Business Requirement Document (BRD)

Project	SolaceTrust Fraud Detection System
Purpose	The purpose is to improve fraud detection in SolaceTrust Bank using advanced data science methods.
Owners	**Your name**
Stakeholders	Sam Carter, CIO, SolaceTrust, Kazuki Murakami, Data Science Team Lead, SolaceTrust
Approved by	**Your stakeholder has to sign this off**
Status	~~Draft /~~ Under Review ~~/ Completed~~

Date: 15/09/2023

Project Overview:

Background: In recent times, we have uncovered irregularities in our financial data that have cast a shadow over our SolaceTrust's performance. After careful examination, it has become evident that these irregularities are not isolated incidents but are indicative of fraudulent activities that have affected our financial standing and eroded our stakeholders' trust.

Project Significance: The SolaceTrust Fraud Detection System project represents a pivotal initiative aimed at fortifying our organization's financial integrity, enhancing security measures, and ensuring regulatory compliance. In response to the growing challenges of fraudulent activities in the financial sector, this project leverages advanced data science methodologies to revolutionize our approach to fraud detection.

Business Objectives

Primary Objectives

1. **Enhanced Fraud Detection:** Our primary objective is to significantly enhance our ability to detect and prevent fraudulent activities within our organization's financial transactions. We strive to minimize financial losses and protect our customers from fraudulent activities.
2. **Operational Efficiency:** By automating and streamlining fraud detection processes, we aim to improve operational efficiency, reduce response times to potential fraud incidents, and enhance our team's productivity.
3. **Regulatory Compliance:** Ensuring strict adherence to industry-specific regulations, such as PCI DSS and GDPR, is paramount. This project will establish robust compliance mechanisms to meet these requirements.

Secondary Objectives

User Experience: In addition to enhancing our fraud detection capabilities, this project places a strong emphasis on improving the user experience for our fraud analysts and investigators. Through the implementation of user-friendly dashboards and early detection features, we aim to empower our team to work more effectively.

Data Privacy and Security: Data privacy and security are non-negotiable aspects of this project. We are dedicated to safeguarding customer information and maintaining their trust by implementing the highest standards of data protection.

Scope:

Backend Development
Data Collection and Preparation: This phase involves the collection of transactional data from multiple sources, comprehensive data preprocessing.
Data Pipelines: Real-time data pipelines will be implemented to ensure the availability of timely data for analysis.
Exploratory Data Analysis (EDA): EDA will provide critical insights into data patterns and inform feature engineering for enhanced fraud detection.
Model Development: Cutting-edge machine learning models will be developed and trained on labeled historical data to achieve superior fraud detection accuracy.

Frontend Development
User Dashboards: User-friendly dashboards will be designed to empower fraud analysts and investigators with real-time monitoring capabilities.
Early Detection Features: Early detection mechanisms will provide alerts and insights into potential fraudulent activities, enabling swift responses.
Fraudulent User Identification: Tools for identifying and flagging fraudulent users will streamline investigative processes.
Reporting: Robust reporting capabilities will facilitate communication and compliance with regulatory authorities.

High-Level Requirements:

1. **High Accuracy Fraud Detection:** The system must achieve a high level of accuracy in detecting fraudulent transactions to minimize financial losses and maintain customer trust.
2. **Real-time Processing and Alerting:** The system should be capable of processing and analyzing transactions in real-time, promptly identifying potential fraud, and alerting relevant stakeholders.
3. **Compliance with Data Privacy Regulations:** The system must adhere to all relevant data privacy and security regulations, including but not limited to GDPR, PCI DSS, and industry-specific standards.

Key Features:

1. **Advanced Machine Learning Models:** Implement and continuously update state-of-the-art machine learning models for fraud detection, including supervised and unsupervised techniques. Particularly, we will focus on autoencoders for anomaly detection, DBSCAN to identify group of transactions and prophet for time-series analysis. If time allows, we will also use NLP to process text data.
2. **Real-time Monitoring and Alerts:** Provide user-friendly dashboards for real-time monitoring of transactions, including early detection features and instant alerts for potential fraud cases.

