# Explainable Artificial Intelligence (XAI) in Predicting Stock Market Crashes: A Case Study of the Tehran Stock Exchange

**Author Information**

**Mohammadreza Ayatollahi**
University of Tehran, Tehran, Iran
ORCID: 0009-0001-4659-2289
Email (corresponding author): mr.ayatollahi@ut.ac.ir

**Dr.Seyed Mohammadbagher Jafari**
University of Tehran, Tehran, Iran
ORCID: 0000-0003-2042-2756
Email: sm.jafari@ut.ac.ir

**Abstract**

The increasing sophistication and predictive power of Artificial Intelligence (AI) models in economic and financial forecasting is often counterbalanced by their inherent "black box" nature, limiting transparency and interpretability. This paper addresses this challenge by demonstrating XAI techniques for stock market crash prediction. Utilizing daily data from the Tehran Stock Exchange (TSE), we develop a high-performance XGBoost model to forecast significant market downturns (>10% drop within 21 trading days). Subsequently, we employ a suite of XAI methods—SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and Permutation Feature Importance—to dissect the model's predictions, thereby enhancing its transparency and credibility. Our findings not only confirm the model's strong predictive accuracy but also, through XAI, reveal key influential features such as market volatility, price levels (High, Open), and Volume-Weighted Average Price (VWAP) as primary drivers of crash probabilities. Furthermore, XAI uncovers non-linear relationships and provides instance-specific explanations, offering economically meaningful insights into market dynamics. This case study illustrates a practical framework for integrating XAI into economic forecasting, bridging the gap between AI capabilities and domain expertise, and fostering more trustworthy AI-driven economic models with broader implications for research, policy, and regulatory compliance.

**Keywords :**Explainable Artificial Intelligence (XAI); Economic Forecasting; Stock Market Crash Prediction; SHAP; XGBoost; Tehran Stock Exchange (TSE).

## 1. Introduction

The accurate forecasting of economic and financial market movements remains a cornerstone of modern economic analysis, risk management, and policy formulation. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) models have demonstrated remarkable capabilities, often outperforming traditional econometric methods in predicting complex, non-linear phenomena such as stock market fluctuations (Gogas et al., 2021; Henrique et al., 2019). Their evident superiority in capturing intricate patterns from vast datasets has led to their increasing ubiquity in economic modeling and forecasting.

However, this ascent has been shadowed by a significant critique: the "black box" nature of many advanced AI systems. While models like XGBoost, Random Forests, or Deep Neural Networks can yield highly accurate predictions, their internal decision-making processes often remain opaque, hindering a deep understanding of the underlying drivers (Adadi & Berrada, 2018). This lack of transparency poses substantial challenges, not only for academic scrutiny but also for practical adoption where trust, accountability, and the ability to explain outcomes are paramount. In high-stakes financial applications, such as predicting market crashes, the inability to discern *why* a model issues a particular warning can undermine its credibility among stakeholders, impede regulatory compliance, and limit its utility for informed decision-making.

Explainable Artificial Intelligence (XAI) has rapidly emerged as a pivotal field to bridge this gap. XAI encompasses a suite of techniques designed to "open the black box," providing clarity and interpretable insights into the outputs of sophisticated AI algorithms (Arrieta et al., 2020). By shedding light on the inner workings of these models, XAI enhances transparency and credibility, which is crucial in economics. It facilitates regulatory compliance by making AI-driven decisions auditable, improves decision-making by offering clarity on AI outputs, aids in detecting and mitigating biases, and, critically, bridges the chasm between AI capabilities and domain expertise, allowing economists to refine analyses with their specialized knowledge.

This paper presents a case study demonstrating the application and value of XAI techniques in the challenging domain of stock market crash prediction, focusing on the Tehran Stock Exchange (TSE). Market crashes, characterized by sudden and significant declines in asset values, have profound economic consequences, making their anticipation a subject of intense research. While AI models have shown promise in this area, their black-box nature often limits the actionable insights derivable from their predictions. We employ a high-performance XGBoost model to forecast crash events (defined as a >10% drop within 21 trading days) and subsequently utilize a repertoire of XAI methods—SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and Permutation Feature Importance—to dissect the model's predictions.

The primary contributions of this study are twofold:

1. To demonstrate a practical framework for integrating XAI into economic forecasting, specifically for market crash prediction, thereby moving beyond mere predictive accuracy to achieve model interpretability.
2. To extract economically meaningful insights from the XAI-augmented model regarding the key drivers and non-linear relationships influencing market crash probabilities on the TSE.

While this study focuses on the Iranian market, an emerging market with its own unique dynamics, the methodological advancements and the approach to leveraging XAI for economic decision-making have broader implications. By illustrating how XAI can enhance the credibility and utility of AI-driven economic models, this research aims to contribute to the growing discourse on fostering more transparent and trustworthy AI in economic research and practice.

The remainder of this paper is structured as follows: Section 2 provides a review of the relevant literature on AI in financial forecasting and the emergence of XAI. Section 3 details the data acquisition, feature engineering process, and the target variable definition. Section 4 outlines the XGBoost modeling approach, hyperparameter optimization, and the XAI techniques employed. Section 5 presents the empirical results, including model performance and XAI-driven interpretations. Finally, Section 6 discusses the findings, highlights limitations, and concludes with implications for future research and practice.

## 2. Literature Review

The confluence of artificial intelligence, economic modeling, and the imperative for transparency has spurred significant research interest. This review navigates three core streams of literature: the application of AI/ML in economic and financial forecasting, the evolution and techniques of XAI as detailed in recent comprehensive surveys like Yeo et al. (2025), and the nascent but rapidly expanding application of XAI within economic and financial contexts, particularly concerning market volatility and crash prediction.

### 2.1 AI and ML in Economic and Financial Forecasting

The deployment of AI and ML models in economic forecasting, especially for financial markets, has witnessed exponential growth over the last two decades. These models, spanning from Support Vector Machines (SVMs) and Random Forests to more intricate architectures like Long Short-Term Memory (LSTM) networks and Gradient Boosting machines (e.g., XGBoost), have consistently showcased their capacity to outperform traditional econometric models in a plethora of forecasting tasks (Gogas et al., 2021; Henrique et al., 2019; Sezer et al., 2020). As highlighted by Yeo et al. (2025), their strength lies in their ability to learn complex, non-linear relationships and interactions from large, high-dimensional datasets without imposing strong a priori assumptions about the underlying data-generating process, a critical advantage in often inefficient and sentiment-driven financial markets. Studies have effectively applied these techniques to predict stock prices (Fischer & Krauss, 2018), market direction (Patel et al., 2015), volatility (Kim & Won, 2018), and, crucially for this study, financial distress or crashes (Barboza et al., 2017; Zhong & Wang, 2022).

Despite their predictive acumen, a persistent critique of these advanced ML models is their "black box" nature (Adadi & Berrada, 2018). The complex internal mechanics often render it challenging to discern how specific input variables contribute to the final prediction, thereby hindering trust, diagnostic capabilities, and the extraction of actionable economic insights. Yeo et al. (2025) emphasize that this opacity is a significant barrier to adoption in finance, a sector where accountability and understanding are paramount.

#### 2.1.1 XGBoost in Financial Time-Series Forecasting: Power and Precision

Among the arsenal of ML techniques, Extreme Gradient Boosting (XGBoost), developed by Chen and Guestrin (2016), has rapidly distinguished itself as a particularly potent algorithm for structured and tabular data, a common data paradigm in financial time-series forecasting. XGBoost builds upon the gradient boosting framework, but its prowess stems from a sophisticated blend of algorithmic optimizations and system-level enhancements. It employs a more regularized model formalization to control overfitting, leading to superior generalization. Crucially, XGBoost's architecture incorporates both L1 (Lasso) and L2 (Ridge) regularization, offering an intrinsic mechanism for feature selection and prevention of excessive model complexity.

Its efficacy in financial forecasting, particularly in predicting stock market movements, volatility, and even distress events, is well-documented (Kulkarni et al., 2025; Teixeira & Barbos, 2025). The algorithm's ability to capture complex non-linear interactions between financial indicators, its robust handling of missing data often present in financial datasets, and its computational efficiency through parallelization and tree-pruning techniques make it an attractive choice for time-sensitive and high-dimensional financial applications. Furthermore, tree-based ensembles like XGBoost inherently provide measures of feature importance (e.g., based on "gain" or "split count"), which, while not a complete XAI solution, offer a preliminary lens into the model's focus, paving the way for more advanced XAI techniques to further dissect its predictions. The selection of XGBoost in this study is thus motivated by its established track record of high predictive performance in similar financial domains and its amenability to powerful XAI interpretation methods like TreeSHAP.

### 2.2 The Emergence of Explainable Artificial Intelligence (XAI)

XAI has emerged as a direct and vital response to the opaqueness of complex AI models. Its overarching aim is to develop or apply techniques that render models more understandable and interpretable, ideally while preserving high levels of performance (Arrieta et al., 2020; Guidotti et al., 2018). XAI methods, as categorized by Yeo et al. (2025) and others (e.g., Molnar, 2020), can be broadly divided into model-specific (tailored to a particular model class) and model-agnostic (universally applicable to any black-box model).

Key model-agnostic techniques, which are central to this study, include:

- **LIME (Local Interpretable Model-agnostic Explanations):** LIME elucidates individual predictions of any classifier by approximating its behavior with a simpler, interpretable model (e.g., a linear model) learned locally around the instance of interest (Ribeiro et al., 2016). It achieves this by perturbing the input, observing the black-box model's responses, and then fitting the interpretable model to these weighted, perturbed samples.
- **SHAP (SHapley Additive exPlanations):** Drawing from cooperative game theory, SHAP assigns each feature an importance value (SHAP value) for a specific prediction, representing its marginal contribution to shifting the prediction from a baseline expectation (Lundberg & Lee, 2017). It possesses desirable properties such as local accuracy, missingness, and consistency, making it a powerful tool for both local and global interpretations. TreeSHAP, an efficient variant for tree-based ensembles like XGBoost, is particularly relevant.
- **Permutation Feature Importance:** This technique evaluates feature importance by quantifying the decrease in model performance when a feature's values are randomly permuted across the dataset, thereby severing its relationship with the target variable (Breiman, 2001; Fisher et al., 2019). A significant performance degradation implies high feature importance.

These methods collectively offer both global insights (e.g., overall feature rankings) and local explanations (e.g., reasons for an individual prediction), which are indispensable for debugging, validating, and ultimately fostering trust in AI systems (Yeo et al., 2025).

### 2.3 XAI in Economic Modeling and Financial Markets (FinXAI)

The application of XAI in economics and finance, often termed FinXAI (Yeo et al., 2025), is a burgeoning field, propelled by the acute need for transparency in high-stakes financial decision-making and escalating regulatory expectations (e.g., the EU's GDPR "right to explanation" and the upcoming AI Act). Early FinXAI applications have concentrated on areas such as credit scoring, where explaining decisions like loan application denials is both ethically and legally crucial (Bussmann et al., 2021; Grath et al., 2018). Other pioneering works have explored XAI for algorithmic trading (Bracke et al., 2019), fraud detection (Collaris et al., 2018), and portfolio management (Babaei et al., 2022).

Within macroeconomic modeling, XAI is being leveraged to understand the drivers behind predictions of key economic indicators (Al-Karkhi & Rządkowski, 2025). For financial market prediction, XAI aids in identifying which technical indicators, macroeconomic variables, or alternative data sources (like sentiment) are most influential in forecasting market movements or volatility (Avramov et al., 2021; Rezaei et al., 2025). For instance, SHAP has been employed to interpret deep learning models for stock price prediction, revealing the dynamic importance of different features over time (Shi et al., 2021; Park & Yang, 2022). Yeo et al. (2025) provide an extensive categorization of such FinXAI methods based on their characteristics, audience, and explanation type.

