CREDIT CARD FRAUD DETECTION

DIMENSIONALITY REDUCTION AND CLASSIFICATION

By Himmler Benitez

DECEMBER 2025

Content

1. Objective

The main objective of this analysis is to apply dimensionality reduction techniques to a highly imbalanced credit card transaction dataset to uncover latent behavioral patterns that improve downstream fraud detection performance.

Its focus is on dimensionality reduction, rather than clustering, with the goal of learning compact representations that preserve meaningful transaction structure while reducing noise and redundancy.

The business value of the project is to improve the robustness of the model in the presence of class imbalances and reduce the computational complexity while enabling a more interpretable classification model which will lead to having a more insightful vision of the risk of each transaction.

2. Dataset Description

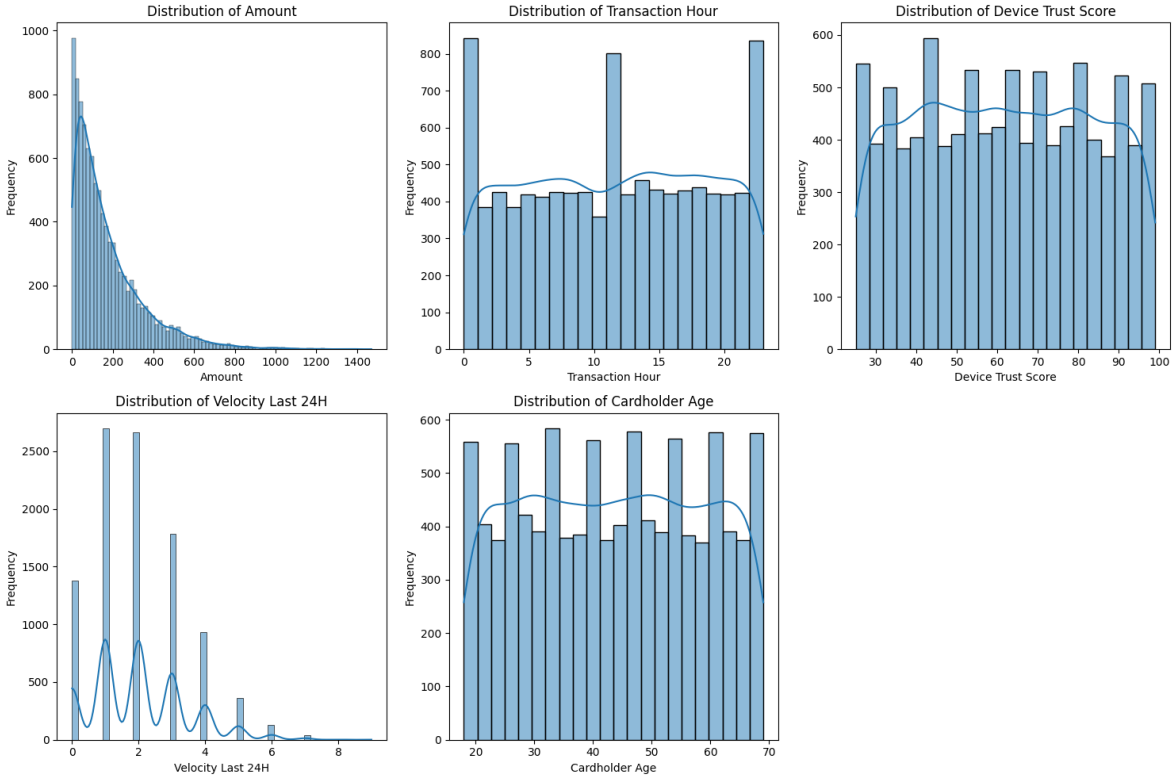
The dataset used in this analysis is the Credit Card Fraud Detection dataset obtained from Kaggle. It contains simulated credit card transactions labeled as either legitimate or fraudulent.

