**Assignment: Anomaly Detection in Credit Card Transactions**

**Objective:**

* To detect fraudulent transactions in credit card data using anomaly detection techniques.

**Dataset:**

* Use a publicly available credit card transaction dataset, such as the Credit Card Fraud Detection dataset from Kaggle.

**Tasks:**

1. **Data Exploration:**
   * Load the dataset and perform exploratory data analysis (EDA) to understand its structure and features.
   * Explore the distribution of transaction amounts, transaction times, and other relevant features.
   * Visualize the distribution of fraudulent and non-fraudulent transactions to understand the class imbalance.
2. **Data Preprocessing:**
   * Perform any necessary data preprocessing steps, such as handling missing values, scaling numerical features, and encoding categorical variables.
   * Split the data into training and testing sets.
3. **Anomaly Detection Techniques:**
   * Implement at least two anomaly detection techniques, such as:
     + Isolation Forest
     + Local Outlier Factor (LOF)
     + One-Class SVM
     + Autoencoders (if you're familiar with neural networks)
4. **Model Training:**
   * Train the anomaly detection models on the training data.
   * Tune hyperparameters if necessary using cross-validation or grid search.
5. **Evaluation:**
   * Evaluate the performance of each model using appropriate evaluation metrics, such as:
     + Precision, recall, and F1-score
     + Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC)
   * Compare the performance of different anomaly detection techniques and select the best-performing model.
6. **Visualization:**
   * Visualize the results of anomaly detection, such as:
     + Plotting the distribution of anomaly scores for fraudulent and non-fraudulent transactions.
     + Visualizing the ROC curve for each model.
7. **Interpretation:**
   * Interpret the results and provide insights into the effectiveness of each anomaly detection technique.
   * Discuss the trade-offs between false positives and false negatives in the context of fraud detection.

**Deliverables:**

* Jupyter Notebook or Python script containing the implementation of anomaly detection techniques, along with detailed explanations and visualizations.
* Report summarizing the findings, including model performance metrics and insights gained from the analysis.

**Bonus Tasks (Optional):**

* Implement feature engineering techniques to improve the performance of anomaly detection models.
* Explore ensemble methods for anomaly detection, such as combining multiple models for better detection accuracy.
* Deploy the best-performing anomaly detection model as a web service using Flask or FastAPI.

This assignment will give you hands-on experience with anomaly detection techniques and their application in real-world scenarios, particularly in fraud detection for credit card transactions. Feel free to modify or extend the tasks based on your interests and learning goals. Good luck with your practice!