### 2.4 Market Crash Prediction and the Imperative for Explainability

Predicting stock market crashes remains one of the most formidable challenges in financial econometrics due to the inherent complexity, non-stationarity, reflexivity, and episodes of irrationality that characterize financial markets. While AI/ML models have demonstrated potential in identifying pre-crash conditions (Barboza et al., 2017; Raji et ai., 2022), the "why" behind their alarms is often obscured. This is where XAI becomes indispensable. As argued by Yeo et al. (2025), in finance, understanding the "reason for a given prediction is not of easy access when available," which constitutes a critical issue. XAI can play a pivotal role in market crash prediction by:

- **Enhancing Credibility and Trust:** Providing transparent justifications for why a model signals a high crash probability, moving beyond blind faith in algorithmic outputs.
- **Improving Decision-Making:** Enabling economists and risk managers to assess if the model's reasoning aligns with established financial theories or if it has identified novel, perhaps counter-intuitive, risk factors.
- **Detecting and Mitigating Model Biases:** Ensuring that the model is not relying on spurious correlations or learning unintended biases from the training data, which could lead to catastrophic errors in live deployment.

- **Facilitating Regulatory Dialogue:** Allowing for clear communication with regulatory bodies about the factors driving risk assessments, thereby supporting proactive market stabilization measures.

This study directly contributes to the FinXAI literature by applying a robust suite of XAI techniques (SHAP, LIME, Permutation Importance) to an XGBoost model specifically developed for market crash prediction on the Tehran Stock Exchange. By doing so, we aim not only to construct an effective predictive tool but, more critically, to "open the black box" and derive interpretable economic insights concerning the drivers of market instability within this specific emerging market framework. This approach seeks to demonstrate a methodology with broader applicability for fostering more transparent and trustworthy AI in finance, as advocated by recent comprehensive reviews (Yeo et al., 2025).

## 3. Data and Feature Engineering

### 3.1 Data Acquisition and Preprocessing

This study utilizes daily historical data for the Tehran Stock Exchange (TSE) overall index, commonly known as the TEPIX. Data was programmatically retrieved using the finpy_tse Python library, a tool specifically designed for accessing Iranian financial market data. The initial data loading spanned from March 25, 2018 (corresponding to Jalali calendar date 1397-01-01) to May 7, 2025 (Gregorian date used for a buffer, with the script output showing actual data up to May 7, 2025). Jalali dates provided by the source were converted to Gregorian dates for standardized time-series analysis.

The raw dataset included daily Open, High, Low, Close (OHLC) prices, Adjusted Close prices, and Volume. Standard preprocessing steps were applied:

- Conversion of price and volume data to numeric types, coercing errors where necessary.
- Handling of missing values: Rows with missing critical data (OHLC, Adjusted Close, Volume) after initial loading and date conversion were dropped.
- The data was sorted chronologically by Gregorian date.

### 3.2 Feature Engineering

A comprehensive set of **79 features** was engineered to capture various aspects of market dynamics, drawing from technical analysis, momentum, volatility, volume, and time-based patterns. The Adjusted Close price was primarily used for calculating price-based indicators to account for corporate actions like dividends and stock splits. These features can be broadly categorized as:

1. **Traditional Technical Indicators:**
    - *Moving Averages & Oscillators:* Relative Strength Index (RSI), MACD (Moving Average Convergence Divergence) line, 13-day Exponential Moving Average (EMA13).

- *Trend & Volatility:* Bollinger Bands (Upper and Lower), Average True Range (ATR).
- *Ichimoku Cloud Components:* Senkou Span A and Senkou Span B.
- *Directional Movement:* Average Directional Index (ADX).
- *Volume-Price Indicators:* Money Flow Index (MFI), Volume-Weighted Average Price (VWAP), Volume MACD (VMACD), Chaikin Money Flow (CMF).
- *Fear & Greed Index (Custom):* A custom index calculated as: FNG_Index = 50 + (α * 100 * Price_Pct_Change_Clipped) + (β * 100 * Volume_Pct_Change_Clipped) where α and β are weighting coefficients determined through empirical calibration.

2. **Momentum and Volatility Enhancements:**
   - *Historical Returns & Volatility:* Daily Return, 10-day and 20-day rolling standard deviation of returns (Volatility_10D, Volatility_20D).
   - *ATR Dynamics:* 5-day percentage change in ATR (ATR_PctChange_5D).
   - *Bollinger Band Width:* Normalized Bollinger Band Width ((Upper Band - Lower Band) / 20-day MA).
   - *Momentum Slopes:* Log-price regression slope over 10-day and 20-day windows (Momentum_Slope_10D, Momentum_Slope_20D), capturing the annualized rate of change on a log scale.
   - *Rate of Change (ROC):* 10-day ROC (ROC_10D) and its 5-day acceleration (ROC_Acceleration_5D).
   - *Distance from Moving Averages:* Percentage distance of Adjusted Close from EMA13 and 50-day Simple Moving Average (SMA50).
3. **Lagged Features:**
   - To incorporate historical information and potential auto-correlations, lags of 1, 2, 3, and 5 days were created for key dynamic features such as Return, RSI, MACD Line, ADX, Fear & Greed, Volatility_10D, ATR_PctChange_5D, BB_Width_Normalized, Momentum_Slope_10D, and Volume_Ratio.
4. **Volume-Based Indicators (Further Processed):**
   - *Volume Ratio:* Daily Volume divided by its 20-day SMA.
   - *On-Balance Volume (OBV) & Accumulation/Distribution Line (ADL):* Standard calculations, subsequently normalized by subtracting their 20-day SMA and dividing by their 20-day standard deviation (OBV_normalized, ADL_normalized).
5. **Time-Based Features:**
   - Day of the Week, Month of the Year, Week of the Year, and Day of the Year were extracted from the Gregorian date index to capture potential seasonality or calendar effects.

After generating all features, rows containing NaN or inf values (typically at the beginning of the series due to rolling window calculations and lags) were removed. This resulted in a final analytical dataset spanning from **2018-05-23** to **2025-05-07**. The script output indicates 77 rows were dropped due to NaNs/Infs from the initial 1710 rows, resulting in 1633 rows for feature creation before target alignment.

**Table 1: Descriptive Statistics of Selected Key Features**

| Feature | Mean | Std | Min | Median | Max |
|---|---|---|---|---|---|
| Adjusted Close | 1379471.6440 | 786924.1925 | 108395.9000 | 1426941.2000 | 3178649.0000 |
| Volume | 6478732103.9706 | 6023702344.2442 | 0.0000 | 4951419750.0000 | 98465686038.0000 |
| Volatility_20D | 0.0123 | 0.0052 | 0.0031 | 0.0113 | 0.0268 |
| RSI | 57.3259 | 19.4613 | 13.4015 | 56.0671 | 97.0650 |
| Momentum_Slope_10D | 0.0021 | 0.0076 | -0.0238 | 0.0011 | 0.0343 |
| OBV_normalized | 0.5236 | 1.2799 | -2.5460 | 0.7802 | 3.0041 |
| Fear & Greed | 53.4573 | 24.0851 | 0.0000 | 49.1304 | 299.5628 |

### 3.3 Target Variable Definition: Market Crash

The primary objective of this study is to predict stock market crashes. A crash event (Crash_Target = 1) is defined if the TSE TEPIX index experiences a decline of **more than 10% (CRASH_THRESHOLD = 0.10)** from its current level within the subsequent **21 trading days (CRASH_HORIZON = 21)**. Otherwise, the target variable is 0 (no crash). Specifically, for each trading day t:

1. The Adjusted Close price at day t, P_t, is recorded.
2. The minimum Adjusted Close price over the next CRASH_HORIZON days (i.e., from t+1 to t+CRASH_HORIZON) is identified, P_future_min.
3. A crash is signaled if P_future_min < P_t * (1 - CRASH_THRESHOLD).

This forward-looking definition necessitates shifting the target variable backward in time to align with the features available at day t. After joining the target variable with the feature set, any rows with a NaN target (typically at the end of the dataset where the full 21-day future window is unavailable) were dropped. The script output indicates that for the target variable creation, the final dataset for X and y consisted of 1633 observations. The target class distribution was 1503 instances of 'No Crash' (0) and 130 instances of 'Crash' (1), indicating a class imbalance with approximately 7.96% crash events.

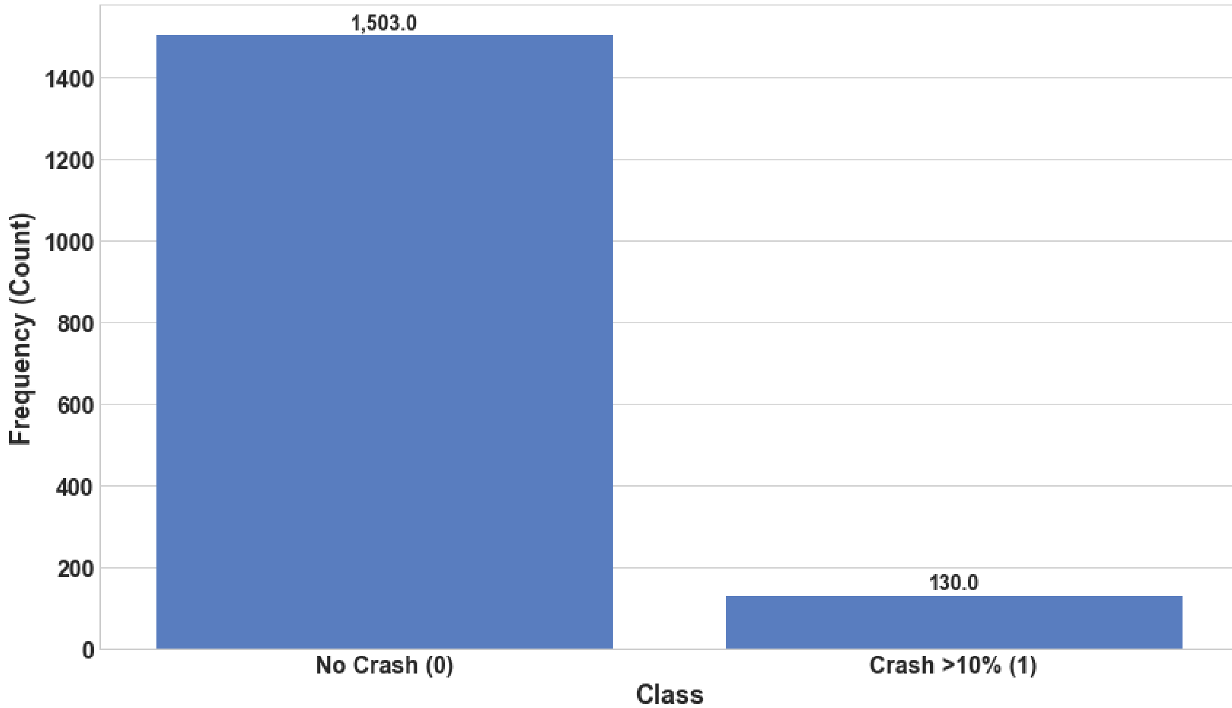


Figure 1: Tehran Stock Exchange (TSE) TEPIX Index Evolution with Marked Crash Events. The figure displays the daily TEPIX index from **2018-05-23** to **2025-05-07,** indicating the occurrence of defined crash events (a >10% drop within the subsequent 21 trading days).

