and recommendations for payment security as well as the enforcement of harsher sanctions for payment fraud.

Finally, it is critical to understand that payment fraud is both a technological and a societal issue. Fraudsters frequently use human weaknesses like fear, greed, and ignorance to deceive their victims. As a result, it is critical to spread knowledge about payment fraud and teach consumers how to spot and prevent typical frauds.

In conclusion, payment fraud poses an increasing risk to both individuals and companies. We can establish a safe and secure payment ecosystem that is advantageous to all parties by

remaining educated, using best practices, and cooperating. Let us be proactive when it comes to payment security so that we can safeguard our personal information, our companies, and the future of the world economy. The goal of the project FRM (Fraud Risk Management System) is to safeguard financial institutions against any UPI fraud transactions.

1.2.MOTIVATION FOR THE WORK

Digital payment techniques have made transactions and payments easier. As technology advances, more and more Indians are using all of these UPI transaction sources, which is assisting India's transformation into a cashless economy [3].

Because of the growing use of smartphones and the internet, digital payments have revolutionized how money is exchanged in India. Several of the most well-liked payment options, like mobile wallets, UPI, and internet banking, enable users to complete transactions quickly and without the need for currency.

The adoption of digital payments has been significantly influenced by the government of India's financial inclusion program. The country's expansion of digital payments was made possible by the demonetization of high-value currency notes in 2016. This act of demonetization served as a catalyst. The government has further supported the cashless transaction trend by promoting digital payments through a number of programs and incentives.[4]

A strong fintech business has emerged in India as a result of the rise in digital payments. Fintech businesses have used technology to provide customers with cutting-edge digital payment options, improving the convenience of financial transactions.

However, as digital payments have grown in popularity, fraud incidents have as well. Fraudsters employ a variety of methods to trick customers into doing their dirty work for them. As a result, there is a growing need for fraud detection and prevention mechanisms.

The creation of the "Fraud Risk Management System" is a critical step in guaranteeing the security of electronic payments. By examining transaction patterns and generating a score that helps identify the type of transaction, this model can identify fraudulent transactions. Financial institutions and fintech firms can successfully prevent fraud and guarantee the security of digital payments by implementing such solutions.

The number of fraud instances is rising along with the rate of digitization. The volume of frauds recorded by financial institutions (FIs) using cards and internet banking was significantly higher than before, according to the Reserve Bank of India's annual report from 21–22 [5].

As I learned more about these problems via my work at the Shivalik bank, I began to develop my own model "Fraud Risk Management System" that will prevent any fraudster from effectively achieving his goals.

Effective fraud risk management systems that can identify and stop such operations are required to counteract the rising trend of fraudulent activity. Banks and financial institutions can create effective systems that can recognize and flag suspicious transactions in real-time with the aid of digital technologies like machine learning and artificial intelligence.

In addition to the monetary loss, payment theft can harm the institution's brand, lose consumer trust, and trigger regulatory investigation. Therefore, it is essential to have a thorough and trustworthy system in place that can give the relevant authorities accurate and timely alerts.

Lack of standardization in the data formats and protocols used by different payment systems is one of the difficulties in establishing fraud risk management systems. Effective data aggregation and analysis are so challenging. But there is a chance to create more integrated and interoperable fraud management systems now that open banking APIs and data standards have emerged.

Customer education and awareness is a crucial component of fraud risk control. People need to be educated on the many fraud kinds and the precautions they can take to protect their transactions. Institutions should also spend money on ongoing training and awareness campaigns to inform their staff and clients about fraud risk reduction.

Overall, managing fraud risk is a crucial component of current banking and financial services. It is possible to reduce the dangers and guarantee a secure payment ecosystem for everyone with the aid of technology, data analytics, and customer awareness.

There are further issues with digital payments in addition to the surge in fraud instances. There is a chance, for instance, that technical issues and system breakdowns will cause transactions to be delayed or to fail altogether. In order to protect sensitive financial and personal data from fraudsters, strong security measures are also required.

A significant emphasis has been placed on using cutting-edge technologies like blockchain and artificial intelligence (AI) to enhance the security and effectiveness of digital payments in order to address these issues. By producing an unchangeable record of each transaction, blockchain, for instance, can offer a safe and transparent platform for transactions. By examining vast amounts of data, AI can assist in identifying and stopping fraud.

Increasing public awareness and education is a crucial component in assuring the success of digital payments. This includes educating individuals about safe and secure online behaviors as well as the many forms of digital payment they can use. It's also crucial to support and help those who might be less tech-savvy and need direction while using these electronic payment options.

Overall, even though digital payments have undoubtedly revolutionized the way we conduct business, it is crucial to be attentive and take the appropriate security measures to guard against fraud and protect our personal and financial data. We can continue to make progress towards a more seamless and safe digital payment ecosystem by utilizing cutting-edge technologies and encouraging awareness and education.

