**Smarter Reconciliation and Anomaly Detection using Gen AI**

**1. Introduction**

The project focuses on leveraging Generative AI to enhance financial reconciliation processes and anomaly detection. Traditional reconciliation methods are often manual, time-consuming, and prone to errors. By incorporating AI-driven techniques, we aim to automate matching processes, detect anomalies efficiently, and improve financial accuracy.

The Anomaly Detection AI Model is designed to analyse banking transactions and identify anomalies in balance differences for each account over time. The model detects unusual spikes or deviations from historical transaction patterns, enabling proactive fraud detection and financial reconciliation.

**2. Objectives**

The purpose of this document is to outline the functional and technical requirements for implementing an AI-based anomaly detection system for banking transactions. The system will help the support team upload transaction data, process anomalies, and generate reports.

* Automatically detect data anomalies by comparing real-time data against historical baselines.
* Provides insights into potential root causes of detected anomalies.
* Integrates with existing reconciliation tools to streamline the anomaly identification process.
* Reduces manual effort and minimizes human error in anomaly detection.

**3. System Architecture**

**Anomaly Detection**

Anomaly Identified

Data Processing

Anomaly Detection Model

Input Real Time Data

Predicted Anomaly Label

Prediction Model

* Ingest banking transaction data (Excel format)
* Analyse balance differences over time
* Detect anomalies based on deviations from historical patterns
* Generate reports with anomaly classification and comments
* Provide a web-based interface for users to upload files and retrieve results

**3.1 Data Sources**

* General Ledger (GL) Systems
* Subledger and Transactional Databases
* External Financial Systems (e.g., Banking Systems, ERP)
* Historical Reconciliation Data

**3.2 Data Ingestion and Preprocessing**

* Data Extraction using APIs and batch processing
* Data Standardization and Normalization
* Handling Missing or Incomplete Data

**3.3 AI Model for Reconciliation**

* **Matching Algorithm:** AI-driven fuzzy matching for transaction pairing
* **Anomaly Detection:** Utilizing machine learning models such as Isolation Forest
* **Predictive Insights:** Forecasting reconciliation patterns and trends

**3.4 User Interface & Reporting**

* Web-based dashboard for real-time reconciliation status
* Alert mechanisms for detected anomalies
* Drill-down capabilities for investigation

**4. AI Model Selection**

* **Supervised Learning:** Training on historical labelled reconciliation data
* **Unsupervised Learning:** Identifying anomalies using clustering and anomaly detection techniques
* **LLMs & Gen AI:** Generative AI models to auto-explain discrepancies and suggest corrections

**5. Implementation Plan**

**5.1 Overview of the Model**

We have developed an AI-based Anomaly Detection System to identify unusual transactions in financial data.

The model automatically analyses each account's transactions and flags anomalies based on transaction patterns.

The system is built using Flask (backend API) and Machine Learning (Isolation Forest).

**5.2 Model Used for Anomaly Detection**

We use Isolation Forest, an unsupervised machine learning model designed to detect anomalies in large datasets.

It learns normal transaction behaviour for each account and flags transactions that deviate significantly.

Additionally, we apply Median-Based Thresholding to detect sudden spikes or drops in transaction amounts.

**5.3 How the Model Detects Anomalies**

1: Learn Normal Transaction Patterns

The model first calculates the median balance difference for an account.

It establishes a baseline of what normal transactions look like.

2: Identify Sudden Spikes or Drops

If a transaction changes more than 5x the median, it is flagged as an anomaly.

Example: If the previous balance difference was -0.83 and suddenly jumps to -40,560.95, it is an anomaly.

3: Machine Learning-Based Detection

Isolation Forest scans the entire dataset and assigns an "anomaly score" to each transaction.

If a transaction is statistically rare, it is marked as an anomaly.

4: Assign a Reason for Each Anomaly

Each anomaly is labelled with why it was detected (e.g., "Sudden Spike in Balance" or "Unusual Transaction Pattern").

This helps in investigation and audit purposes.

**Technologies Used to Build the Anomaly Detection AI Model**

To ensure a scalable, efficient, and secure anomaly detection system, the following technologies and frameworks are used:

**1. Backend Development**

* **Python** – Core language for data processing, AI model implementation, and backend logic.
* **Flask/FastAPI** – Web framework for exposing APIs to support file uploads and result retrieval.
* **Pandas & NumPy** – Libraries for data manipulation and numerical calculations.
* **Scikit-learn** – For anomaly detection model development and statistical analysis.

**2. Machine Learning & AI**

* **Unsupervised Learning (Isolation Forest, DBSCAN, or Custom Threshold-based Detection)** – To identify anomalies in transaction patterns.
* **Statistical Analysis (Median, Standard Deviation)** – For defining anomaly thresholds dynamically.

**3. Data Storage & Processing**

* **SQLite / PostgreSQL / MySQL** – Database options for storing transaction history and anomaly detection logs.
* **Excel (.xlsx) Processing** – Handled via openpyxl and pandas for seamless file operations.

**4. Challenges & Mitigation Strategies**

* **Data Quality Issues:** Implement robust data cleansing techniques
* **Model Accuracy:** Continuous retraining with updated data
* **Interpretability:** Use Explainable AI (XAI) techniques for better understanding

**7. Expected Benefits**

* Faster and more accurate reconciliation
* Early detection of fraud and errors
* Reduction in operational costs
* Enhanced decision-making with AI-driven insights

**8. Test Cases for Anomaly Detection Using Gen AI**

**8.1 Data Anomaly Detection**

* **Test Case 1:** Detect missing transaction entries
  + *Input:* Dataset with randomly removed transactions
  + *Expected Outcome:* AI model flags missing entries
* **Test Case 2:** Identify duplicate transactions
  + *Input:* Dataset with intentional duplicate records
  + *Expected Outcome:* AI flags duplicates with high confidence

**8.2 Reconciliation Discrepancies**

* **Test Case 3:** Mismatched transaction amounts
  + *Input:* Transactions with modified amounts
  + *Expected Outcome:* AI detects and highlights discrepancies
* **Test Case 4:** Currency conversion errors
  + *Input:* Transactions with incorrect currency conversions
  + *Expected Outcome:* AI identifies incorrect exchange rates and mismatches

**8.3 Fraud Detection**

* **Test Case 5:** Suspicious account activity
  + *Input:* Transactions with unusual volume or frequency
  + *Expected Outcome:* AI flags potential fraud cases
* **Test Case 6:** Unauthorized account modifications
  + *Input:* Changes in critical account details without authorization
  + *Expected Outcome:* AI detects and reports unauthorized changes

**8.4 Model Performance & Validation**

* **Test Case 7**: False positive and false negative analysis
  + *Input:* Balanced dataset with labelled anomalies and normal transactions
  + *Expected Outcome:* AI maintains optimal precision and recall
* **Test Case 8**: Model adaptability to new anomaly patterns
  + *Input:* Introduce new types of discrepancies not seen in training
  + *Expected Outcome:* AI generalizes well and adapts to new patterns

**9. Conclusion**

This project aims to revolutionize financial reconciliation by leveraging Generative AI to automate processes and detect anomalies. With an AI-powered approach, organizations can achieve greater accuracy, efficiency, and financial integrity.