Anomaly Detection for Time Series Data

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Jesse Li, Jingwen Zhang

Prof. Ali Hirsa; Advisor: Yossi Cohen

WELLINGTON MANAGEMENT

Overview

Our objective is to detect anomalies within financial data using daily transaction counts and amounts from a dataset of 51 companies spanning 2016–2022. By employing regime detection, a voting-based anomaly detection system, and refining a deep learning model, we aim to develop a robust anomaly detection framework that adapts to different economic regimes and enhances anomaly detection accuracy.

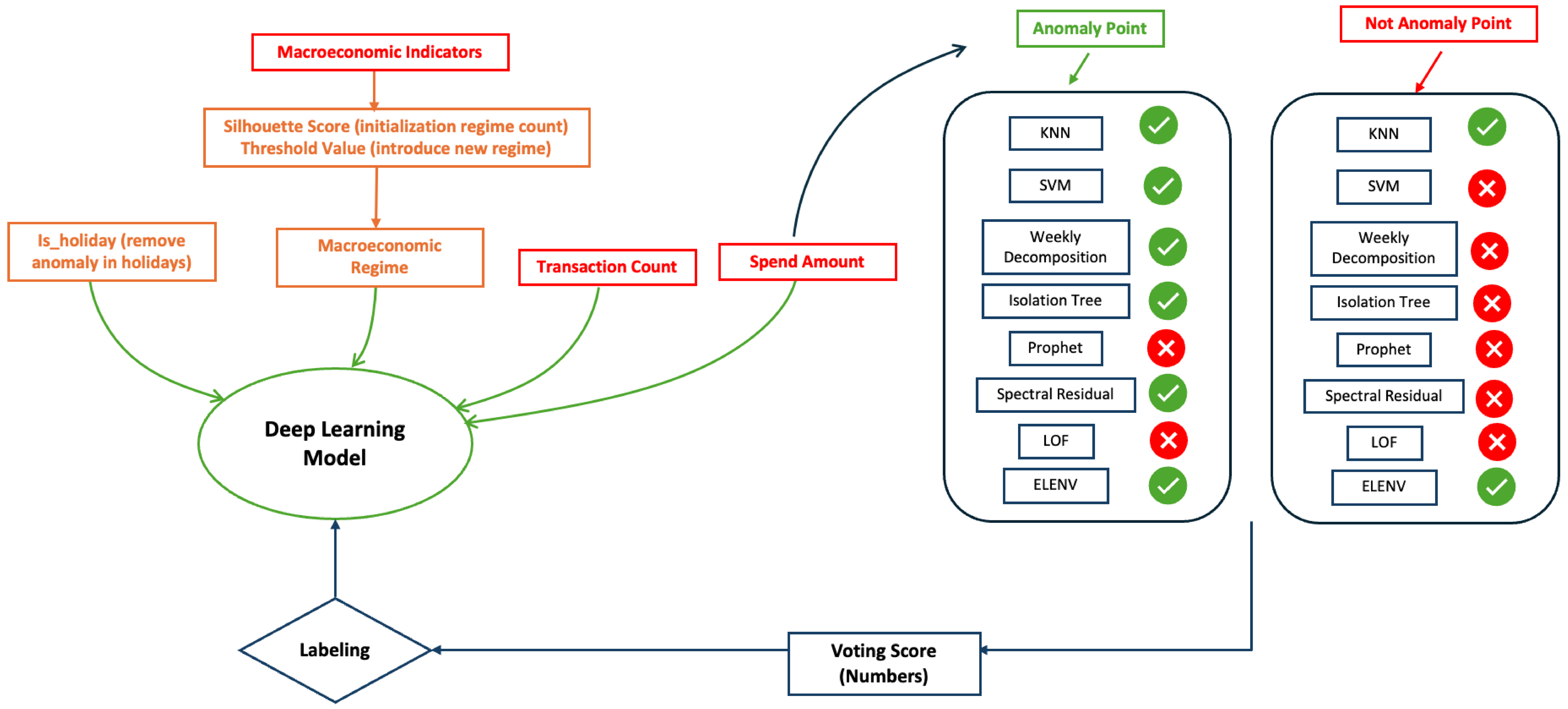
- **Regime Detection:** Building on a Robust Rolling Regime Detection(R2-RD) framework, we introduced a silhouette score to determine the optimal regime count and a threshold value function to assess the need for new regimes. We also incorporated six macroeconomic indicators, including GDP, Unemployment, and CPI, covering the period from 2016 to 2022 as a supplementary dataset for regime detection.
- **Anomaly Detection Algorithms and Voting System:** We integrated the Local Outlier Factor (LOF) and Elliptic Envelope algorithms in addition to the formerly selected six algorithms into a voting system to enhance the detection accuracy. We analyzed the distribution of anomalies across different regimes, plotting their percentages and amounts for each algorithm, providing a deeper understanding of the anomalies. An evaluation system was also developed, calculating precision, recall, and F1 scores to assess and refine the algorithms, ensuring the replacement of unfit models.
- **Deep Learning Model:** The results of the eight models were used as labels to train a deep learning model which was further refined through parameter tuning, experimenting with different epochs, batch sizes, learning rates, and optimizers.