1.3.OBJECTIVES

Finding all the fraudulent transactions is the project's main objective. Preventing fraudsters from carrying out fraudulent UPI transactions and stealing the people's hard-earned money is the main goal of the entire endeavor. UPI transactions are becoming the most popular form of payment among consumers all throughout India in the current digital age. However, as UPI transactions have become more widely used, the prevalence of fraudulent operations has also increased. Customers suffer enormous financial harm as a result of fraudsters' constant invention of new schemes to conduct fraudulent transactions, undermining their faith in electronic payment systems.

The project's goal is to create a trustworthy fraud detection and prevention system that can identify and stop all fraudulent transactions in real-time in order to address this problem. The main goal of the system is to resist attempts by fraudsters to carry out unauthorized transactions and steal peoples' hard-earned money. The technology uses complex algorithms to analyze transaction data in real-time, allowing it to spot any questionable activity and only perform authorized transactions.

The solution can increase customers' trust in electronic payment systems in addition to improving their financial security. The technology can offer a safe environment for users to make digital transactions without worrying about being scammed by preventing fraudulent behaviors. Customers are not the only beneficiaries of the scheme; financial institutions might also gain from decreased fraud-related losses. The method can spare financial institutions millions of dollars that would otherwise be wasted to fraudsters by stopping fraudulent transactions.

The main objective of the initiative is to make it easier and safer for people to conduct digital transactions. The technology can assist in moving the nation towards a cashless economy, making financial transactions more convenient and effective for everyone, by avoiding fraudulent activity. The system's effects may be widespread, fostering a climate of trust and confidence in digital payments and boosting the nation's overall economic development.

1.4.METHODOLOGY

Two different strategies are used for the entire project. The first method only detects the fraudulent UPI transaction for the one person it was intended for. The second method will be effective for all consumers because it will identify fraudulent UPI transactions based on each customer's usual pattern of transactions.

Approach 1: Generating the fraud score for a single person using Machine Learning

- The project begins with fetching the demo transaction details of a person.
- The required excel sheet of one person's transaction history (the data was manipulated was not the real data of a person) was received from the Shivalik bank.
- Jupyter notebook and google colab were used to write the python code for the Machine Learning Model.
- The excel sheet was fetched from the system onto a jupyter notebook and it was converted into a data frame.
- Only debit transaction details were kept in the excel sheet, and credit transaction details, as well as balance inquiry, were removed because only debit transactions could have frauds.
- The Payer account number was kept as a reference to visit any transaction details as the payer account number is unique for every customer.
- One extra column was introduced named target variable which contained the two outputs, fraud and not fraud.
- Pre-processing of data was done. Using label encoding, all the labels were converted into numeric form so that the machine learning algorithms can be applied to them.

- Sklearn library was imported along with the necessary classes such as train-test split, Logistic Regression, accuracy score.
- Finally, we trained and evaluated our model.
- The attributes were divided into an 80:20 ratio. 80 percent of the data was used to train the model and the rest 20 percent of data was used to test the model.
- After knowing the target variable and necessary attributes, Logistic regression was applied to train and test the model.
- The final accuracy of the model was 93.7 %. Finally, the data frame containing all the numerical values was stored in an excel sheet for future reference.

Approach 2: Generating the fraud score for all persons without using Machine Learning. The model that works for all the customers is divided into two codes.

Code 1:

- Used an excel sheet containing around 9 lakhs of transaction data of every customer for the past month. (Generated account numbers using random function so the actual data of the customers won't be leaked)
- This excel sheet was uploaded to the MySQL database.
- MySQL database was integrated with google colab using the python module called pyodbc.

- Using the driver's and MySQL software's information, the entire transaction details were fetched and was converted into a data frame. The code to fetch the entire information from the database was: `query=select* from dbo.may order by s_no.`
- Reason to integrate code with MySQL database: the bank has its own database named UAT, which gets updated daily on the basis of latest transaction of the person. From that we can analyze the latest habit of the person. This will help to create a fully fledged working model which will directly fetch information from the database of the bank and there will be no need to upload the excel sheet again and again.
- Blocks of time (4 hours each), amount (difference of Rs. 5000 in each block) were created. Each block contained the frequency of transactions belonging to them. Furthermore, a dictionary was created for every customer which stores the frequency of a particular payer address.
- Fraud score was calculated based on the frequency of transaction in every block. The blocks which contained more transactions were given less score while the blocks which contained less transactions were given more score.
- $\text{Score for each block} = \text{max score (10)} - (\text{max score (10)} / \text{max frequency of transaction}) \times \text{current frequency of transaction.}$
- Now the data frame which contained the score for every block for every customer was assigned the account number as the index and this data frame was saved in a pickle format which is provided by python. This pickle file was uploaded on AWS (a cloud service platform powered by amazon). This pickle file will be used to generate the score for real time transactions. The sole reason to store the pickle file and use it in the next code to generate score was to make the second code run as fast as possible.

