

**MAHARASHTRA EDUCATION SOCIETY’S**

**Abasaheb Garware College**

**(Autonomous)**

**NAAC -Re-accredited ‘A’ Grade**

**Best Collge Award By Savitribai Phule Pune University**

**Karve Road, Pune, 411004**

Academic Year 2023-24

**“CREDIT CARD FRAUD DETECTION”**

**Project Develooped By :-**

**Khandekar Vaishnavi Anil**

**Shinde Ekta Kakasaheb**

**INDEX TABLE**

**INDEX**

|  |  |  |
| --- | --- | --- |
| Sr No. | Topic | Page No |
|  | **Company Certificate** | 1 |
|  | **College Certificate** | 2 |
|  | **Acknowledgement** | 3 |
| 1 | **Introduction**   * 1. Company Profile   2. Existing System   3. Proposed System | 4 |
| 2 | **System Requirements Specification**  2.1 Hardware requirements  2.2 Software requirements | 8 |
| 3 | **Algorithm** | 9 |
| 4 | **Feasibility Study**  4.1 Technical Feasibility  4.2 Economic Feasibility  4.3 Operational Feasibility | 10 |
| 5 | **UML Diagrams** | 12 |
| 6 | **Limitation/Drawbacks** | 12 |
| 7 | **Future Enhancement** | 13 |
| 8 | **Conclusion** | 14 |
| 9 | **Bibliography** | 15 |

**Company Certificate**

**College Certificate**

**ACKNOWLEDGEMENT**

I am extremely thankful to our project guide Mrs. Bharti Karapatti. for her valuable suggestions while the developing project. I am thankful to our head of Department Mrs. Chitra Nagarkar. for her constant motivation and encouragements.

There are so many people who contributed either directly or indirectly to complete this project. I shall mention few of them, who personally or professionally encouraged and assisted us in the entire duration of project a very pleasant endeavor.

It was a fantastic and knowledgeable experience for both of us to work together on project topic given to us. We learnt that ‘A team Spirit nature makes any difficult task easier and joyful’.

We would also like to thank all the Teaching and Non-Teaching staff members of Computer Science Department who have helped us in this project without which this project was an absolute dream for us.

**Introduction**

Brief overview of credit card fraud and its impact.

Purpose of the project: to develop a fraud detection system using machine learning algorithms.

Explanation of the algorithms used: Isolation Forest, Local Outlier Factor (LOF), and Support Vector Machine (SVM).

Importance of fraud detection in financial transactions.

The **Credit Card Fraud Detection Project** addresses this pressing concern by employing advanced machine learning algorithms to identify suspicious activities and fraudulent patterns in credit card transactions.

By leveraging large datasets containing historical transaction records, coupled with sophisticated anomaly detection techniques, the project aims to enhance fraud detection capabilities and minimize the impact of fraudulent activities on both financial

Through this project, we aspire to contribute to the ongoing efforts to enhance security in the realm of electronic payments, ultimately fostering trust and confidence in the global financial system.

**1.1Company Profile**

**1.2Existing System**

Credit card fraud detection systems typically employ a combination of rule-based techniques, anomaly detection, machine learning algorithms, and sometimes even artificial intelligence to detect fraudulent activities.

These systems analyze various factors such as transaction history, spending patterns, geographical locations, and behavioral anomalies to flag suspicious transactions for further investigation. Many financial institutions also use real-time monitoring to quickly identify and respond to fraudulent activities as they occur.

**1.3Proposed System**

A proposed system for credit card fraud detection could integrate advanced machine learning models, such as deep learning algorithms, to improve accuracy and efficiency. This system could also incorporate real-time transaction monitoring, anomaly detection, and behavioral analysis to swiftly identify and respond to potential fraud.

Additionally, leveraging big data analytics and incorporating data from multiple sources, including transaction histories, device fingerprints, and user behavior patterns, can enhance the system's ability to detect fraudulent activities accurately while minimizing false positives. Implementing a robust fraud detection system also involves continuous monitoring, updating, and refining algorithms to stay ahead of evolving fraud techniques.