By integrating these three components, our approach successfully highlights anomalies across various time periods and also offers a refined method for detecting anomalies in different economic scenarios. Future work could focus on these areas to further enhance the model's performance.

Data Description

We utilized two distinct datasets:

1. **Financial Institution Dataset:** This dataset provided by Wellington Management Company comprises daily transaction data from 51 companies, covering the period from January 1, 2016, to December 31, 2022. For each company, the dataset includes two transaction count and transaction amount on daily basis.
2. **Macroeconomic Indicator Dataset:** Sourced from the Federal Reserve Economic Data (FRED), this dataset includes a selection of macroeconomic indicators relevant to the period from 2017 to 2022 on monthly basis. The indicators include: GDP, CPI, Unemployment, Interest Rate, Retail Sale and Trade Balance. It was used to supplement the financial institution dataset for regime detection.

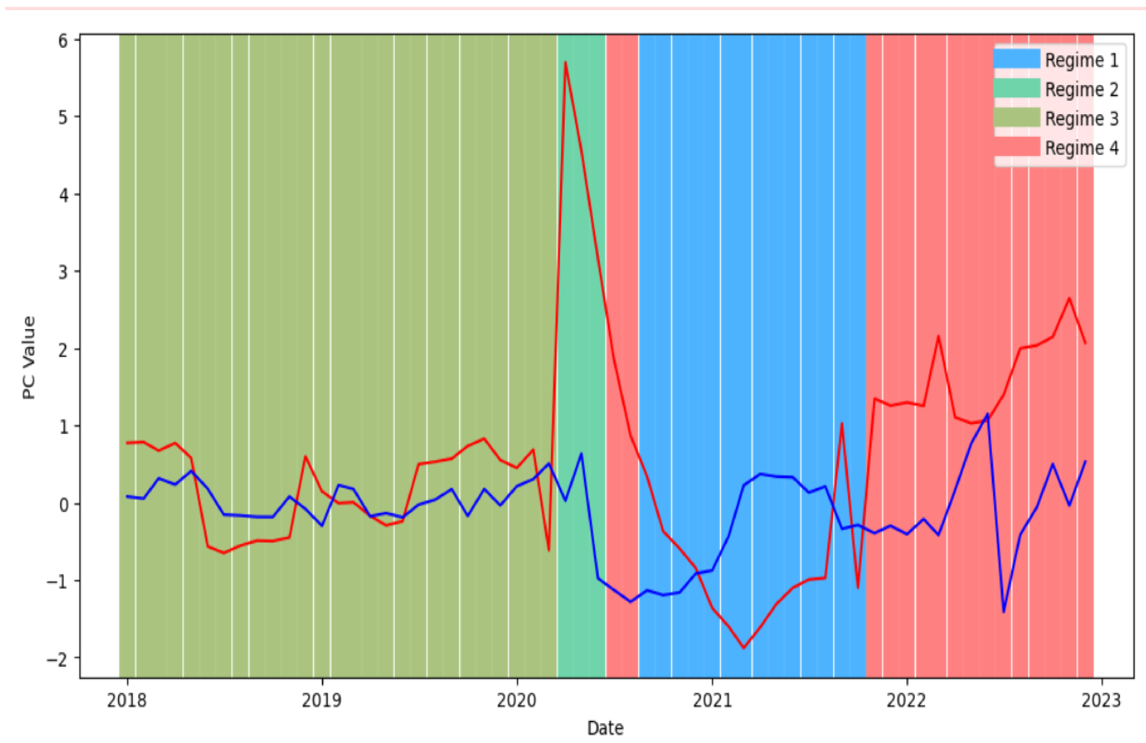


Phase 1: Regime Detection

1. Silhouette Score: We used the silhouette score to determine the optimal number of regimes. This metric measures how well each data point fits within its assigned regime compared to other regimes, helping us select the most appropriate number of regimes. Our analysis indicated that two regimes were optimal based on the highest silhouette score.