Code 2:

- Pickle file was fetched from the cloud and was converted into data frame. Whenever a real time transaction occurs, the API call gets triggered and the necessary information about the transaction, such as account number, amount, date, and time is sent to this code. Now this information will get compared with the slots in which they lie and a score variable gets added up for the value stored inside particular blocks. Finally, this score value is sent to the bank server via API and if the score value is greater than the threshold value, the transaction gets blocked.

Entire workflow of the model-

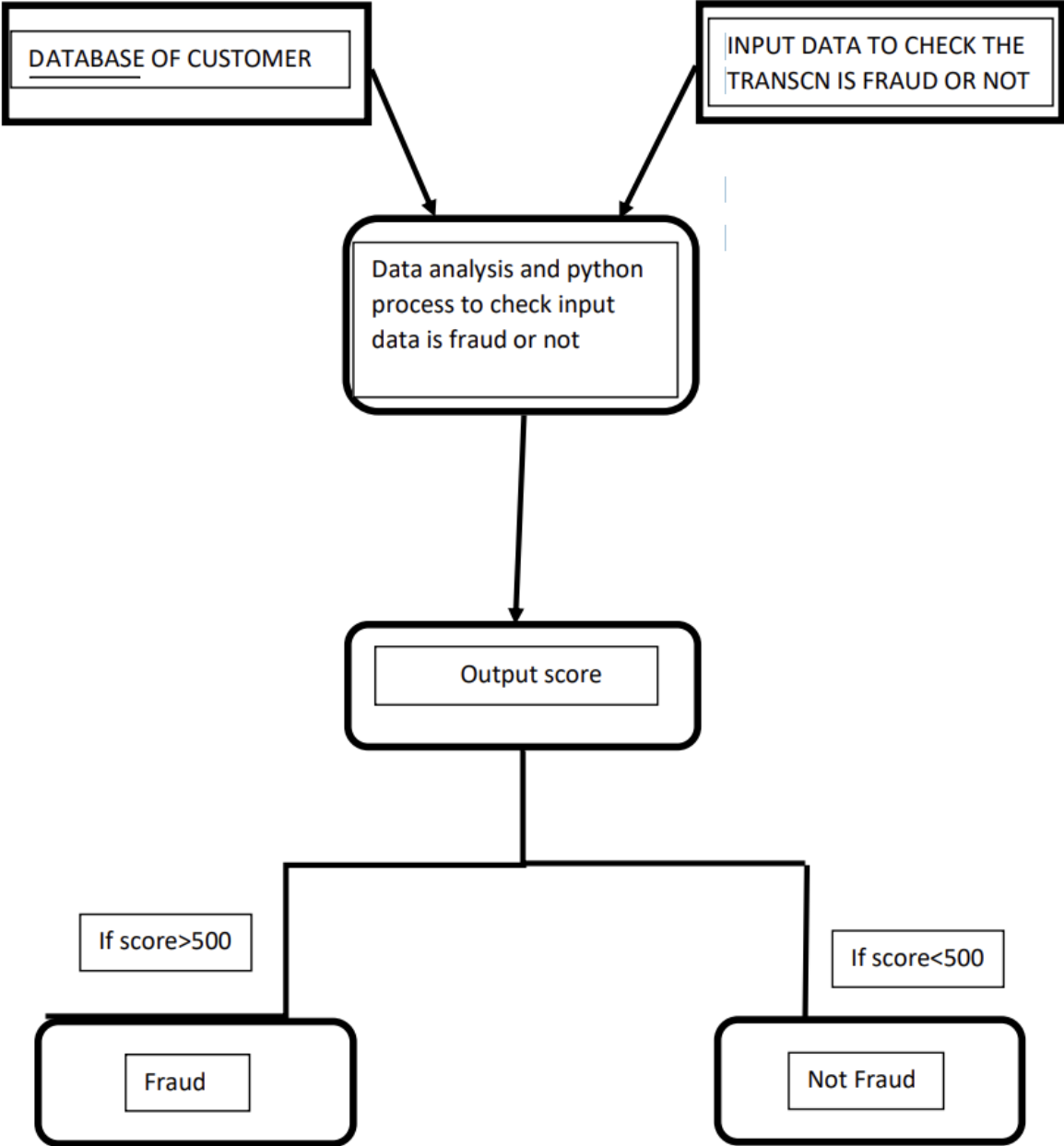


Figure 1 Flowchart

1.5.SOFTWARE REQUIREMENTS

1.5.1. PYTHON

Python is a computer language which is used to create software. It contains several built-in libraries and is simple to learn. Python is an object-oriented, high-level programming language with dynamic semantics that is interpreted.