It is an extremely unbalanced dataset where there are only 151 samples of the fraudulent class, in contrast to the 9849 samples for the legitimate class. Making it a realistic and challenging fraud detection scenario.

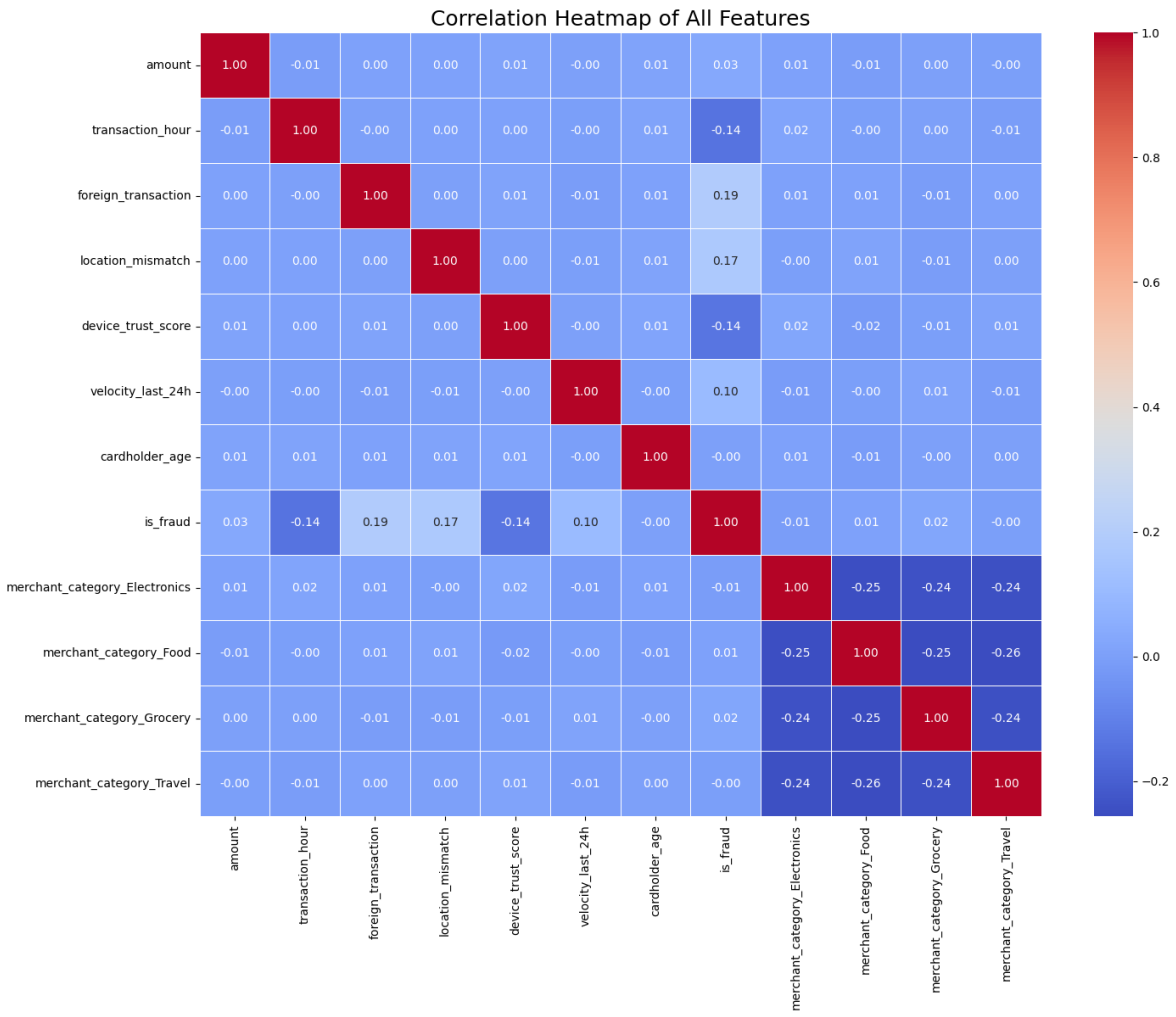
3. Data Exploration and Preparation

Data Exploration

- Highly right-skewed distributions for transaction amount and velocity.
- Near-uniform distributions for contextual and demographic features.