2. Threshold Value Function: To dynamically decide when to introduce new regimes, we implemented a threshold value function. This function calculates the assignment costs for adding a new regime and compares them to predefined thresholds (set at 95% and 105% of the initial assignment cost). If the assignment cost falls within this range, a new regime is introduced, allowing the model to adapt to new data.

3. Macroeconomic Regime Detection: We extended our regime detection analysis by using six macroeconomic indicators (GDP, CPI, unemployment rate, interest rates, retail sales, and trade balance) from 2017 to 2023. Through this analysis, we identified four distinct economic regimes.



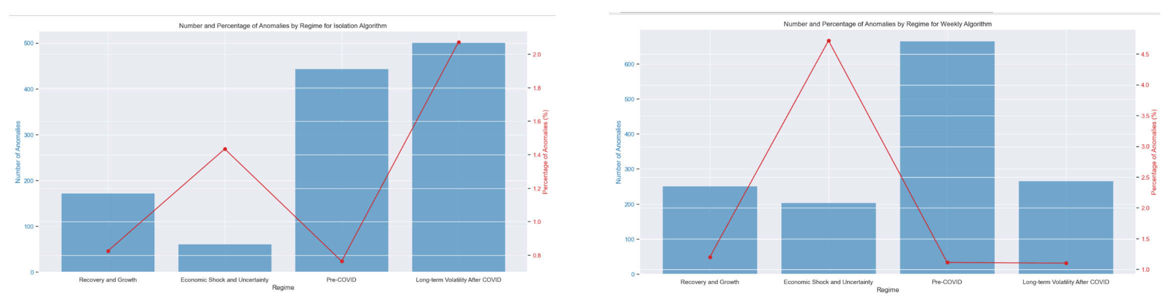
Phase 2: Voting-based Anomaly Detection

1. Algorithm Integration: Based on the 6 algorithms from the last group, we added 2 anomaly detection algorithms—Local Outlier Factor (LOF) and Elliptic Envelope (ELEVN)—and incorporated multiple algorithms, including LOF, ELEVN, Isolation Forest, KNN, and more, into a voting system. The system determined anomalies by consensus—if a data point was flagged by the majority of algorithms, it was considered an anomaly.