Python is frequently used for Rapid Application Development and as a scripting or glue language to join together already existing components. Python is a preferred choice for many developers due to its built-in data structures and dynamic binding. The grammar of the language is straightforward and simple to learn, which lowers the cost of program maintenance. Python also provides modules and packages that promote modularity and code reuse. The Python interpreter and extensive standard library are both available in source or binary form for all major platforms at no cost and can be distributed freely[6].

The rich library support offered by Python is one of its important characteristics. Numerous built-in libraries in Python offer prewritten code to carry out a variety of functions. Because programmers may use these libraries to quickly implement functionality in their applications without having to create code from scratch, they are intended to simplify and accelerate the development process.

1.5.2. PANDAS

Python users can work with data frames using the PANDAS library. A free and open-source library called Pandas is made for quickly and easily dealing with relational or labelled data. It is constructed on top of the NumPy library, which makes it quick and high-performing[7]. Its data structures and operations are targeted towards handling numerical data and time series.

Wes McKinney created the library in 2008 while he was employed by AQR Capital Management. He was later joined by Chang She in 2012, who became the project's second major contributor. Pandas have since been issued in numerous versions, the most recent of which being version 1.5.3 as of January 18, 2023[7].

Pandas users can benefit from a variety of benefits. It can import data from various file objects and manipulate and analyze data quickly and effectively. Additionally, it enables users to represent missing data as NaN in both floating-point and non-floating-point data, making it simple to handle missing data. Pandas also has the benefit of being size mutable, which allows for the addition and deletion of columns in Data Frames and other higher dimensional structures. Additionally, it offers time-series capabilities, customizable reshaping and pivoting of data sets, and merging and joining of data sets. Additionally, it has a strong group-by functionality for splitting, applying, and combining data sets[7].

1.5.3. NUMPY

Numpy is a library for dealing with numerical equations in Python. Python has a robust and popular package for scientific computing called NumPy. It gives a multidimensional array object with great performance and a wide number of utilities for working with these arrays, together with an array-processing package. Numpy is a great option for executing intricate mathematical computations and statistical analysis due to its effective data storing and manipulation capabilities[8].

Numpy is a multi-dimensional container that can be used to store and analyse general data in addition to being utilised for scientific applications. A tuple of positive integers serves as the index into the array in Numpy, which is a table of elements that are typically numerical and of the same type. The shape of the array is a tuple of numbers that represents the size of the array along each dimension, and the rank of the array is the number of dimensions[8].

A variety of techniques for manipulating arrays are available through the Numpy ndarray class, including arithmetic operations, slicing, indexing, and reshaping. Square brackets are used to access Numpy array elements, and nested Python Lists can be used to initialise Numpy arrays. Additionally, vectorized operations are supported by Numpy arrays, allowing for the manipulation of enormous arrays with a single operation and producing faster and more effective code[8].

Overall, Numpy is a crucial tool for scientific computing and data analysis in Python because to its strong array-processing capabilities, user-friendliness, and wide variety of functionality[8].

1.5.4. SKLEARN

Sklearn is a Python package that comes with the language that handles machine learning techniques. A Python library called Scikit-learn, commonly referred to as Sklearn, offers a variety of tools for machine learning, data mining, and data analysis. Since it was initially developed as an extension of the SciPy library, its name is a mix of "SciPy Toolkit" and "learn," and it was first developed by French research scientist David Cournapeau as part of the Google Summer of Code project. Other programmers gradually rewrote the basic software, with the French Institute for Research in Computer Science and Automation serving as the project's lead in 2010[9].

Since its initial public release on February 1, 2010, Scikit-learn has become one of the most well-liked machine learning libraries for Python. In fact, it ranks among the most popular open-source projects on GitHub[9].

Sklearn is mostly developed in Python and makes extensive use of NumPy for array and linear algebra computations. Some of the fundamental algorithms in the library were also developed in Python to improve its efficiency. SciPy, Pandas, Matplotlib, seaborn, and plotly, to name a few, are just a few of the many Python programs that work flawlessly with Scikit-learn[9].

Algorithms for decision-making, classification, regression, grouping, and predictive analysis are some of scikit-learn's key ideas and features. It supports a number of algorithms, such as simple linear regression and neural networks for pattern recognition and predictive analysis. It is compatible with the matplotlib, pandas, and numPy libraries. Without explicit programming, a predictive model can be created or trained on input data by computers using machine learning. Numerous industries, including marketing, banking, and healthcare, use Scikit-learn extensively[9].

1.5.5. DATETIME

Datetime is a Python module that is used to work with date and time-based data. A built-in module in Python called datetime offers classes for working with dates, timings, and intervals of time. When handling and modifying time and date data in Python programs, these classes are helpful. You don't need to install the module individually because it is built into Python by default[10].

The fact that the datetime module offers classes for representing dates and times as simply formattable and manipulable objects is one of the key benefits of using it. The date class, time class, datetime class, timedelta class, tzinfo class, and timezone class are some of these classes[10].