3. **Comprehensive Reporting and Audit Trail:** Develop reporting capabilities for the generation of comprehensive reports for regulatory authorities and maintain a detailed audit trail of all transactions and system activities.

Risks and Mitigation:

- **Data Privacy Concern:** handling of sensitive customer transaction data for fraud detection purposes could raise privacy concerns among customers
- **Mitigation:** Implement robust data encryption and transparent privacy communication to maintain customer trust.
- **Risk 2:** Technical Complexities Delaying Project Milestones
- **Mitigation:** Embrace Agile methodologies, ensure an experienced team, and plan for contingencies to prevent project delays.

Timeline: (you can use a Gantt Chart instead, link to excel file)

- **Start Date:** [Start Date]
- **End Date:** [End Date]
- **Milestones:**
    1. Planning and Data Prep (Week 1-12):
        a. Define objectives and gather data.
        b. Prepare data and set up the team.
    2. Model Development and Dashboards (Week 13-20):
        a. Create fraud detection models.
        b. Build real-time dashboards.
    3. Integration and Security (Week 21-28):
        a. Integrate with external systems.
        b. Ensure data security and compliance.
    4. Testing and User Training (Week 29-38):
        a. Test system and train users.
    5. Deployment and Ongoing Maintenance (Week 39 onwards):
        a. Deploy system.
        b. Monitor, update, and maintain.

Key Deliverables:

1. **Advanced Predictive Models Demonstrating Improved Accuracy**
2. **User-friendly Dashboard Prototypes Showcasing Data Visualization of transaction data**
3. **Functional API Connecting our models to External Platforms**

APPENDIX

A more comprehensive timeline:

Project Phases and Tasks:

1. Project Initiation (Week 1-2):
   - Define project objectives and scope.
   - Identify stakeholders and establish a project team.
   - Develop a project plan and timeline.
2. Data Preparation (Week 3-6):
   - Gather transaction data and relevant datasets.
   - Perform data cleaning and data transformation.
   - Create a data processing pipeline.
3. Model Development (Week 7-16):
   - Research and select appropriate machine learning algorithms.
   - Develop and train initial fraud detection models.
   - Conduct exploratory data analysis (EDA) and feature engineering.
   - Fine-tune models for optimal performance.
4. Real-time Monitoring and Dashboard Development (Week 17-20):
   - Design and develop real-time monitoring interfaces and dashboards.
   - Implement user-friendly data visualization tools.
   - Ensure the dashboard's responsiveness and interactivity.
5. API Integration (Week 21-24):
   - Develop and test a functional API for external platform integration.
   - Establish data synchronization capabilities with external systems.
   - Conduct thorough API testing and validation.
6. Data Privacy and Compliance (Week 25-28):
   - Implement data encryption and security measures.
   - Ensure compliance with data privacy regulations (e.g., GDPR, PCI DSS).
   - Communicate transparently with customers about data usage and privacy.

7. User Training and Testing (Week 29-32):
   - Train fraud analysts and investigators on system usage.
   - Conduct user acceptance testing (UAT) to validate system functionality.
   - Address any usability issues and gather user feedback.
8. Comprehensive Reporting (Week 33-36):
   - Develop reporting capabilities for generating regulatory reports.
   - Ensure audit trail completeness and integrity.
   - Test report generation processes and templates.
9. Final Testing and Quality Assurance (Week 37-38):
   - Conduct end-to-end testing of the entire system.
   - Address and resolve any identified issues or bugs.
   - Perform security testing and penetration testing.
10. Deployment and Go-Live (Week 39-40):
    - Prepare for the production environment deployment.
    - Monitor system performance in the production environment.
    - Officially launch the Data Science Fraud Detection System.
11. Ongoing Monitoring and Maintenance (Week 41 onwards):
    - Establish continuous monitoring of system performance and fraud detection rates.
    - Implement regular model updates and improvements.
    - Conduct periodic security audits and compliance checks.