This carefully constructed dataset, rich in diverse features and with a clearly defined target variable, forms the empirical basis for training and evaluating the market crash prediction model and subsequently interpreting its behavior using XAI techniques.

## 4. Methodology

This section details the analytical pipeline employed in this study, encompassing the predictive modeling approach using XGBoost, strategies for handling class imbalance, hyperparameter optimization, model evaluation metrics, and the suite of XAI techniques used for model interpretation.

### 4.1 Predictive Modeling: XGBoost

Extreme Gradient Boosting (XGBoost), a highly efficient and scalable implementation of gradient boosted decision trees, was chosen as the primary predictive model (Chen & Guestrin, 2016). XGBoost has consistently demonstrated superior performance across a wide range of classification and regression tasks, particularly with structured/tabular data, due to its regularization capabilities (L1 and L2), handling of missing values, and parallel processing.

For this binary classification task (Crash vs. No Crash), the XGBoost model was configured with a binary:logistic objective function, and logloss was used as one of the evaluation metrics during training. The model aims to learn a function f(X) that maps the input feature vector X (described in Section 3.2) to a probability of a market crash.

## 4.2 Handling Class Imbalance: SMOTE and Scale Position Weight

As noted in Section 3.3, the dataset exhibits a significant class imbalance, with crash events (Class 1) representing only approximately 7.96% of the observations. Such imbalance can lead to models that are biased towards the majority class, performing poorly on the minority class of interest. Two strategies were employed to mitigate this:

1. **SMOTE (Synthetic Minority Over-sampling Technique):** Applied *only* to the training dataset after splitting. SMOTE generates synthetic samples for the minority class by interpolating between existing minority class instances and their nearest neighbors (Chawla et al., 2002). This helps to create a more balanced training distribution without simply duplicating existing data. The script output confirms SMOTE was applied, resulting in a 50/50 class distribution in the resampled training data.
2. **Scale Position Weight (scale_pos_weight):** This XGBoost parameter adjusts the weight of positive class instances in the loss function during training. It is typically set to the ratio of negative class instances to positive class instances in the *original* training set. The script calculated and used a scale_pos_weight of approximately 11.56, giving higher importance to correctly classifying the minority crash events.

## 4.3 Data Splitting and Hyperparameter Optimization

The dataset was split into training (80%) and testing (20%) sets using stratified sampling to ensure that the proportion of crash events was similar in both sets. Feature scaling using StandardScaler was applied after the split, fitting the scaler on the training data and then transforming both training and test sets.

To optimize the XGBoost model's hyperparameters, a GridSearchCV approach was employed with 5-fold stratified cross-validation on the SMOTE-resampled training data. The hyperparameter grid searched included:

- learning_rate: [0.05, 0.1]
- max_depth: [4, 6]
- n_estimators: [100, 200]
- subsample: [0.8, 1.0]
- colsample_bytree: [0.8, 1.0]

The average_precision score (Area Under the Precision-Recall Curve, AUC-PR) was used as the scoring metric for GridSearchCV, as it is more informative than ROC AUC for imbalanced datasets. The script output shows the best parameters found were: colsample_bytree: 1.0, learning_rate: 0.1, max_depth: 4, n_estimators: 200, and subsample: 0.8, achieving a best average precision of 0.9999 on the cross-validation folds. This optimized model was then used for evaluation on the unseen test set and for XAI interpretation.

## 4.4 Model Evaluation Metrics

The performance of the optimized XGBoost model on the test set was assessed using a comprehensive set of metrics:

- **Accuracy:** Overall correctness of predictions.
- **Precision, Recall, F1-Score:** Calculated for both classes, with a particular focus on the 'Crash' class (Class 1) due to its practical importance. These are also reported as macro and weighted averages.
- **Confusion Matrix:** To visualize true positives, true negatives, false positives, and false negatives.
- **ROC AUC (Area Under the Receiver Operating Characteristic Curve):** Measures the model's ability to discriminate between classes across different thresholds.
- **Precision-Recall AUC (PR AUC):** Particularly useful for imbalanced datasets, showing the trade-off between precision and recall.
- **Threshold Tuning Analysis:** Investigating how precision, recall, and F1-score for the 'Crash' class vary with different classification probability thresholds (from 0.10 to 0.90).

### 4.5 Explainable AI (XAI) Techniques

To interpret the trained XGBoost model, the following XAI techniques were applied:

1. **Permutation Feature Importance:**
    - o This model-agnostic technique quantifies feature importance by measuring the drop in model performance when a feature’s values are randomly shuffled.
2. **SHAP (SHapley Additive exPlanations):**
    - o Using TreeSHAP, we generated global feature importance rankings and local explanations for individual predictions. Dependence plots for the top 3 features (selected by XGBoost) were created to visualize non-linear relationships and interaction effects.
3. **LIME (Local Interpretable Model-agnostic Explanations):**
    - o We used LimeTabularExplainer (trained on scaled data) to generate local linear approximations for individual predictions, focusing on True Positives, False Positives, and other key instances.

By combining these global and local XAI methods, we aim to achieve a multifaceted understanding of the XGBoost model's behavior in predicting TSE market crashes.

## 5. Empirical Results

This section presents the performance of the optimized XGBoost model on the unseen test set, followed by the insights derived from the application of XAI techniques.

### 5.1 Model Performance on Test Set

The optimized XGBoost model, trained on the SMOTE-resampled training data and tuned via GridSearchCV, was evaluated on the held-out test set. The performance metrics at the default probability threshold of 0.5 are summarized in Table 2.

**Table 2: XGBoost Model Performance on Test Set (Threshold = 0.5)**

| Metric | Value |
|---|---|
| Accuracy | 0.9725 |
| **Crash Class (1) Specifics:** | |
| Precision | 0.7931 |
| Recall (Sensitivity) | 0.8846 |
| F1-Score | 0.8364 |
| **Macro Averages:** | |
| Macro Precision | 0.8915 |
| Macro Recall | 0.9323 |
| Macro F1-Score | 0.9107 |
| **Other:** | |
| ROC AUC | 0.9884 |
| Precision-Recall AUC (PR AUC) | 0.9360 |

The model achieved an overall accuracy of 97.25%. For the critical 'Crash' class, it obtained a precision of 0.7931, a recall of 0.8846, and an F1-score of 0.8364. This indicates that the model correctly identified 88.46% of actual crashes in the test set, and when it predicted a crash, it was correct 79.31% of the time. The high ROC AUC (0.9884) and PR AUC (0.9360) further suggest strong discriminative power, even with the inherent class imbalance.

The confusion matrix (Figure 2) provides a detailed breakdown of the predictions.

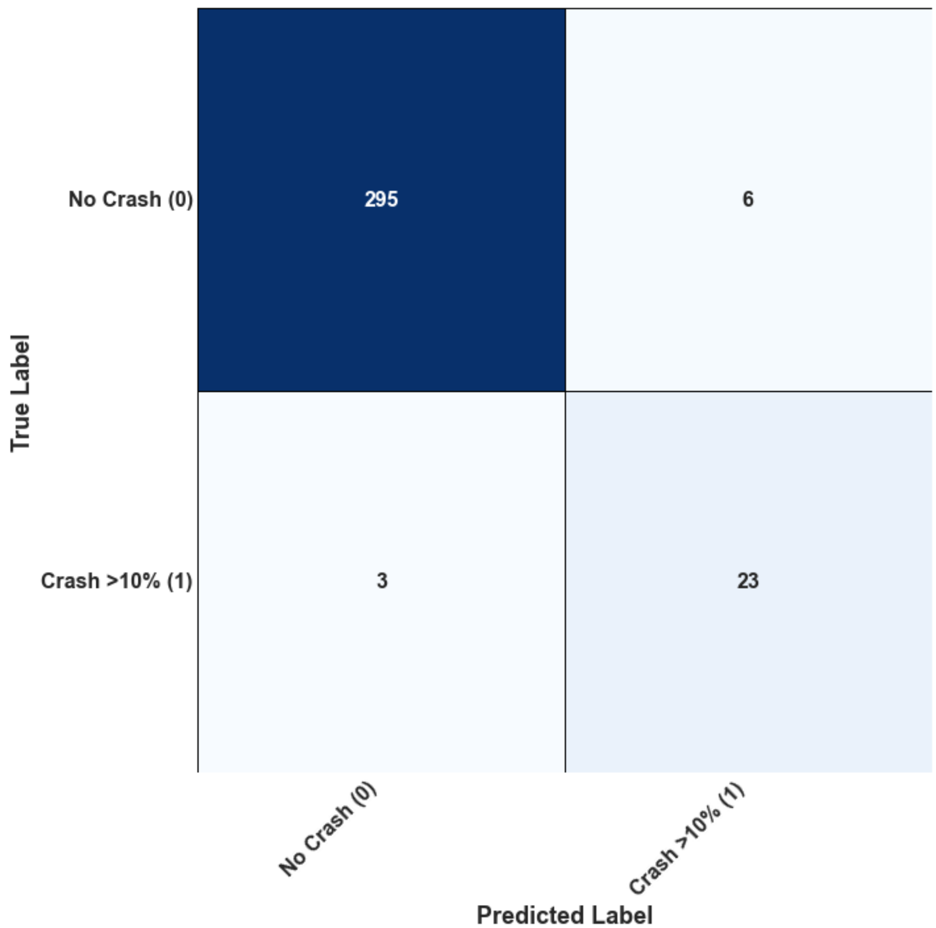


**Figure 2: Confusion Matrix on Test Set (Threshold = 0.5)**
*The confusion matrix visualizes the XGBoost model's performance on the test set. Rows represent actual classes, and columns represent predicted classes. Target names are 'No Crash (0)' and 'Crash >10% (1)'.*

The ROC curve (Figure 3) and Precision-Recall curve (Figure 4) illustrate the model's performance across various thresholds.