The date class provides a hypothetical date with properties for the year, month, and day. It presumes that the Gregorian calendar is constantly in use. The time class, which has properties for the hour, minute, second, microsecond, and tzinfo, depicts an idealised time. Assuming that every day has exactly $24*60*60$ seconds, it is independent of any specific day. The year, month, day, hour, minute, second, and microsecond properties are included in the datetime class, which combines the date and time classes[10].

The timedelta class represents a duration or difference between two dates, times, or datetime instances. It is useful for calculating the difference between two dates or times, and can be used for addition or subtraction operations. The tzinfo class provides time zone information objects, which can be used to handle time zones and daylight saving time. Finally, the timezone class is a new class in version 3.2, which implements the tzinfo abstract base class as a fixed offset from UTC. Overall, the datetime module is an essential part of Python for working with dates and times, and it provides a wide range of functionality to handle date and time data in Python programs[10].

1.5.6. PYODBC

A python library that is used to integrate SQL with our python code. Python is a high-level programming language that has gained popularity due to its readability, ease of use, and versatility. It can be used for a wide range of applications, including web development, data analysis, machine learning, and artificial intelligence. Guido van Rossum created Python in 1991, and since then it has become one of the most widely used programming languages[11].

Pyodbc is a powerful and efficient open source Python module that provides an interface for connecting to ODBC databases. It adheres to the DB API 2.0 specification, which means that it is compatible with a wide range of databases. With pyodbc, Python developers can easily connect their applications to a database using an ODBC driver. The ODBC driver manager and ODBC driver work together to establish a connection to the database server, which can be located either locally or remotely. The ODBC driver manager is platform-specific, while the ODBC driver is specific to the database that you are connecting to. By using pyodbc, Python developers can leverage the power of ODBC to access and manipulate data in a wide range of databases, including Microsoft SQL Server, Oracle, and MySQL[11].

2. LITERATURE REVIEW

2.1.LITERATURE REVIEW

As per the data that I have fetched out Fraud Risk Management is a method through which frauds are systematically reduced in an organization. If the fraud is detected at an early stage only then it can be reduced. After the detection of fraud, we need to take measures to stop it. This entire approach is known as Fraud management.

As per the survey, many organizations could not develop due to increasing fraud, this makes it crucial to focus on the Fraud risk management part of the organization. For an organization Fraud Detection is important. If there is a protocol for detecting such fraud, then an organization can effectively tackle its way during a crisis. Hence it is important to have such a framework related to fraud risk management[12].

Precautions that one should take to prevent UPI Frauds:

- While dealing with any payment request double-check all details.
- Don't enter your pin while receiving funds, there is no need for it.
- Take the warning message that you get on your UPI app seriously.
- Don't download any suspicious app on your device.
- Don't share your PIN or OTP to some random calls.
- Don't open any suspicious link or message.

It takes a lot of time to discover UPI fraud. UPI fraud transactions should be handled with extreme caution[13].

One of the study papers I read promotes the idea that customers will receive a warning message and the choice of whether or not to report fraud following each transaction. According to their proposal, the consumer will have two choices for a questionable transaction: "OK" for a genuine transaction and "Fraud" for a fraudulent transaction[14].

According to a joint study by DSCI and PayPal, the tools made for fraud prevention should have the features that are mentioned below:

- Machine learning: Using real-time capabilities offered by both monitored and unmonitored anomaly detection approaches, uncover fraud patterns using UPI transaction data or user behavior patterns. Complex computations can be carried out by machine learning algorithms more quickly than by human validation[15].

Finding patterns that point to fraudulent behavior is one of the biggest hurdles in fraud detection. This problem has grown more difficult with the emergence of digital payment systems like UPI. Modern fraud risk management systems, on the other hand, have access to real-time features that let them closely monitor user behavior and transaction data.

To identify probable fraud trends, these systems employ both monitored and unmonitored anomaly detection techniques. Unmonitored approaches rely on machine learning algorithms to find deviations from predicted behavior, whereas monitored methods use predetermined rules and patterns.

- Automated workflow: Speed up workflow by automating checks for payment fraud, processing order details, blocking suspicious devices, canceling fraudulent orders, and more[8].

Implementing automation solutions can greatly improve your processes and save you significant time. You can speed up your workflow by automating a number of processes, such as processing order details automatically, blocking any suspicious devices, cancelling fraudulent orders, and checking for payment fraud. This not only lessens the workload on your team but also lowers the possibility of mistakes and fraud. By utilizing automation, you can streamline your business operations and concentrate your efforts on more crucial duties like expanding your company and providing remarkable client experiences.

- Insights dashboard: Highlights suspicious activity and shows relevant anti-fraud data in a single interface without switching between multiple screens. This allows you to effectively execute and simplify the fraud screening process[8].

