Case Study: FraudDetectXNet – A Hybrid Model for Fraud Detection

---

1. Problem Statement and Objectives

Credit card fraud has become a serious concern with increasing digital transactions. Due to the rarity of fraudulent cases in large datasets, most traditional models fail to detect these minority events accurately. The objective of this case study is to design a robust and intelligent fraud detection model that:

- Accurately detects fraudulent transactions in real-world, imbalanced datasets
- Uses sequence-aware learning to capture transaction patterns
- Achieves high recall and low false-positive rate for business applicability

---

2. Data Preprocessing

- **Dataset**: Kaggle Credit Card Fraud Detection dataset (284,807 transactions, 492 frauds)
- **Missing Values**: None
- **Features**: 28 anonymized features (PCA components), 'Time', and 'Amount'
- **Class Distribution**: 0.17% fraud cases

Steps Taken: - Dropped 'Time' as it was not useful for modeling - Standardized 'Amount' feature using StandardScaler - Resampled data using **SMOTE** to balance fraud and non-fraud classes

---

3. Model Selection and Development

Proposed Model: FraudDetectXNet

- **Step 1**: Use **SMOTE** to oversample minority class in training set
- **Step 2**: Feed data to an **LSTM network** to capture temporal relationships
- **Step 3**: Extract intermediate representations (features) from LSTM output
- **Step 4**: Use **XGBoost** on these features for final classification

Justification: - LSTM handles sequential dependencies in time-series data - XGBoost provides robust classification with gradient boosting - The combination enables superior fraud detection performance

---

4. Visualizations and Insights

- **Class Distribution**: Visualization before and after applying SMOTE shows balanced data
- **Confusion Matrix**: Highlights a strong true positive rate for frauds
- **ROC Curve**: AUC Score ~0.97 indicating excellent classifier performance

Insights: - Traditional models struggle with fraud recall - Hybrid architecture significantly improves performance - Balanced training data is essential for minority class detection

5. Recommendations

- **Deploy hybrid models** like FraudDetectXNet in production for fraud-prone sectors
- Continuously update models with new transaction data
- Use cost-sensitive learning to minimize business impact of false negatives
- Integrate with real-time detection systems for faster response

6. Conclusion

FraudDetectXNet demonstrates that combining temporal learning and gradient boosting yields significant improvements in detecting rare fraudulent transactions. This architecture serves as a promising framework for real-world fraud detection systems.