The feature distributions reveal strong skewness, discrete behavioral patterns, and weak standalone predictors, reinforcing the need for dimensionality reduction to uncover latent transaction behaviors critical for effective fraud detection.



The correlation analysis shows no dominant linear drivers of fraud, indicating that fraudulent behavior is governed by subtle, multivariate patterns.

Overall Observation

- Most feature pairs exhibit very low linear correlation ($|\rho| \approx 0.00–0.10$).
- This indicates low redundancy and limited linear dependence across features.
- Implication: The dataset is well-suited for dimensionality reduction to capture latent structure rather than simple feature elimination.

Data Preparation

A pipeline was created to allow clean preparation of the data and its input to the model. Actions taken:

- Applied standardization to numerical features to normalize scale.
- Log-transformed skewed variables to stabilize variance.
- Performed stratified train-test splitting to preserve fraud ratios.
- Clean NA values.

To address the poorly represented minority class, SMOTE was selected to create synthetic data and compensate the representation in the training split.

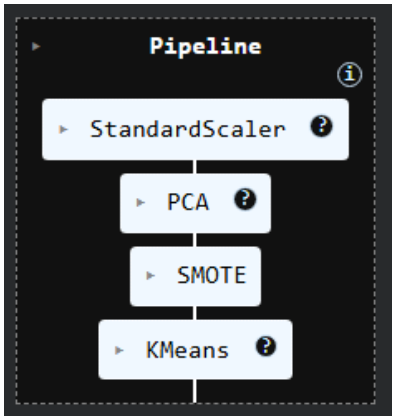
4. Unsupervised Model Variations

Model 1: PCA (95% Explained Variance):

- Linear dimensionality reduction.
- Reduced the original feature space to a smaller set of principal components.
- Served as the foundation for all subsequent modeling.

Model 2: K-Means Clustering on PCA-Reduced Data:

- Applied to the PCA-transformed feature space.
- Used to explore latent groupings of transactions.
- Cluster assignments were analyzed in relation to fraud labels (not used during training).



5. Classifier Model

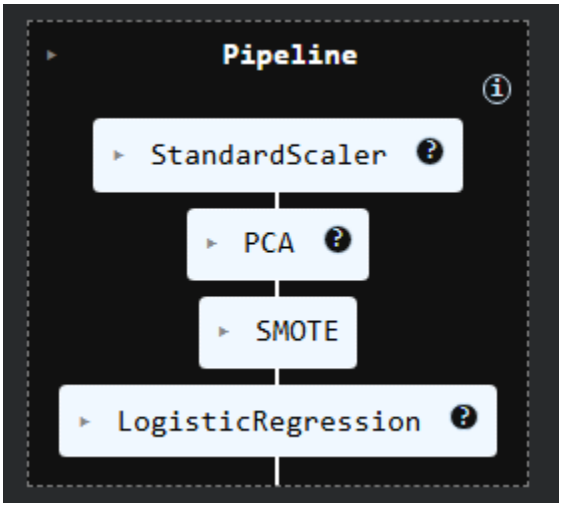
A function was defined to create the pipeline and the parameters distribution statement to execute a random search of the best hyperparameters comparing PCA (90% and 95%) and Kernel PCA.

```python
def get_pipeline(clf):
  pipeline = Pipeline(steps=[
      ("scaler", StandardScaler()),
      ("pca", PCA()),
      ("smote", SMOTE(random_state=42)),
      ("clf", clf)
  ])
  return pipeline

param_distributions = [
    # PCA lineal
    {
        "pca": [PCA(random_state=42)],
        "pca__n_components": [0.90, 0.95],
        "clf__C": loguniform(1e-2, 1e1),
        "clf__class_weight": [None, "balanced"]
    },
    # Kernel PCA
    {
        "pca": [KernelPCA(kernel="rbf", fit_inverse_transform=False)],
        "pca__n_components": [10, 15, 20],
        "pca__gamma": loguniform(1e-3, 1e-1),
        "clf__C": loguniform(1e-2, 1e1),
        "clf__class_weight": [None, "balanced"]
    }
]
```

```python
from sklearn.model_selection import GridSearchCV, StratifiedKFold
from sklearn.model_selection import RandomizedSearchCV, StratifiedKFold

# Creating the pipeline for logistic regresion
log_pipeline = get_pipeline(log_clf)

cv = StratifiedKFold(n_splits=3, shuffle=True, random_state=42)

random_search = RandomizedSearchCV(
    estimator=log_pipeline,
    param_distributions=param_distributions,
    n_iter=20,
    scoring="roc_auc",
    cv=cv,
    n_jobs=-1,
    verbose=1,
    random_state=42
)

random_search.fit(X_train, y_train)

print("Best parameters found:")
print(random_search.best_params_)

print(f"Best CV ROC-AUC: {random_search.best_score_:.4f}")
```