You can efficiently detect and reduce fraudulent behavior in real-time by utilizing cutting-edge fraud detection technology. Your fraud screening procedure can be streamlined and made more effective with the use of a single interface that flags questionable activity and offers pertinent anti-fraud data. With this method, you may obtain pertinent information without wasting time switching between various screens. Instead, you can quickly obtain all the relevant information in one location and act quickly to stop fraudulent transactions from happening. This not only improves your fraud protection abilities but also lowers the chance of expensive chargebacks and business reputational harm. Overall, investing in a fraud detection solution that offers a simplified and streamlined interface can help you protect your business and enhance your customers' trust in your brand.

- Chargeback Guarantee: In case of fraud, such services fully cover approved orders, making it a risk-free investment and keeping your money safe[15].
- Device Fingerprint: Technologies that record information about the devices you use to conduct online transactions. Several properties such as browser, operating system, location, and language are analyzed to see if the device being used is associated with fraud and can be blocked[15].
- Customization: You can manually customize fraud checks and data points to suit your business type and personal preferences[15].

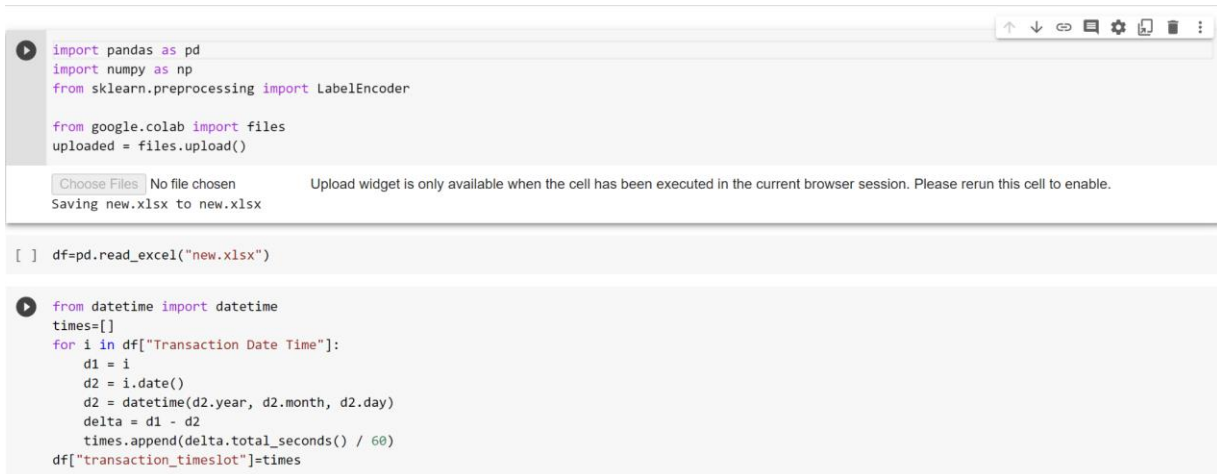
All UPI transactions, according to the data I've studied, are NPCI rule-based, but the model I'm presenting here is habit-based and operates in accordance with the person's most recent transaction habits. The majority of the data I've fetched out was related to fraud prevention, based on how to prevent the UPI fraud transaction, but my model is related to detection that how to identify that fraud transaction and doesn't permit that transaction to occur.

3. SYSTEM DEVELOPMENT

3.1. ALGORITHMS

Approach 1:

Snapshot of uploading the excel sheet in the python code is attached below.



```
import pandas as pd
import numpy as np
from sklearn.preprocessing import LabelEncoder

from google.colab import files
uploaded = files.upload()

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving new.xlsx to new.xlsx

[ ] df=pd.read_excel("new.xlsx")

from datetime import datetime
times=[]
for i in df["Transaction Date Time"]:
    d1 = i
    d2 = i.date()
    d2 = datetime(d2.year, d2.month, d2.day)
    delta = d1 - d2
    times.append(delta.total_seconds() / 60)
df["transaction_timeslot"]=times
```

Figure 2 Uploading the excel sheet

After importing the necessary libraries, the spreadsheet file was uploaded to Google Colab. Each transaction's time was converted to seconds so that Python's datetime module could be used on it.

It's crucial to import the essential libraries into your Python environment before working with the data in an Excel file. You may quickly submit the Excel file to Google Colab, a cloud-based platform that offers a variety of tools and resources for data analysis and machine learning, after the libraries have been imported.

It's crucial to convert the time of each transaction into seconds before doing any datetime operations on the transaction data. This enables you to take advantage of Python's robust datetime module, which offers a variety of functions and techniques for working with date and

time data. You may quickly edit and analyze the data to acquire useful insights into client behavior and trends by converting the transaction times to seconds.

Overall, a variety of potent data analysis tools and methods are available to you by importing the required libraries and uploading the Excel file to Google Colab. You may take full advantage of Python's datetime module to acquire useful insights and improve your business operations by converting transaction timings to seconds.