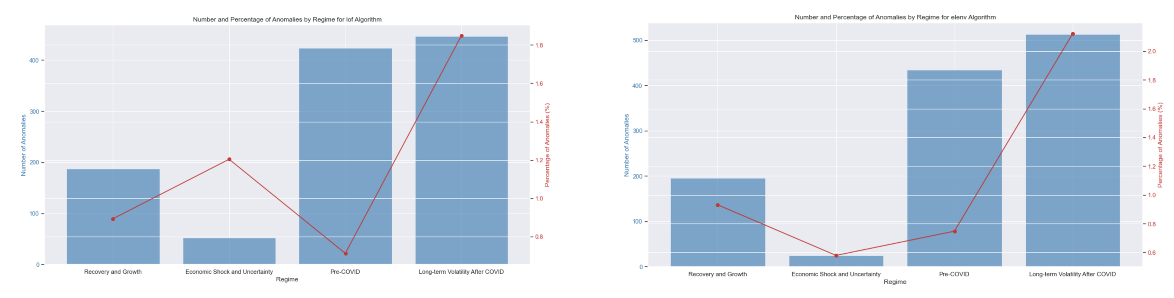
2. Evaluation of Algorithms: We considered the results from the voting system as the true anomaly points and then evaluated the performance of each algorithm using precision, recall, accuracy, F1 score, and AUC. According to the results, all algorithms are applicable while isolation tree, KNN, SVM, LOF, ELEVN and weekly multiplicative decomposition have good performance:

Algorithm	Precision	Recall	Accuracy	F1 Score	Applicable
Isolation Tree	0.95	0.92	0.97	0.94	Good
Prophet	0.52	0.17	0.85	0.26	Moderate
KNN	0.93	0.91	0.97	0.92	Good
Weekly Decomp.	0.97	0.90	0.98	0.93	Good
SVM	0.93	0.91	0.97	0.92	Good
Special Residual	0.93	0.91	0.97	0.92	Good
LOF	0.90	0.88	0.97	0.89	Moderate
Elliptic Envelope	0.95	0.92	0.98	0.93	Good

3. Macroeconomic Regime Application: We applied the detected regimes (from our regime detection phase) to analyze the distribution of anomalies across different economic contexts. This step allowed us to observe how different algorithms performed under varying economic conditions.