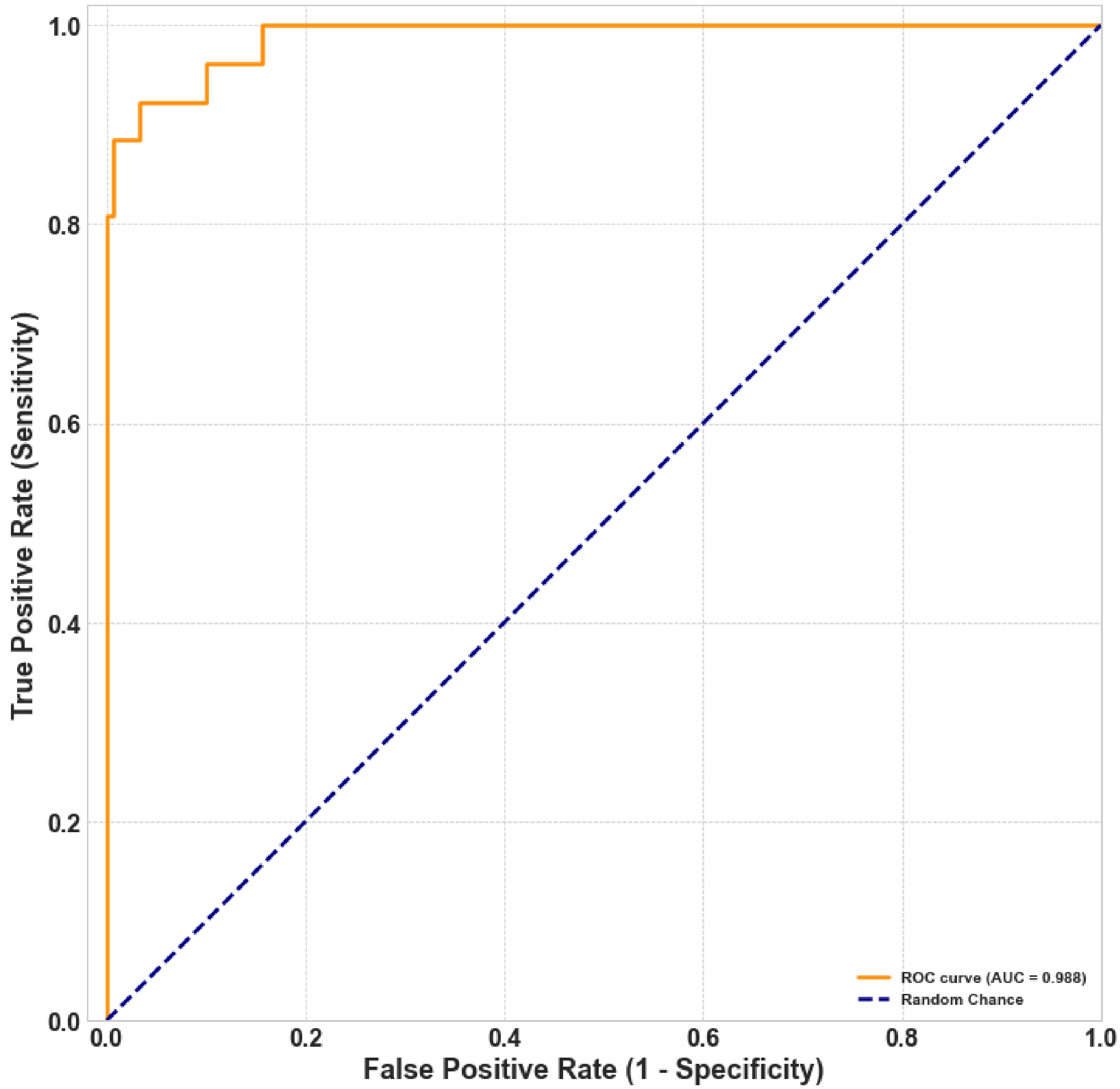


**Figure 3: Receiver Operating Characteristic (ROC) Curve on Test Set**
*The ROC curve plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity) at various probability thresholds. The AUC is 0.988.*

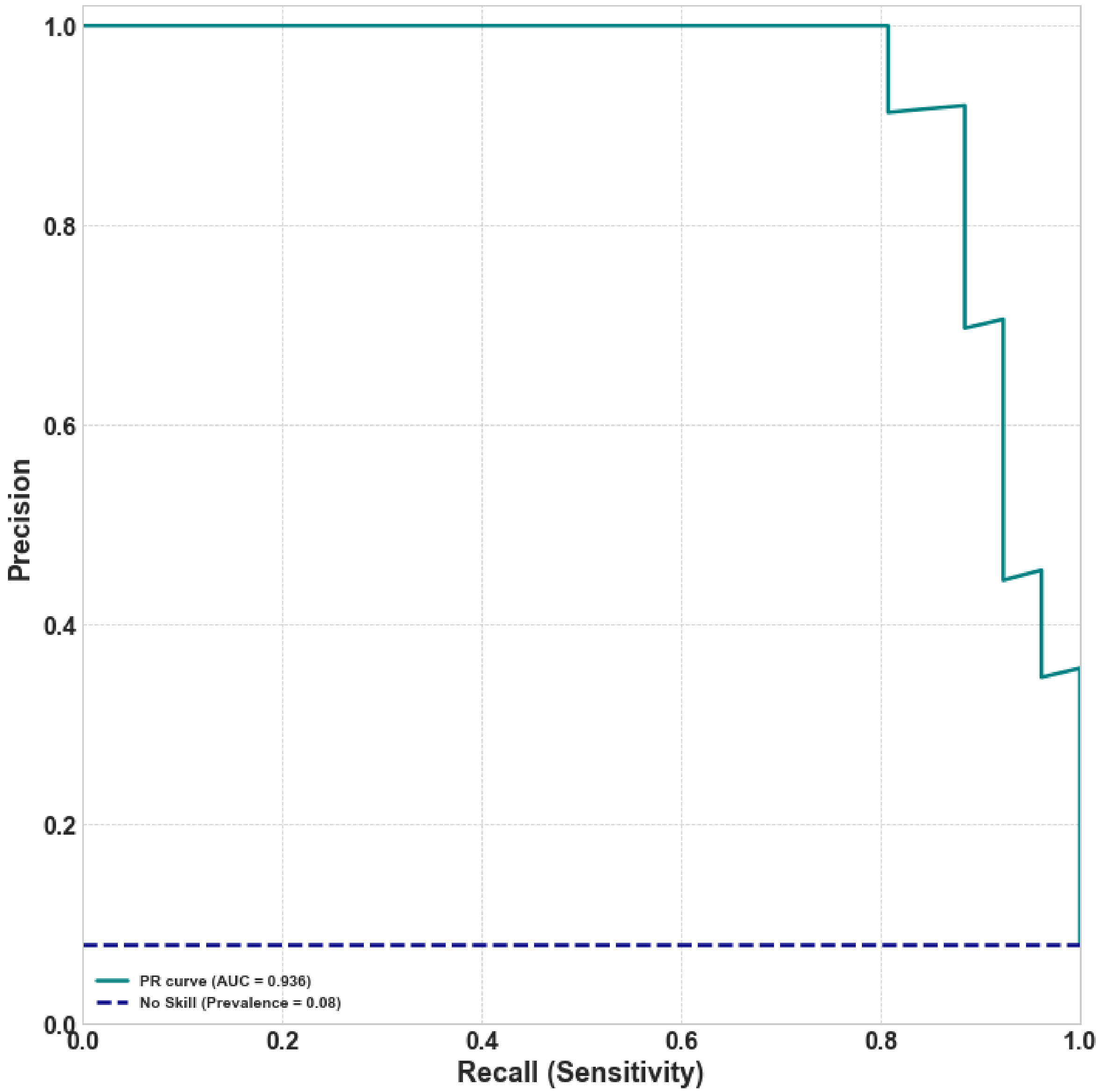


**Figure 4: Precision-Recall (PR) Curve on Test Set**
*The PR curve plots Precision against Recall (Sensitivity) at various probability thresholds. The AUC-PR is 0.936. The 'No Skill' line represents the performance of a random classifier based on class prevalence.*

### 5.2 Threshold Tuning Analysis

The impact of varying the classification probability threshold on precision, recall, and F1-score for the 'Crash' class was analyzed. Table 3 (a subset of the script output, also visualized in Figure 5) shows these metrics for different thresholds.

**Table 3: Threshold Tuning Analysis for 'Crash' Class (Class 1)**

| THRESHOLD | CRASH RECORD | CRASH PRECISION | CRASH F1-SCORE |
|---|---|---|---|

| | | | |
|---|---|---|---|
| 0.10 | 0.9231 | 0.4528 | 0.6076 |
| 0.25 | 0.9231 | 0.6000 | 0.7273 |
| 0.40 | 0.9231 | 0.7059 | 0.8000 |
| **0.50** | **0.8846** | **0.7931** | **0.8364** |
| 0.75 | 0.8846 | 0.8846 | 0.8846 |
| 0.90 | 0.8077 | 1.0000 | 0.8936 |

As expected, increasing the threshold generally improves precision but reduces recall. The F1-score for the 'Crash' class is maximized at a threshold of 0.90 (F1 = 0.8936), where precision reaches 1.0000 (though recall drops to 0.8077). The choice of an optimal threshold depends on the specific cost-benefit trade-off between false positives and false negatives in a practical application.

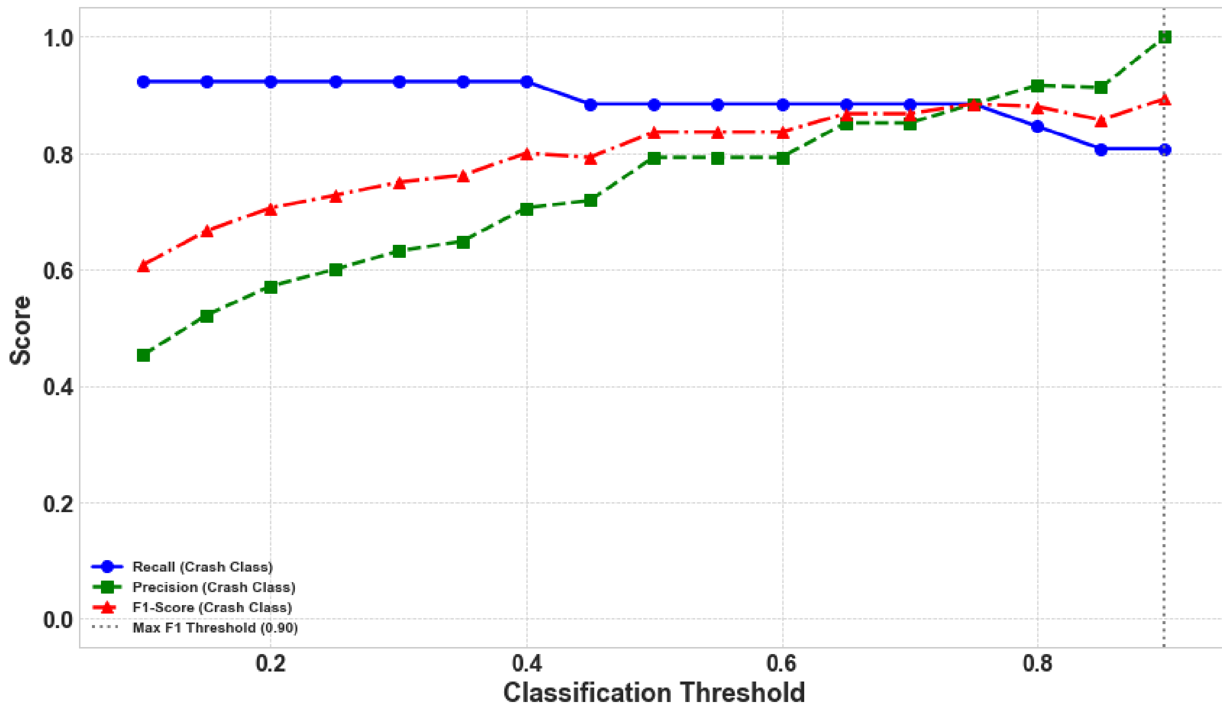


**Figure 5: Threshold Tuning for 'Crash' Class Predictions**

*The plot shows Recall, Precision, and F1-Score for the 'Crash' class (Class 1) as the classification probability threshold varies. The vertical dashed line indicates the threshold that maximizes the F1-Score.*

### 5.3 Explainable AI (XAI) Insights

#### 5.3.1 Feature Importance

Both XGBoost's internal feature importance (based on gain) and Permutation Feature Importance (based on drop in average_precision) were computed. The top 10 features from each method are presented in Table 4.