Snapshot of training the model with our dataset

```
[ ] temp = []
    for i in df["target_var"]:
        if i == "not fraud":
            temp.append(0)
        else:
            temp.append(1)
    df["target"] = temp

[ ] predictors = ["transaction_timeslot", "Amount"]
    X = df[predictors]
    Y = df['target']

[ ] from sklearn.model_selection import train_test_split
    from sklearn.linear_model import LogisticRegression

[ ] x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2,random_state=5)

[ ] model = LogisticRegression(solver='liblinear',random_state=10)
    model.fit(x_train,y_train)
    y_pred = model.predict(x_test)
```

Figure 3 Training the model

The Sklearn library was loaded to use logistic regression. The data that were available were split in an 80:20 ratio. The remaining 20% of the data was utilized to test the model once it had been trained with the remaining 80% of the data. The model was trained and tested using the logistic regression algorithm.

It is crucial to use a methodical approach in order to create a reliable predictive model. In this instance, a classification model was created using the logistic regression algorithm, a well-liked and often used machine learning technique, which was imported from the Scikit-learn module. The supplied data was divided into training and testing datasets using an 80:20 ratio after the

algorithm was imported. This indicates that although 20% of the data was set aside for testing the model's performance, the remaining 80% was used for training.

The training step involved applying the logistic regression algorithm to the training dataset and tuning the model's parameters to reduce error rates. This procedure was repeated until the model was able to accurately represent the relationships and patterns in the training data.

The remaining 20% of the data was utilized to test the model's precision and generalizability once it had been trained. The model was fed the testing data, and its effectiveness was assessed based on how well it foresaw the results of the test data.

Overall, the 80:20 split ratio and the logistic regression algorithm allowed for the building of a trustworthy and accurate predictive model. This strategy ensures that the model can effectively predict outcomes on fresh, unforeseen data and is a generally acknowledged and employed practice in the field of machine learning.

Snapshot of finding the accuracy of the algorithm.

```
[ ] import numpy as np
    from sklearn.metrics import accuracy_score
    print('accuracy:', accuracy_score(y_test, y_pred))
```

accuracy: 0.9375

```
[ ] import datetime
    transaction_datetime = datetime.datetime.now()
    d1 = transaction_datetime
    d2 = transaction_datetime.date()
    d2 = datetime.datetime(d2.year, d2.month, d2.day)
    delta = d1 - d2
    transaction_datetime = delta.total_seconds() / 60
    amount = 5000.0

    tdf = pd.DataFrame(columns=["transaction_timeslot", "Amount"])
    tdf.loc['t1'] = [transaction_datetime, amount]
```

```
▶ model.predict(tdf)
```

```
↳ array([0])
```

Figure 4 Accuracy of the algorithm

The accuracy of the model was 93.7 percent. It was calculated using the `accuracy_score` module of `sklearn` library

Snapshot of the result we got after giving the input.

```
[ ] from math import floor
    score = floor(model.predict_proba(tdf)[0][1]*1000)
    print(score)
```

130

Figure 5 Trained model result

Finally, the model was used to calculate and display the fraud score for any given transaction.

4. CONCLUSION

4.1. CONCLUSION

We can infer from all the data in this report that we will receive a numerical output to determine whether a transaction is fraudulent or not. According to this model, if the score is higher than 500, it can be classified as fraudulent, and if it is lower than 500, it can be classified as not fraudulent. The bank can set the threshold to block the transaction based on its own preferences. This model is based on a person's transaction habits, such as how they typically transact in terms of amount, timing, and payer address. Based on these attributes, we will generate an output that will enable us to determine whether a transaction is legitimate or fraudulent.

It is evident from the data in this report that the fraud detection model generates a numerical output to identify whether or not a transaction is fraudulent. The methodology categorizes transactions into two categories: legitimate and fraudulent. Transactions with a score of 500 or higher are considered fraudulent. It's important to remember that banks can modify this threshold in accordance with their own preferences and levels of risk tolerance.

The different parameters relating to the transaction habits of the person conducting the transaction are taken into consideration by this fraud detection model. These consist of the number and time of transactions as well as the address of the payment. The model creates an output that aids in determining the legitimacy of a transaction by examining these properties and their relationships.

The ability to identify fraudulent behavior in real-time, which aids banks and financial institutions in preventing losses and upholding their reputation, is one of the main advantages of utilizing such a fraud detection methodology. The model is a flexible and useful instrument in the fight against financial fraud because it can also be adjusted and tailored to meet the unique demands of a firm.

This fraud detection model also considers the account holder's past data, such as their transaction history and account balance, in addition to the aforementioned parameters. The model can detect aberrant patterns of behavior and flag possibly fraudulent transactions for examination by analyzing this data combined with real-time transaction data.

This fraud detection model is also quite effective and capable of analyzing huge amounts of data in real-time, giving a quick and precise evaluation of transaction risk. This is essential in the fast-paced digital world of today, when fraudulent transactions can happen instantly and without prior notice.

