Project Design Phase - I

Solution Architecture

Date	23 October 2023
Team ID	Team-592613
Project Name	Online Payments Fraud Detection Using ML
Maximum Marks	4 Marks
Team Members	Srajal Agarwal (Leader) Deepesh Dhruv Paras Garg

Solution Architecture

1. Data Collection and Preparation: We received a dataset that has been thoroughly pre-processed, which includes cleaning it by removing duplicates, dealing with missing values, and resolving outliers. Categorical data was converted to numerical format using techniques such as one-hot encoding, and numerical characteristics were scaled and normalised to maintain consistency.

2. Data Splitting: We divided our dataset into training, validation, and test sets. We used an 80-10-10 split ratio to ensure that both fraudulent and lawful transactions are distributed consistently across these groupings.

3. Model Selection: We investigated many machine learning methods. In addition, we compared methods such as Gradient Boosting, Neural Networks, and Support Vector Machines. The choice is based on the unique needs of our fraud detection system.

4. Model Training: Using techniques and methodologies given by machine learning libraries such as scikit-learn, the selected models were rigorously trained on the training dataset. Cross-validation is used to evaluate model performance and reduce the danger of overfitting. To improve model performance, hyperparameter tuning approaches such as grid search and random search have been utilised.

5. Threshold Adjustment: We've adjusted thresholds depending on system requirements to create a balance between false positives and false negatives. To identify an optimum threshold, methods such as ROC analysis have been used.

6. Model Evaluation: To guarantee the efficacy of our models, we thoroughly examined them using measures such as accuracy, precision, recall, F1-score, and ROC AUC.

7. Model Deployment: To collect transaction data, our model is deployed using a Flask-based interface, with cloud services such as AWS Lambda, Azure Functions, or Google Cloud Functions for real-time predictions. For effective deployment and administration, we've also used containerization technologies like Docker and Kubernetes.

8. Scalability and cloud integration: For scalability and resource management, we put the system on cloud platforms such as AWS, Azure, or Google Cloud, which provide auto-scaling. Scalable data storage is achieved through the usage of cloud databases and data lakes.

9. Monitoring and Alerting: To ensure continuous performance, we've integrated strong monitoring technologies like Prometheus or DataDog, as well as alerting systems that warn administrators of performance concerns as soon as they occur.

10. Security: Security is a top priority for us, therefore we use data encryption, access limits, and secure communication protocols to protect critical information. Furthermore, we closely follow data protection standards, including GDPR compliance via data anonymization and user consent processes.

11. Reporting and Analysis: We've created reporting tools like dashboards and visualisation tools to gather insights and data, giving you a complete picture of the system's performance and fraud detection findings.

12. Compliance: We are committed to adhering to all legal and regulatory obligations. To guarantee full compliance, we regularly monitor and upgrade our system to match with current standards and consult with legal experts.

13. Education and Training: Our crew is knowledgeable about how to use and maintain the system, evaluate findings, and respond to fraud situations. Our plan includes regular training and instruction.

14. Maintenance and Improvement: We've built a disciplined plan for regular upgrades and model fine-tuning to react to emerging fraud strategies and patterns. Through ongoing research and

participation with industry journals, we remain up to date on the newest advances in fraud detection and machine learning.

Solution Architecture Diagram

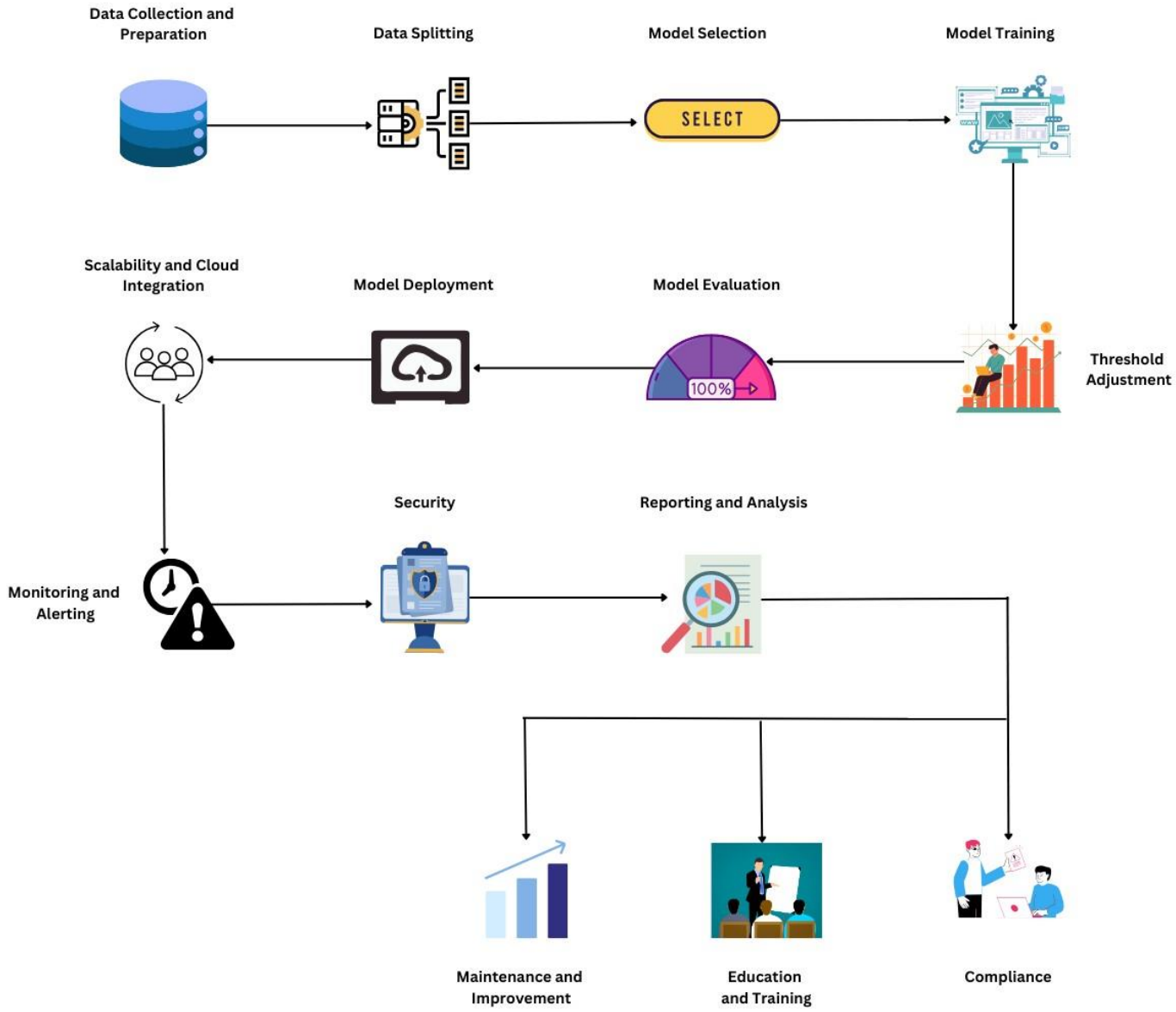

Figure 1: Solution Architecture Diagram for “Online Payments Fraud Detection Using ML”