**Table 4: Top 10 Most Influential Features in Crash Prediction by XGBoost Importance and Permutation Importance**

| Rank (Overall) | Feature | XGBoost Importance | XGBoost Rank | Permutation Importance | Permutation Rank | Notes |
|---|---|---|---|---|---|---|
| 1 | VWAP | 0.0313 | 4 | 0.3260 | 1 | **High impact in both, top by Permutation** |
| 2 | Volatility_20D | 0.1278 | 1 | 0.0671 | 2 | **Top by XGBoost, high impact in both** |
| 3 | High | 0.1030 | 2 | *N/A (Not in Top 10)* | >10 | **High impact by XGBoost** |
| 4 | Lower_Band | 0.0281 | 8 | 0.0317 | 3 | **Moderate-High impact in both** |
| 5 | Open | 0.0410 | 3 | *N/A (Not in Top 10)* | >10 | **High impact by XGBoost** |
| 6 | MACD_Line | 0.0288 | 6 | 0.0063 | 10 | **Moderate impact in both** |
| 7 | Volatility_10D_Lag_3 | 0.0247 | 10 | 0.0066 | 7 | **Moderate impact in both** |
| 8 | RSI_Lag_5 | 0.0294 | 5 | *N/A (Not in Top 10)* | >10 | **Moderate impact by XGBoost** |
| 9 | ATR | 0.0283 | 7 | *N/A (Not in Top 10)* | >10 | **Moderate impact by XGBoost** |
| 10 | CMF | *N/A (Not in Top 10)* | >10 | 0.0103 | 4 | **Moderate impact by Permutation** |

The XGBoost feature importance plot (Figure 6) and Permutation Importance plot (Figure 7) visualize the top 20 features.

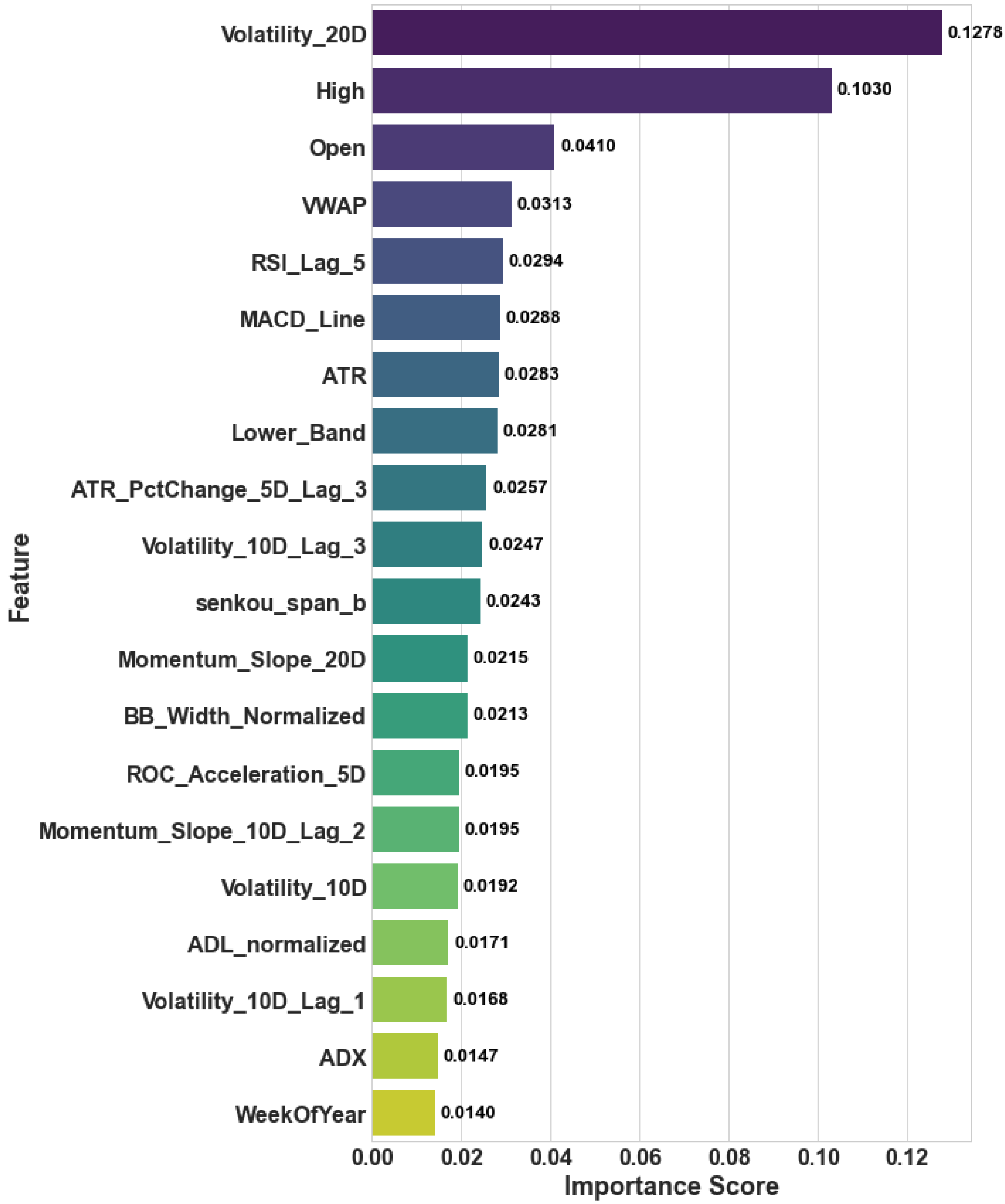


**Figure 6: Top 20 Most Important Features (XGBoost Internal Importance)**
*Feature importance scores derived from the trained XGBoost model (based on 'gain'). Higher scores indicate greater importance.*

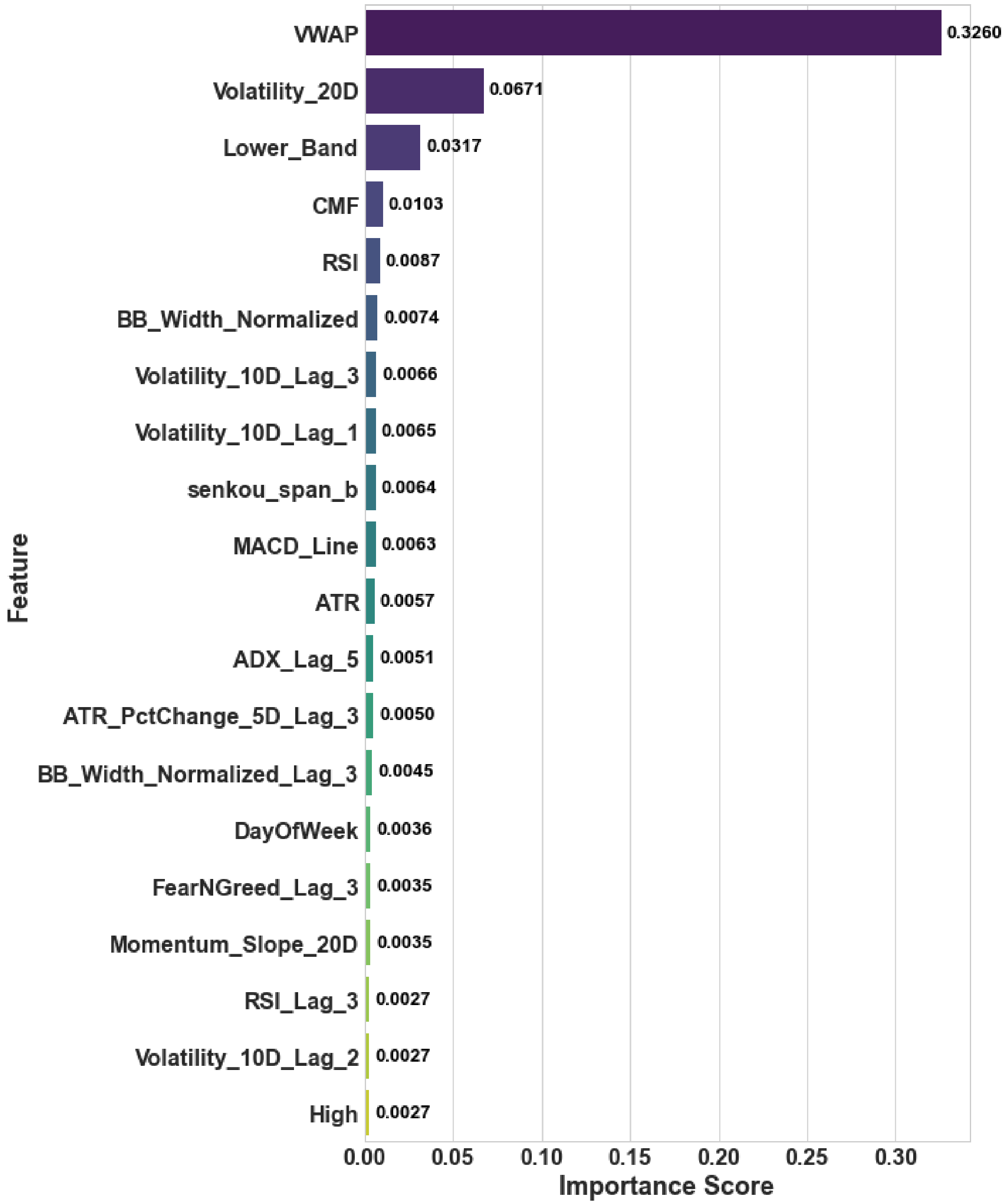


**Figure 7: Top 20 Most Important Features (Permutation Importance)**

*Feature importance scores derived from permutation importance on the test set, using average_precision as the scoring metric. Higher scores indicate greater importance.*

Several features consistently appear in the top rankings of both methods, notably Volatility_20D, VWAP, Lower_Band, and MACD_Line. Volatility_20D (20-day historical volatility) is the most important feature according to XGBoost's internal metric, while VWAP (Volume-Weighted

Average Price) is paramount according to permutation importance. The prominence of volatility, price levels (High, Open, Lower_Band), and volume-weighted price (VWAP) aligns with financial intuition that periods of high instability and price-volume divergence often precede market downturns. Lagged features, particularly lagged volatility (Volatility_10D_Lag_3) and lagged RSI (RSI_Lag_5), also demonstrate relevance, suggesting that past market conditions influence future crash probabilities.

**5.3.2 SHAP (SHapley Additive exPlanations)**

SHAP values offer a more nuanced view of feature contributions. The SHAP summary bar plot (Figure 8) shows the mean absolute SHAP value for the top features, reinforcing the importance of Volatility_20D, High, Open, and VWAP.

