**Method**

**Research Design** This study adopts a mixed-methods research design, integrating quantitative analysis of transactional datasets with qualitative insights from domain experts. The primary objective is to develop and evaluate an advanced fraud detection framework that combines the strengths of supervised machine learning and unsupervised anomaly detection algorithms. This hybrid approach aims to enhance detection accuracy while minimizing false positives, critical for maintaining customer trust and operational efficiency in financial transactions.

**Data Collection & Preparation**

* **Dataset Acquisition**: Transactional data is sourced from a leading financial institution, spanning a three-year period. The dataset encompasses approximately 10 million transactions, including attributes such as transaction amount, timestamp, merchant category code (MCC), user demographics, and a binary flag indicating fraudulent or legitimate transactions.
* **Data Cleaning**: Data is preprocessed to handle missing values through imputation techniques, excluding transactions with incomplete data necessary for analysis. Outliers are identified using statistical methods (e.g., Z-score, IQR) and either corrected or removed based on their impact on model performance.
* **Feature Engineering**: In addition to basic transaction attributes, engineered features include time-series analysis (e.g., transaction velocity, periodicity), behavioral profiling (e.g., deviation from usual purchase patterns), and network analysis (e.g., links between users and merchants). Expert interviews guide the selection and creation of these features to capture known fraud patterns.

**Model Development & Selection**

* **Supervised Learning Models**: A comparative assessment of machine learning classifiers—Random Forest, Gradient Boosting Machines (GBM), and a Long Short-Term Memory (LSTM) neural network—is performed. Hyperparameter tuning is executed via grid search or randomized search, optimizing for the area under the ROC curve (AUC-ROC).
* **Unsupervised Anomaly Detection**: Two techniques, Isolation Forest and One-Class Support Vector Machines (OCSVM), are implemented to detect anomalies in transaction patterns that do not rely on labeled data. Parameters are fine-tuned for maximum detection sensitivity without overwhelming false positive rates.

**Validation Strategy**

* **Cross-Validation**: A 10-fold cross-validation scheme is employed to robustly estimate model performance, ensuring each transaction is used for testing exactly once across all folds.
* **Performance Metrics**: Model evaluation focuses on precision, recall, F1-score, and the AUC-ROC, with emphasis on balanced detection given the class imbalance. Cost-sensitive learning is incorporated to assign higher penalties to false negatives (missed frauds).

**Implementation & Testing** The best-performing models from both supervised and unsupervised categories are integrated into a hybrid detection system. A threshold optimization strategy is devised based on a cost-benefit analysis to determine the optimal point for flagging transactions as potentially fraudulent.

**Ethics & Privacy**

* All data handling adheres strictly to GDPR and local data protection regulations. Personal identifiers are fully anonymized, and access to data is restricted to authorized researchers.
* Informed consent was secured from all interview participants, and their insights were anonymized before inclusion in the study.
* The research protocol was reviewed and approved by the institutional ethics committee, ensuring adherence to ethical standards throughout the research process.

By incorporating these detailed steps, the Method section now provides a comprehensive roadmap for conducting the study, enhancing its transparency, replicability, and adherence to best practices in research ethics and data handling.