Logistic Regression

- Trained in PCA-reduced and SMOTE-balanced data.
- Used to evaluate whether the reduced feature space retains fraud-discriminative power.
- Performance assessed using recall, precision, F1-score, and ROC-AUC.



The logistic regression model demonstrates strong discriminative capability when evaluated on the test set, particularly in its ability to identify fraudulent transactions. It reports a 97% accuracy and a 0.993 ROC-AUC which translates into proper classification for fraudulent transactions.

```
ROC-AUC en test: 0.9934
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.97      0.98      1970
           1       0.31      0.93      0.47        30

    accuracy                           0.97      2000
   macro avg       0.66      0.95      0.73      2000
weighted avg       0.99      0.97      0.98      2000
```

It is a critical requirement in fraud detection systems where missed fraud is costly. However, the relatively low precision indicates a higher rate of false positives, reflecting a trade-off between fraud recall and alert accuracy.

6. Conclusion

This project examined whether unsupervised dimensionality reduction can improve fraud detection in a highly imbalanced credit card transaction dataset. Principal Component Analysis (PCA) and Kernel PCA were used as unsupervised feature learning methods, and their representations were evaluated using a downstream logistic regression classifier. Also, a K Means model was used for comparison.

A RandomizedSearchCV approach was applied to compare linear PCA (90% and 95% explained variance) with Kernel PCA using an RBF kernel. Pipelines and stratified cross-validation were used to prevent data leakage and ensure a fair and efficient comparison.

The results indicate that Kernel PCA consistently outperformed linear PCA, suggesting that fraudulent behavior is driven by non-linear patterns not captured by linear projections. The non-linear latent space learned by Kernel PCA led to improved classification performance.

Logistic regression achieved 97% accuracy and an ROC-AUC of 0.993, indicating strong discrimination between fraudulent and legitimate transactions. Importantly, the model reached a fraud recall of 0.93, showing effective detection of most fraudulent cases. Although fraud precision was moderate (0.31), this trade-off is expected and acceptable in fraud detection, where recall is a higher priority.

Overall, the findings show that unsupervised dimensionality reduction is an effective tool for detection of fraud in imbalanced settings. By learning compact and non-linear representations, Kernel PCA improves performance and robustness without relying on complex supervised models.

In conclusion, I recommend combining Kernel PCA with a simple and interpretable classifier that provides a strong and scalable baseline for fraud detection, balancing performance, and interpretability for practical deployment.

For next research, I suggest implementing a 1D CNN or SVM model that could get a better performance for this type of data.