Fraud Detection Using Machine Learning

1. Problem Statement

Financial fraud, especially through unauthorized fund transfers and cash withdrawals, is a critical challenge for banks and payment systems. The objective of this project is to build a **machine learning model** that can identify fraudulent transactions based on transaction details and account balance movements.

---

2. Dataset Overview

The dataset (Fraud.csv) contains transaction-level details such as:

- **step**: Time (hour/day of simulation)
- **type**: Transaction type (TRANSFER, CASH_OUT, PAYMENT, etc.)
- **amount**: Transaction amount
- **oldbalanceOrg / newbalanceOrig**: Source account balances before and after transaction
- **oldbalanceDest / newbalanceDest**: Destination account balances before and after transaction
- **isFraud**: Target variable (1 = fraud, 0 = legitimate)

---

3. Feature Engineering

To better capture fraudulent behavior, new features were engineered:

- **errorBalanceOrig** – Discrepancy in originating account balance.
- **errorBalanceDest** – Discrepancy in destination account balance.
- **balance_diff_orig** – Change in source account balance.
- **balance_diff_dest** – Change in destination account balance.
- **One-hot encoding** of transaction type.

These engineered features highlight unusual patterns like sudden balance drops or mismatches in expected balances.

---

4. Methodology

- **Data Preprocessing**: Cleaned dataset, created new features, and applied one-hot encoding.
- **Model Selection**: Random Forest Classifier was chosen due to its robustness and ability to handle imbalanced datasets.
- **Evaluation Metrics**: Classification report (Precision, Recall, F1-score) and Confusion Matrix were used to measure performance.

---

5. Results

- The model achieved strong **recall** on fraudulent transactions, meaning it successfully detected the majority of fraud cases.
- **Feature importance analysis** revealed the most influential predictors:
    1. **errorBalanceOrig** – Strong signal of inconsistencies in source balances.
    2. **balance_diff_orig** – Sudden, unexplained decreases indicate drained accounts.
    3. **amount** – Large amounts are highly suspicious.
    4. **type_TRANSFER & type_CASH_OUT** – High-risk transaction types linked to fraud.
    5. **errorBalanceDest** – Discrepancies in destination account balances often signal fraudulent inflows.

---

6. Insights & Interpretation

- **Balance Discrepancies**: Fraudsters often create mismatches in balances to quickly move funds; this shows up as strong predictive signals.
- **Transaction Types**: TRANSFER and CASH_OUT dominate fraudulent cases, aligning with known fraud patterns. PAYMENT and CASH_IN act as contrasting, low-risk categories.
- **Transaction Amounts**: Fraudsters aim for maximum profit, often transferring unusually large sums.
- **Timing (step)**: Some fraudulent activities are concentrated at specific times, indicating strategic attempts to avoid detection.

---

7. Recommendations

1. **Real-time Monitoring**: Implement alerts for transactions with large errors in balances (errorBalanceOrig, errorBalanceDest).
2. **High-Risk Transaction Types**: Apply stricter verification for TRANSFER and CASH_OUT above a certain threshold.
3. **Adaptive Rules**: Use time-based risk scoring to increase scrutiny during off-peak hours.
4. **Continuous Model Training**: Retrain with updated data to adapt to evolving fraud techniques.
5. **Hybrid Approach**: Combine machine learning with rule-based systems for better fraud coverage.

---

8. Conclusion

This project demonstrates that **machine learning, supported by engineered balance features and transaction type indicators, can effectively identify fraudulent behavior**. By focusing on discrepancies, transaction patterns, and amounts, financial institutions can significantly reduce fraud losses while maintaining smooth operations for legitimate customers.