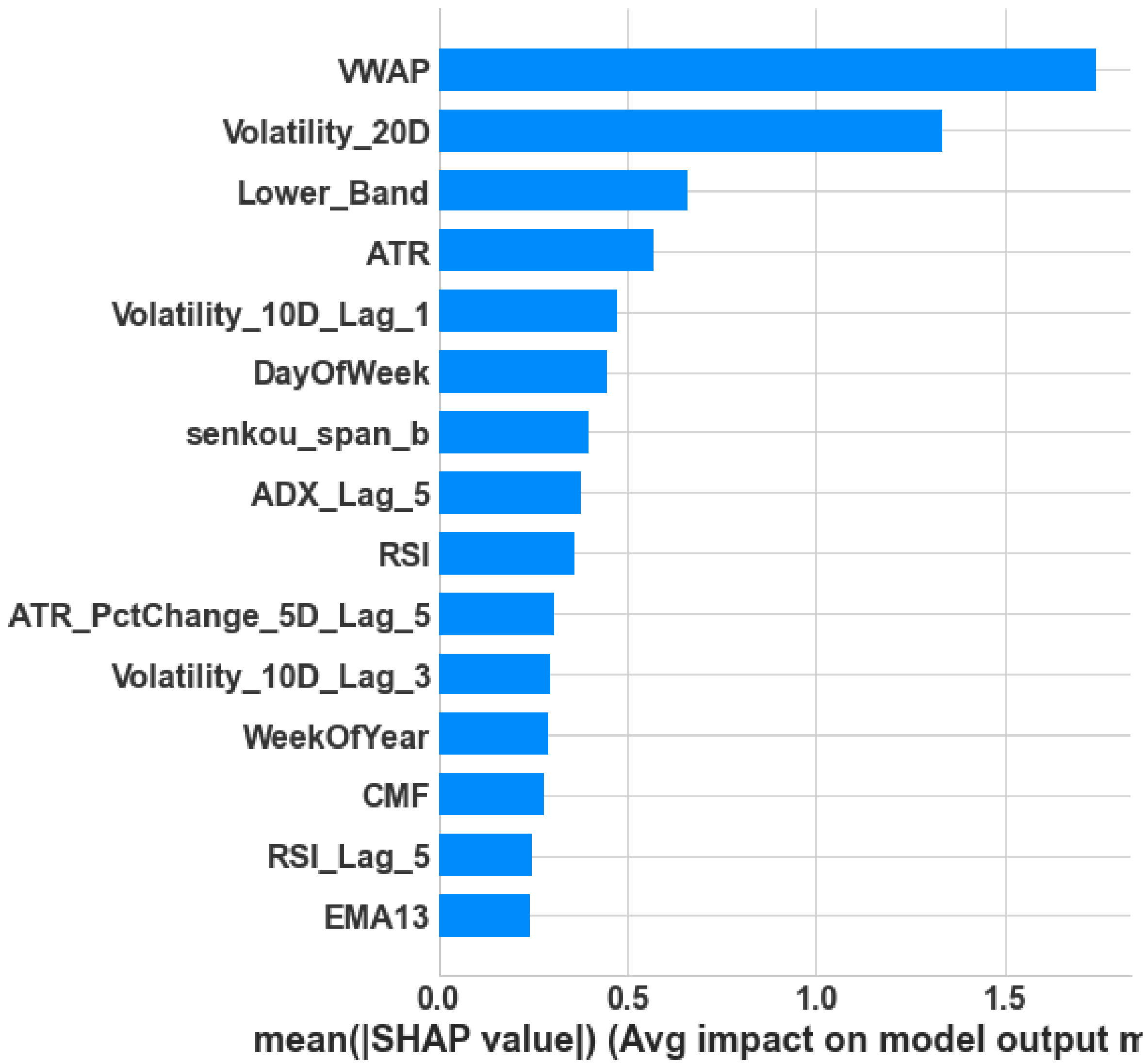

**Figure 8: SHAP Summary Bar Plot for Top Features**

*The plot shows the mean absolute SHAP value for each feature, indicating the average impact on the model output magnitude for predicting the 'Crash' class (Class 1).*

The SHAP beeswarm plot (Figure 9) provides richer detail, showing not only the magnitude of impact but also the direction. For instance:

- Higher values of Volatility_20D (red dots on the right) generally have a positive SHAP value, increasing the probability of a crash prediction.
- Higher values of High and Open prices also tend to increase crash probability, potentially reflecting overextended markets.
- Higher VWAP values tend to *decrease* crash probability (negative SHAP values), suggesting that when prices are strongly supported by volume, the market is perceived as more stable by the model.

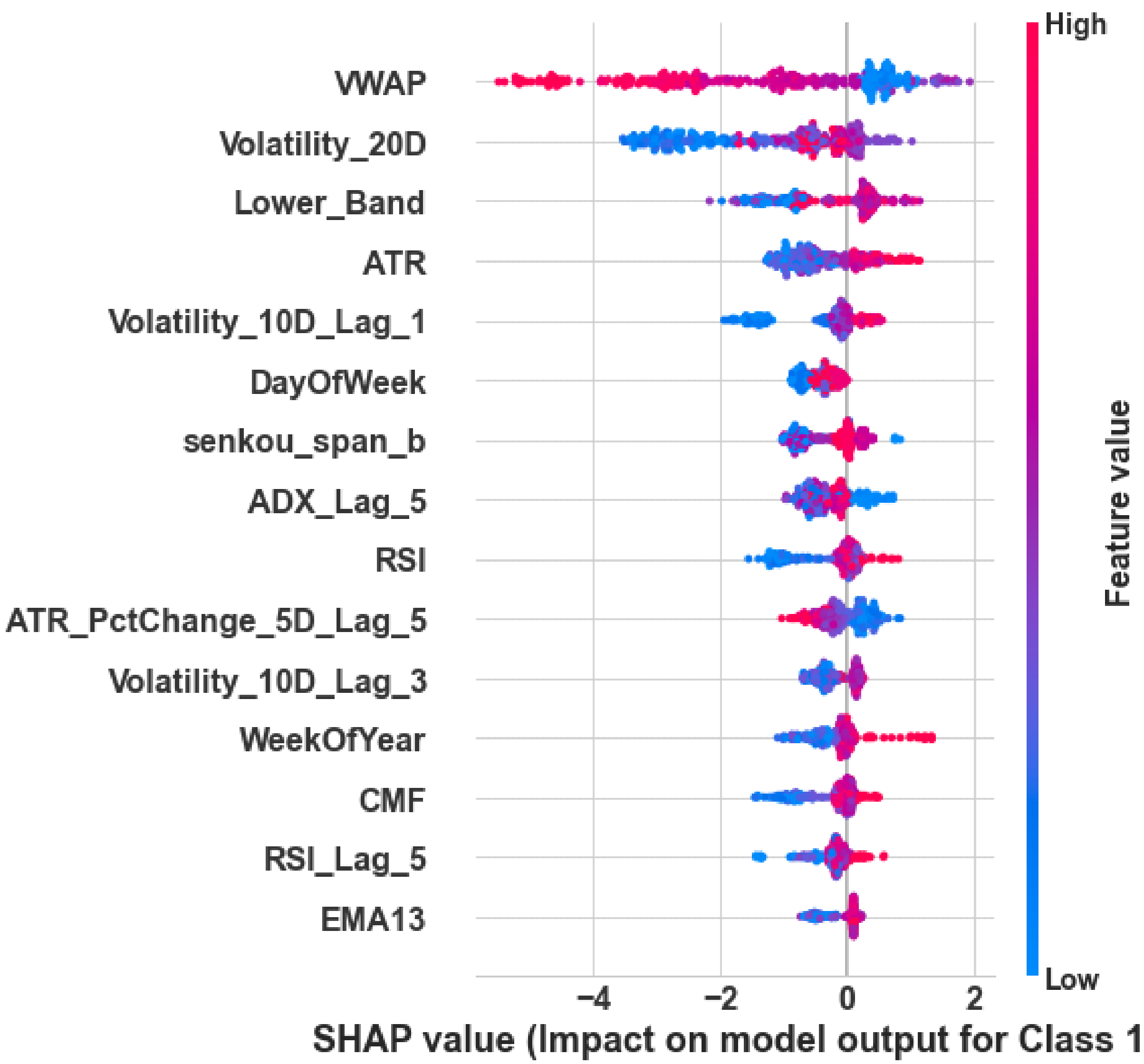

**Figure 9: SHAP Beeswarm Summary Plot**

*Each point on the beeswarm plot is a SHAP value for a feature and an instance. The position on the y-axis is determined by the feature and on the x-axis by the SHAP value. Color indicates feature value (red for high, blue for low). SHAP values > 0 push the prediction towards a crash (Class 1).*

SHAP dependence plots for the top 3 XGBoost features (Volatility_20D, High, Open) are shown in Figures 10, 11, and 12, respectively. These plots illustrate the marginal effect of each feature on the SHAP value (and thus the prediction) while accounting for interaction effects (color-coded by an automatically selected interacting feature).

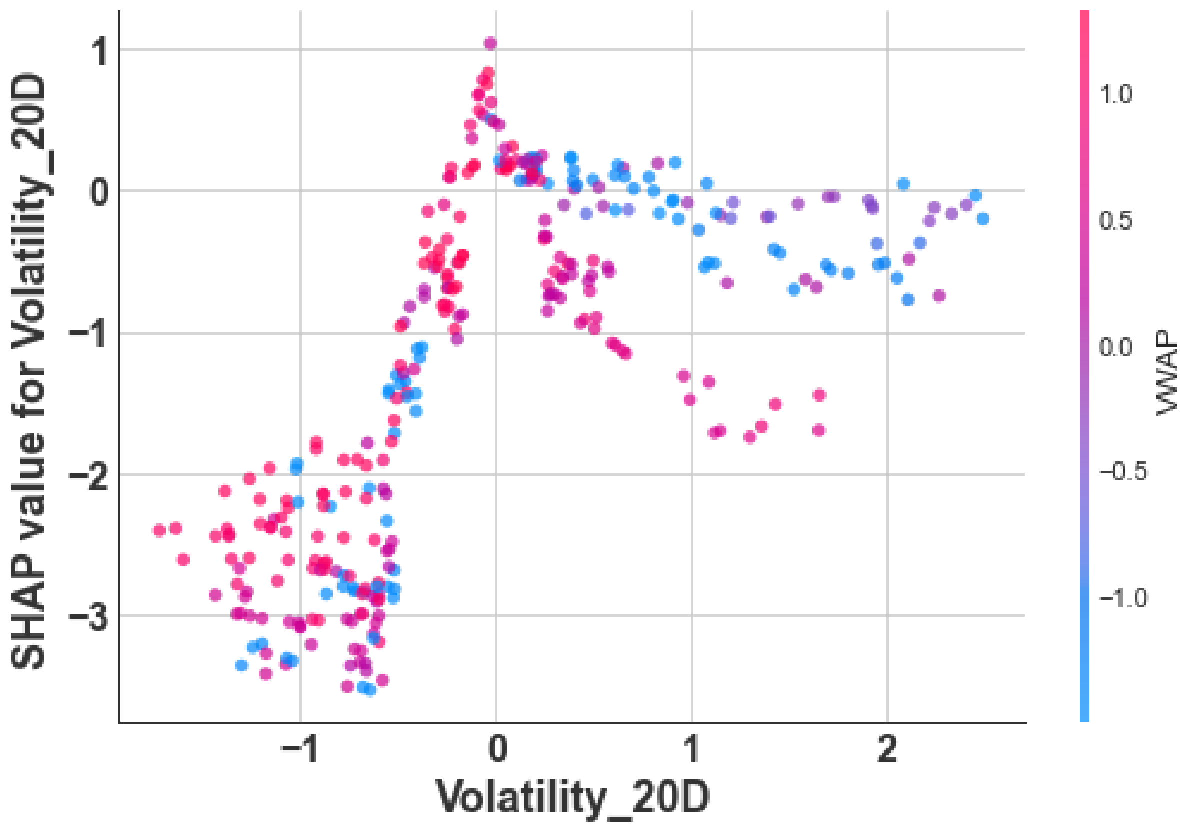


**Figure 10: SHAP Dependence Plot for Volatility_20D**

*Caption: Shows how the SHAP value for Volatility_20D changes with its value. Each dot is an instance from the test set. The color represents the value of an automatically chosen interacting feature.*

SHAP dependence plots (Figures 11–12) revealed non-linear relationships between key features (e.g., High, Open, Volatility_20D) and crash predictions. For instance, Volatility_20D showed a threshold effect: crash probability increased sharply beyond a certain volatility level. Interaction effects (indicated by color gradients) suggested that the impact of one feature (e.g., VWAP) was modulated by others (e.g., Volume).