**System Requirement Specification**

**2.1 Hardware Requirement :**

* + - * Lenovo IdeaPad gaming 82K1
      * **Processor:** intel core i5
      * **RAM:** 8 GB
      * **ROM:** SSD 512 GB

**2.1 Software Requirement :**

* **Operating System:** Microsoft Windows 11 Home
* **Development Environment:** python on Visual studio code
* **Data Analysis and Visualization:**
* Pandas for data manipulation and analysis.
* Matplotlib and/or Seaborn for data visualization.
* NumPy for numerical computations

**ALGORITHMS**

* **Isolation Forest Algorithm :**

One of the newest techniques to detect anomalies is called Isolation Forests. The algorithm is based on the fact that anomalies are data points that are few and different. As a result of these properties, anomalies are susceptible to a mechanism called isolation.

This method is highly useful and is fundamentally different from all existing methods. It introduces the use of isolation as a more effective and efficient means to detect anomalies than the commonly used basic distance and density measures. Moreover, this method is an algorithm with a low linear time complexity and a small memory requirement. It builds a good performing model with a small number of trees using small sub-samples of fixed size, regardless of the size of a data set.

Typical machine learning methods tend to work better when the patterns they try to learn are balanced, meaning the same amount of good and bad behaviors are present in the dataset.

* **How Isolation Forests Work :**

The Isolation Forest algorithm isolates observations by randomly selecting a feature and then randomly selecting a split value between the maximum and minimum values of the selected feature. The logic argument goes: isolating anomaly observations is easier because only a few conditions are needed to separate those cases from the normal observations. On the other hand, isolating normal observations require more conditions. Therefore, an anomaly score can be calculated as the number of conditions required to separate a given observation.

The way that the algorithm constructs the separation is by first creating isolation trees, or random decision trees. Then, the score is calculated as the path length to isolate the observation.

* **Local Outlier Factor(LOF) Algorithm :**

The LOF algorithm is an unsupervised outlier detection method which computes the local density deviation of a given data point with respect to its neighbors. It considers as outlier samples that have a substantially lower density than their neighbors.

The number of neighbors considered, (parameter neighbors) is typically chosen

**1)** greater than the minimum number of objects a cluster has to contain, so that other objects can be local outliers relative to this cluster, and

**2)** smaller than the maximum number of close by objects that can potentially be local outliers. In practice, such information are generally not available, and taking neighbors=20 appears to work well in general.

**FEASIBILITY STUDY**

**4.1 Technical Feasibility**

Technical feasibility for credit card fraud detection involves assessing whether the proposed system can efficiently analyze transaction data, employ machine learning algorithms to detect fraudulent patterns in real-time, integrate seamlessly with payment systems, scale to handle increasing volumes, and comply with security and regulatory standards.