Banks and other financial institutions must continually update and enhance their fraud detection programs to achieve optimal efficacy. This can be accomplished by including fresh and pertinent data sources, such as IP address tracking and social media activity, and by continuously enhancing the algorithms of the model to increase accuracy and reduce false positives.

In conclusion, every bank or financial institution seeking to safeguard itself and its clients from financial fraud must adopt a sophisticated fraud detection methodology. This model can successfully spot and stop fraudulent conduct by fusing historical data, real-time transaction data, and sophisticated algorithms, protecting the institution's finances and reputation.

4.2.FUTURE SCOPE

In the not-too-distant future, every bank would have access to software for fraud risk management. Unsupervised machine learning would be used to create additional enhancements that would considerably improve the outcome and increase its efficiency and dependability. We would improve the software's UI/UX to make it more user-friendly.

Totally automatic software would be available. Every day, it would gather the data from the updated server and train itself automatically. The generated dataset would be stored in the MySQL database itself as opposed to being created as a pickle file and hosted on the cloud, rendering the sensitive data vulnerable to cyber attacks. Unsupervised machine learning will be utilised to adapt the model to the customer's new transactional behaviours, and the model will continue to improve on its own. As a result, Fraud Risk Management System would become the most dependable and well-liked software to address UPI and other types of transaction-related frauds.

We may add one more characteristic to this model so that we can determine if a transaction is fraudulent or legitimate based on its location. For example, if a transaction is made at 4:00 p.m. from Solan, it can be determined that a transaction made at 4:10 p.m. from Delhi is fraudulent.

With the emergence of cutting-edge software solutions and the further advancement of technology, the future of fraud risk management appears promising. Banks will have access to sophisticated fraud detection software in the upcoming years that makes use of unsupervised machine learning algorithms to greatly improve the effectiveness and dependability of fraud protection. We will work on the software's user interface and user experience (UI/UX), making it more user-friendly and intuitive, to provide a seamless user experience.

With the introduction of autonomous data fetching, the software will train itself without human input by retrieving data from regularly updated servers. The generated dataset will be safely kept in a MySQL database, lowering the danger of cyber attacks, instead of being stored in a cloud-based pickle file. The Fraud Risk Management System will become an increasingly

dependable and well-liked program for managing UPI and other sorts of transaction-related frauds as a result of the unsupervised machine learning model's ability to adapt to the customer's transaction behaviors and continue to improve on its own.

Furthermore, we can add location-based attributes to the system to further strengthen the fraud detection model. When a transaction is made from two separate places quickly after one another, such as from Mumbai at 5:50 pm and Vadodara at 6:00 pm, the model may flag it as potentially fraudulent. The Fraud Risk Management System will be significantly more successful at preventing financial fraud by being able to identify suspicious activities based on geography.

5. REFERENCES

- [1] Pushpita Dey: Online Banking Frauds Doubled Post-Covid, Hyderabad Records Highest Jump, 14 Dec 2021.
- [2] Livemint.com: UPI payment fraud: How to ensure safety of your money? 5 safety measures for you, 11 June 2022
- [3] DSCI.in: Fraud & Risk Management in Digital Payments, 26 August 2020
- [4] [Digital-Journeys: India embraces mobile money \(imf.org\)](#)
- [5] PWC.in: Combating Fraud in the era of digital payments, May 2022
- [6] [What is Python? Executive Summary | Python.org](#)
- [7] [Introduction to Pandas in Python - GeeksforGeeks](#)
- [8] [Python Numpy - GeeksforGeeks](#)
- [9] [What is Sklearn in Python - Javatpoint](#)
- [10] [Python datetime module - GeeksforGeeks](#)
- [11] [Tutorial: Connecting to ODBC Data Sources With Python and pyodbc - DZone](#)
- [12] enterslice: Fraud Risk Management- An Overview 2022.
- [13] Dhruv:How To Prevent UPI Fraud And Improve Financial Security. May 14 2022.
- [14] Neha Priyal, Jawed Ahmed, M. Afshar Alam: Digital Payments: A scheme for Fraud Data Collection and Use in Indian Banking Sector.

[15] A DSCI-PayPal Joint Study: Fraud & Risk Management in Digital Payments Fraud & Risk Management in Digital Payments July 2020

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
PLAGIARISM VERIFICATION REPORT

Date:

Type of Document (Tick): PhD Thesis M.Tech Dissertation/ Report B.Tech Project Report Paper

Name: _____ Department: _____ Enrolment No _____

Contact No. _____ E-mail. _____

Name of the Supervisor: _____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor)

Signature of HOD

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	<ul style="list-style-type: none"> • All Preliminary Pages • Bibliography/Images/Quotes • 14 Words String 		Word Counts	
Report Generated on			Character Counts	
		Submission ID	Total Pages Scanned	
			File Size	

Checked by
Name & Signature

Librarian

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com