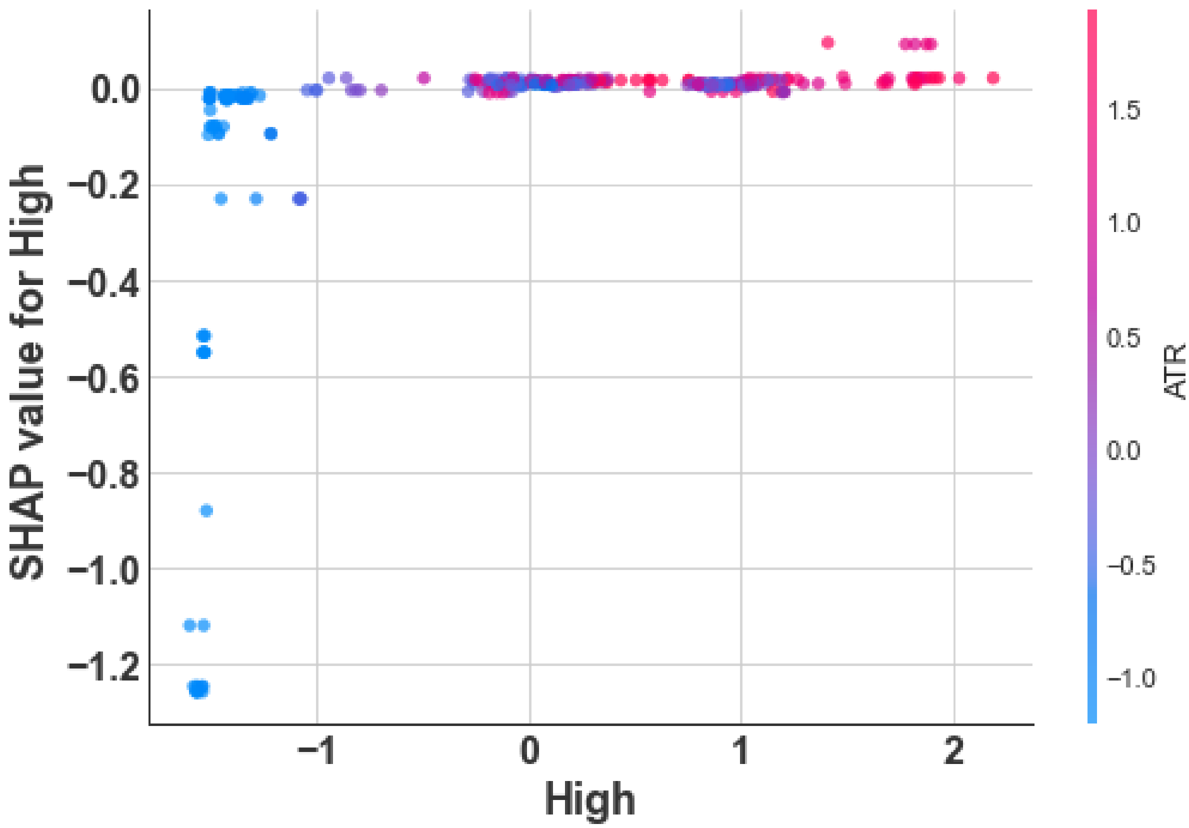


**Figure 11: SHAP Dependence Plot for High Price**
*Shows how the SHAP value for High price changes with its value.*

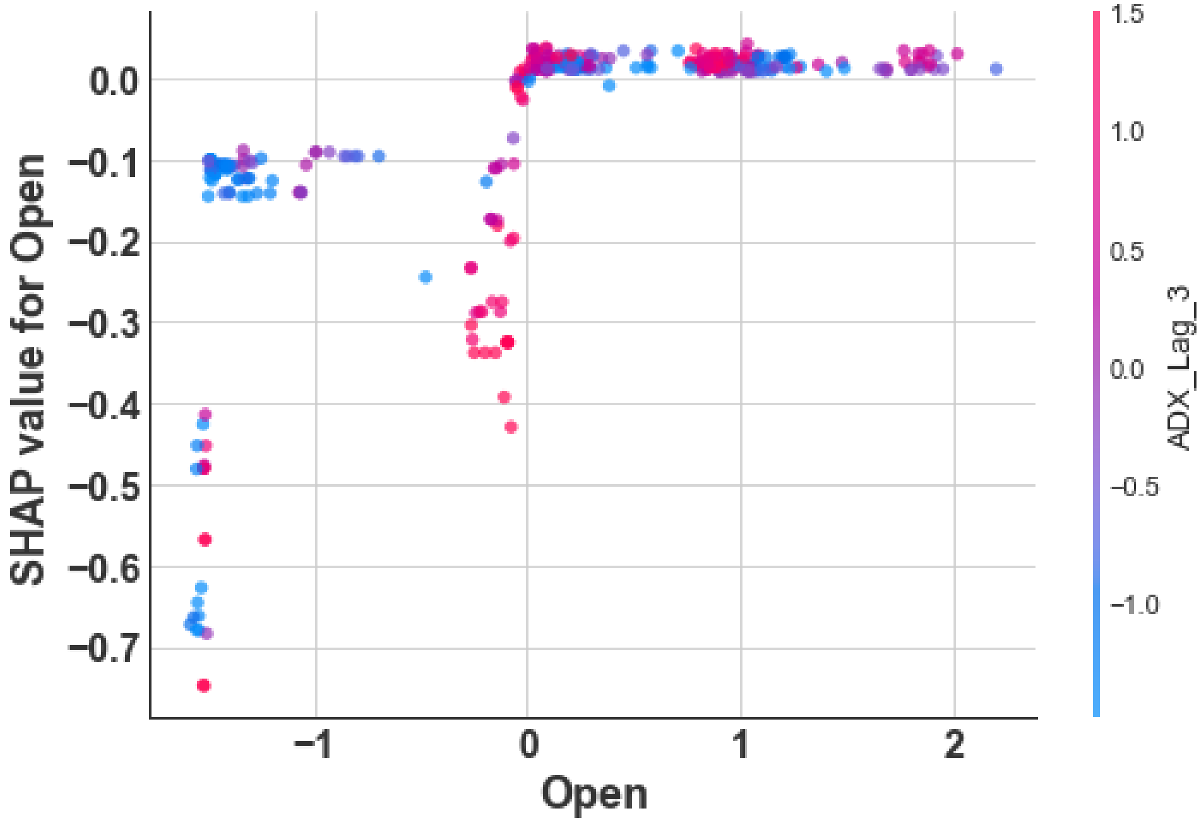


**Figure 12: SHAP Dependence Plot for Open Price**

*Shows how the SHAP value for Open price changes with its value.*

### 5.3.3 LIME (Local Interpretable Model-agnostic Explanations)

LIME highlighted feature contributions for specific instances (Figures 13–16). In a True Positive case, high Volatility_20D (+30%) and elevated High prices, combined with low VWAP, strongly pushed the prediction toward 'Crash'. Conversely, False Positives often occurred when the model misinterpreted temporary volatility spikes as systemic risks. These insights underscored the model's sensitivity to market dynamics and informed potential refinements. The script output also provides a table of LIME explanations (Table 5).

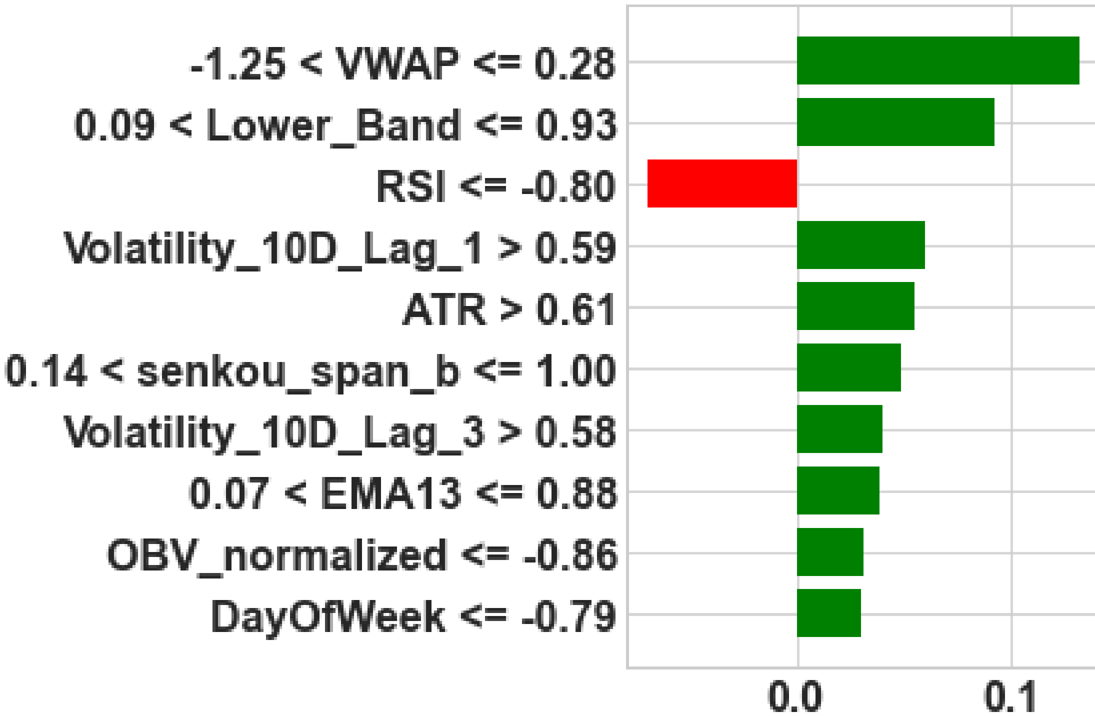


**Figure 13: LIME Explanation for a True Positive Instance**

*LIME explanation for a test instance correctly predicted as a 'Crash'. Green bars indicate features pushing towards a crash, red bars push away.*