**4.2 Economic Feasibility**

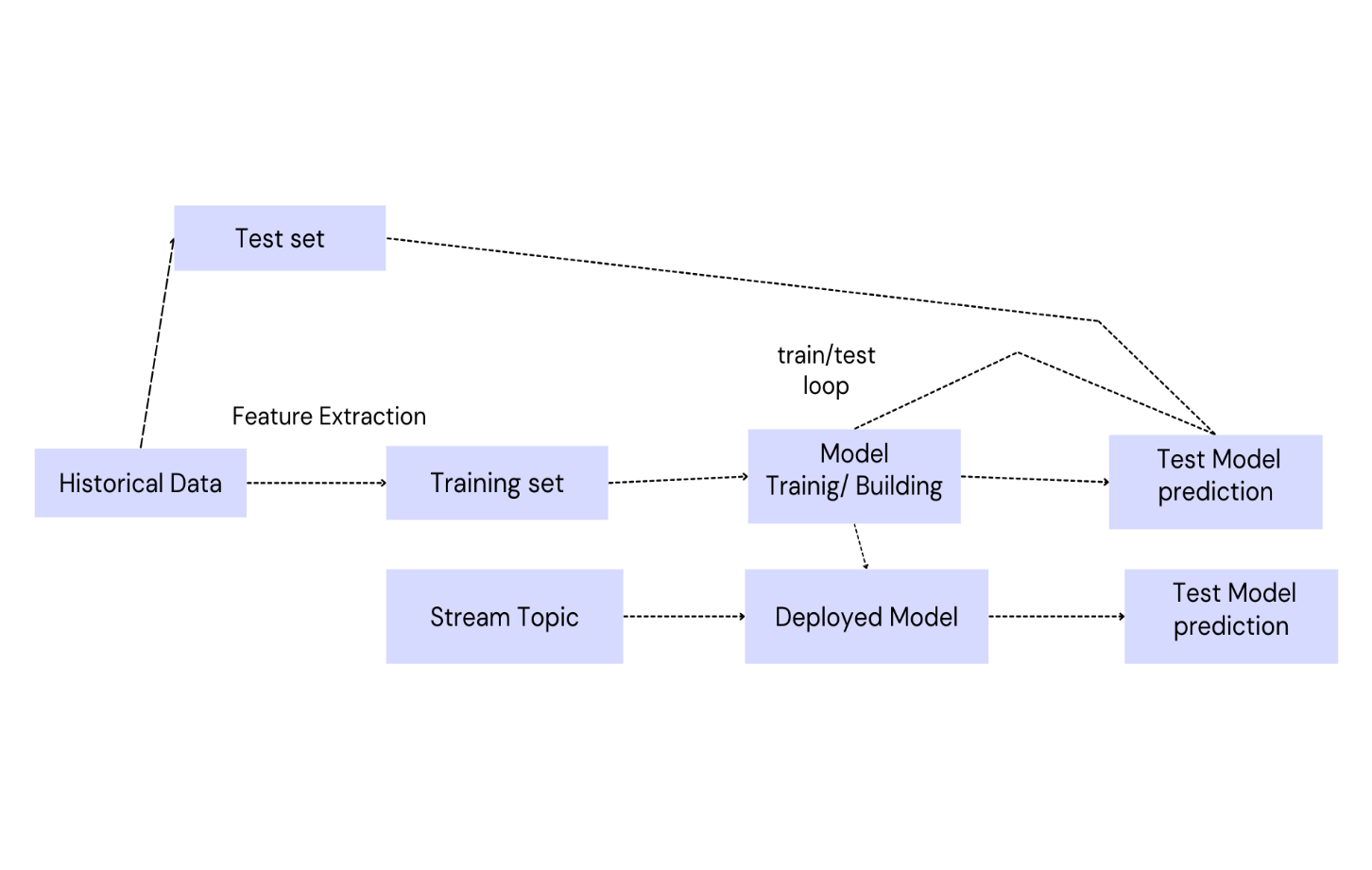
Economic feasibility for credit card fraud detection involves evaluating whether the benefits of implementing a fraud detection system, such as reduced financial losses from fraud, outweigh the costs of development, implementation, and maintenance, including factors like data processing and infrastructure expenses.

**4.3 Operational Feasibility**

Operational feasibility for credit card fraud detection involves assessing if the proposed system can be smoothly integrated into existing processes, accepted by users, adequately supported, and effectively measured for performance, considering available resources and managing potential risks.

**UML Diagram :**

**1.1 E-R Diagram**



**Limitations/Drawbacks**

Discussion of potential limitations of the implemented algorithms.

Challenges in real-time fraud detection.

Data quality issues and their impact on model performance.

Scalability concerns with increasing transaction volume.

**Future Enhancements**

Proposed enhancements to improve the system's performance and capabilities.

Integration of additional machine learning algorithms

or techniques.

Incorporation of advanced anomaly detection methods.

Enhancement of the GUI for better user experience.

**Conclusion**

Isolation Forest detected 73 errors versus Local Outlier Factor detecting 97 errors vs. SVM detecting 8516 errors

Isolation Forest has a 99.74% more accurate than LOF of 99.65% and SVM of 70.09

When comparing error precision & recall for 3 models , the Isolation Forest performed much better than the LOF as we can see that the detection of fraud cases is around 27 % versus LOF detection rate of just 2 % and SVM of 0%.

So overall Isolation Forest Method performed much better in determining the fraud cases which is around 30%.

We can also improve on this accuracy by increasing the sample size or use deep learning algorithms however at the cost of computational expense. We can also use complex anomaly detection models to get better accuracy in determining more fraudulent cases.

**Bibliography**

https://www.youtube.com

https://[www.kaggle.com](http://www.kaggle.com)

**Following books are used to SVM, Isolation Forest, LOF**

• Python books

• Fundamental mathematics and statistics

• Machine learning notes