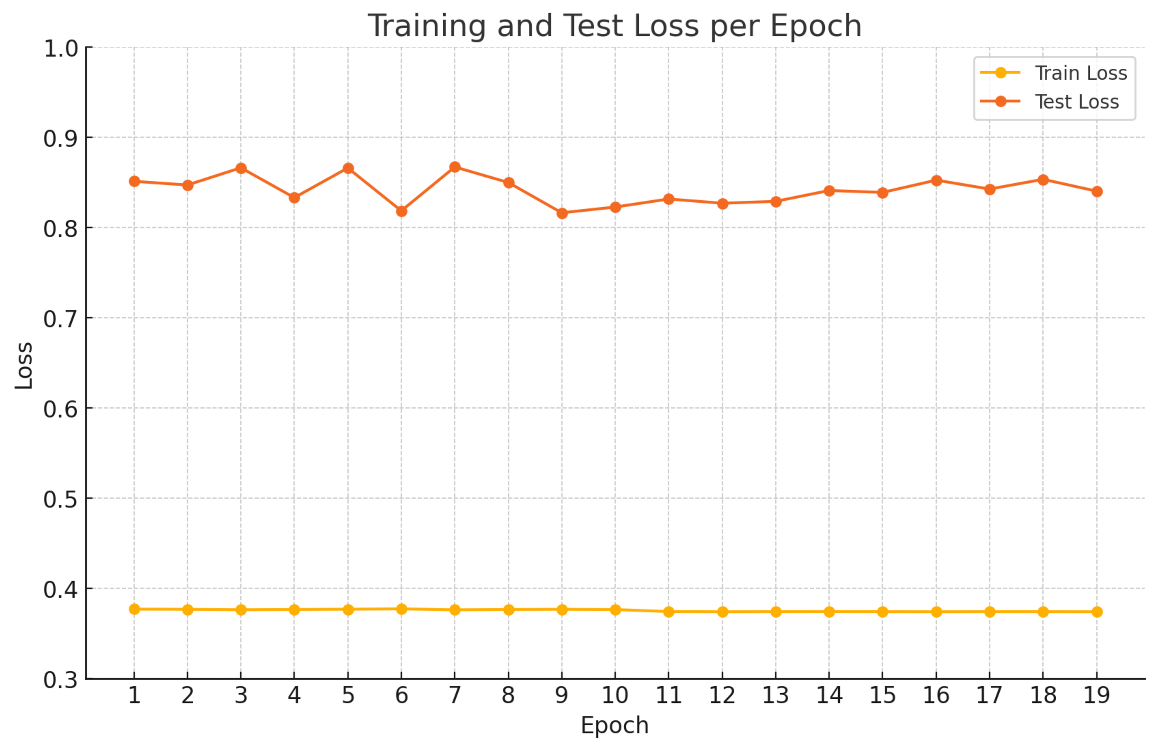
Isolation Tree Prophet



LOF Elliptic Envelope

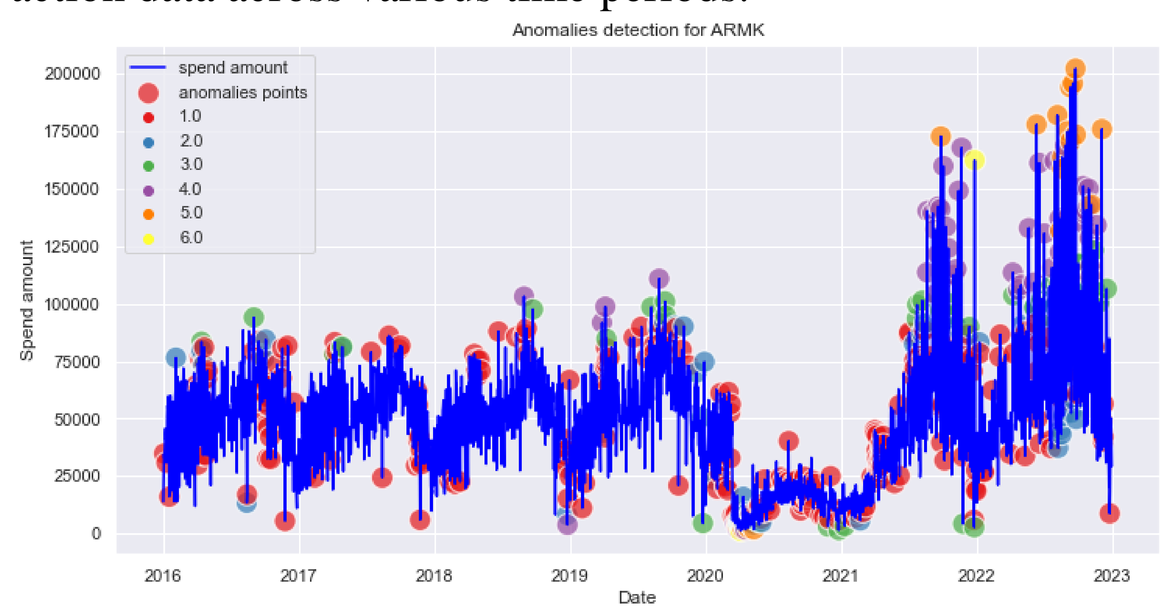
Phase 3: Refined Deep Learning Model

We built upon the framework used by the previous group, specifically their use of Knowledge Distillation model. We focused on refining the deep learning model through extensive parameter tuning to improve model stability and performance. After experimenting with various hyperparameters, such as learning rates, batch sizes, and optimizers, we determined that the optimal setup for our model includes a learning rate of 0.01 with a scheduler, 200 epochs with early stopping, a batch size of 64, and the Adam optimizer.



Conclusion

Through the integration of regime detection, voting-based anomaly detection, and deep learning refinement, our approach effectively identifies anomalies within financial transaction data across various time periods.



The sample results demonstrate that our methodology is capable of adapting to different economic scenarios, providing valuable insights into financial patterns and enabling more accurate anomaly detection. Future work could explore further model enhancements, such as including additional data sources and experimenting with other machine learning algorithms to improve performance.