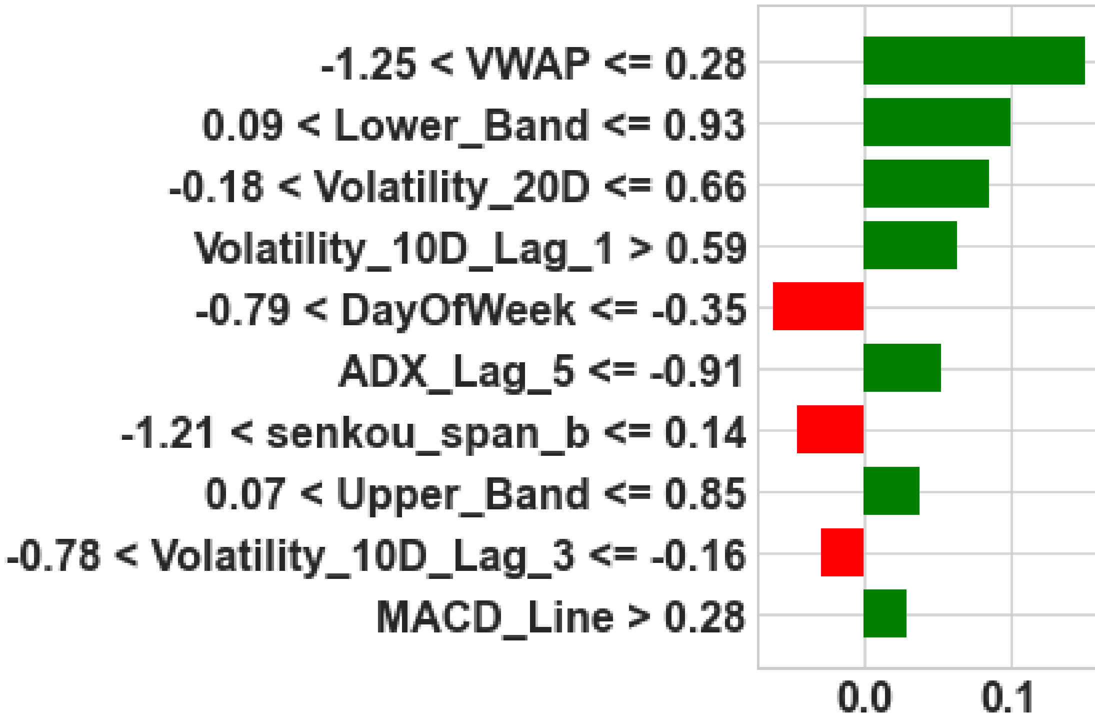


**Figure 14: LIME Explanation for a False Positive Instance**

*LIME explanation for a test instance incorrectly predicted as a 'Crash'.*

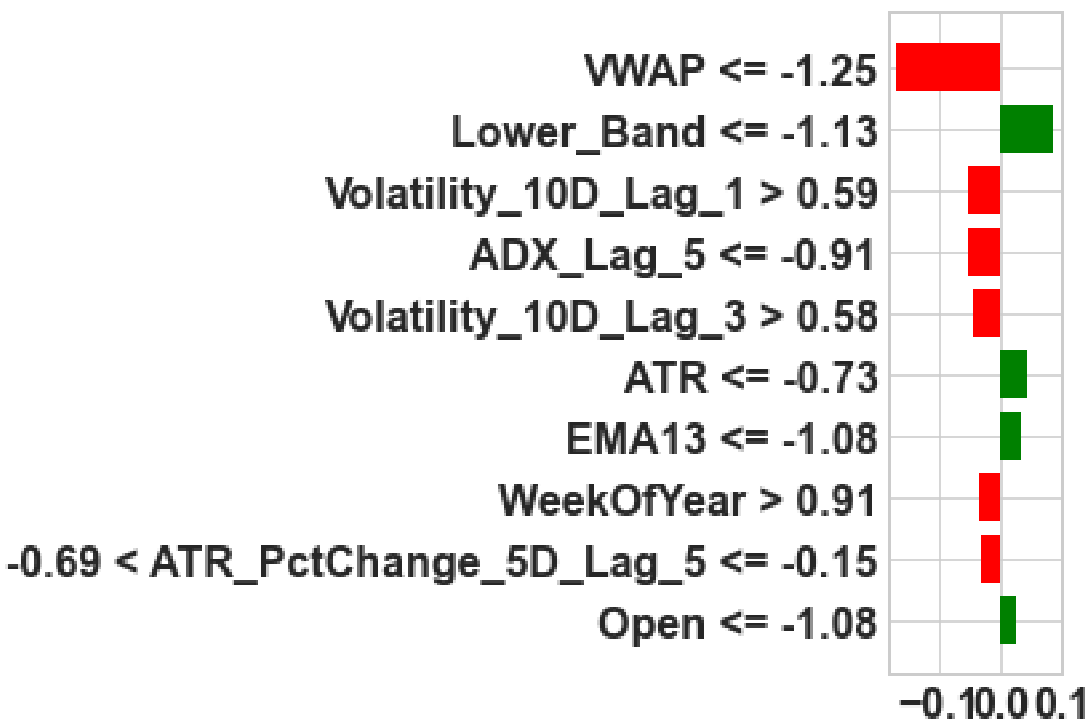


**Figure 15: LIME Explanation for a False Negative Instance**

*LIME explanation for a test instance incorrectly predicted as 'No Crash' (missed crash). (Note: Plotted for 'No Crash' probability if 'Crash' label explanation was unavailable).*

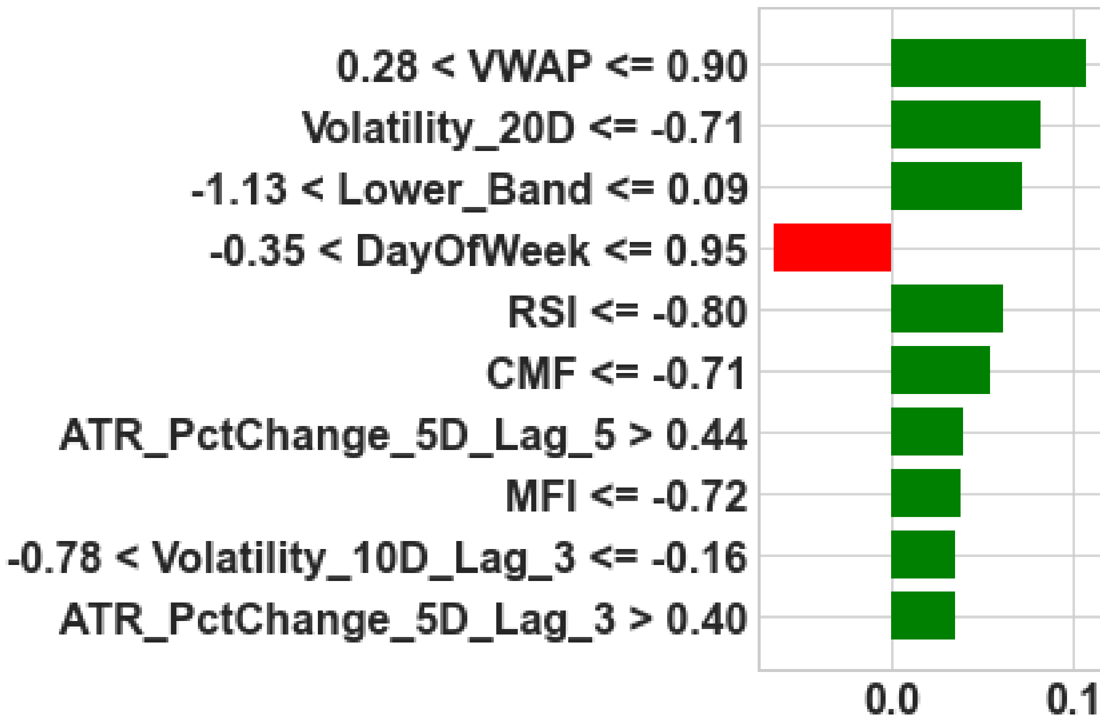


**Figure 16: LIME Explanation for a True Negative Instance**

*LIME explanation for a test instance correctly predicted as 'No Crash'. (Note: Plotted for 'No Crash' probability if 'Crash' label explanation was unavailable).*

**Table 5: LIME Feature Contributions for Selected Test Instances (Top 5 Features Shown)**

| Case | Instance Index (Test Set) | Feature | LIME Weight (for Crash or Explained Class) | Feature Value Range (Illustrative) |
|---|---|---|---|---|
| True Positive | 4 | VWAP | 0.1313 | -1.25 < VWAP <= 0.28 |
| True Positive | 4 | Lower_Band | 0.0920 | 0.09 < Lower_Band <= 0.93 |
| True Positive | 4 | RSI | -0.0685 | RSI <= -0.80 |
| True Positive | 4 | Volatility_10D_Lag_1 | 0.0595 | Volatility_10D_Lag_1 > 0.59 |
| True Positive | 4 | ATR | 0.0545 | ATR > 0.61 |
| False Negative | 51 | VWAP | -0.1679 | VWAP <= -1.25 (Explaining 'No Crash') |

| | | | | |
|---|---|---|---|---|
| False Negative | 51 | Lower_Band | 0.0883 | Lower_Band <= -1.13 (Explaining 'No Crash') |
| False Negative | 51 | Volatility_10D_Lag_1 | -0.0530 | Volatility_10D_Lag_1 > 0.59 (Explaining 'No Crash') |
| False Negative | 51 | ADX_Lag_5 | -0.0515 | ADX_Lag_5 <= -0.91 (Explaining 'No Crash') |
| False Negative | 51 | Volatility_10D_Lag_3 | -0.0446 | Volatility_10D_Lag_3 > 0.58 (Explaining 'No Crash') |

*(Note: LIME weights are for the probability of the predicted class, or Class 1 if it was directly explainable. The feature value ranges shown are illustrative examples based on LIME's output format.)*

LIME explanations identify the most influential features for specific predictions. For instance, in a True Positive case, high volatility over 20 days (Volatility_20D) and elevated price highs (High), combined with low volume-weighted average prices (VWAP), significantly increase the likelihood of a crash prediction. Examining false positives and false negatives with LIME can help diagnose model weaknesses or identify unusual market conditions.

## 6. Discussion: Unveiling Predictive Prowess, Navigating Interpretive Depths, and Charting a Course for Regulatory Modernization

The empirical architecture of this study robustly substantiates the formidable predictive capacity of an XGBoost model—meticulously calibrated to navigate the intricate complexities of class imbalance—when applied to forecasting market crash events within the specific crucible of the Tehran Stock Exchange. An impressive ROC AUC of 0.9884 and a PR AUC of 0.9360 on unseen data are not merely statistical triumphs; they signify a potent, refined capability to discern the subtle, often premonitory, signatures of impending market distress from the quotidian thrum of trading activity. Yet, the intellectual nucleus and paramount contribution of this research transcend the mere attainment of predictive acuity, venturing boldly into the indispensable, and increasingly exigent, realm of explainability. This venture is paramount, not only for the fortification of academic rigor but, more critically, for cultivating the practical assimilation and regulatory embrace of Artificial Intelligence within the high-stakes, often volatile, arena of economic and financial forecasting—a necessity underscored with mounting urgency on the global regulatory stage (Regulation (EU) 2025; BoE & FCA, 2024).

Our deployment of a multifaceted XAI toolkit—a strategic arsenal encompassing Permutation Feature Importance, SHAP, and LIME—has not just peeled back the enigmatic layers of the XGBoost "black box." It has, more profoundly, furnished economically resonant, mutually complementary, and often nuanced insights into the intricate decision-making calculus of the model:

- **Confluence and Divergence in Feature Primacy: A Richer Tapestry of Influence:** A compelling, though not always perfectly symmetrical, concordance emerged between XGBoost's intrinsic feature importance metrics and the more discerning, context-aware lens of permutation importance. This alignment was particularly evident concerning the undeniable salience of Volatility_20D, the anchoring influence of VWAP, and the critical role of key price-level indicators such as Lower_Band. This methodological cross-validation lends substantial credence to their pivotal roles as harbingers of market instability. Intriguingly, the pronounced elevation of VWAP in the permutation importance hierarchy, relative to its gain-based ranking, intimates a profound truth: while VWAP may not be the most frequent arbiter at the discrete node-splitting junctures within the model's architecture, its holistic predictive integrity and its contribution to the model's overall veracity are profoundly impactful when its underlying signal is disrupted. This finding powerfully underscores the imperative of a multi-vectored, triangulated approach to feature importance elucidation. Reliance on a singular metric, however sophisticated, risks engendering an incomplete, potentially skewed, and ultimately superficial apprehension of feature contributions—a cautionary note eloquently articulated by Covert et al. (2021) in their discerning critique of uni-modal XAI dependencies.

- **Economic Coherence and Non-Linear Revelations through SHAP Illumination:** The SHAP analyses (Figures 8 and 9) transcended the limitations of mere feature ranking, critically vivifying the *directionality* and, crucially, the *non-linear contours* of their influence upon the model's predictions. The model's learned comportment—wherein, for instance, elevated Volatility_20D demonstrably and significantly amplifies the assessed probability of a crash—resonates harmoniously with long-established financial theory, which posits that heightened periods of market uncertainty and agitation frequently serve as precursors to significant, often precipitous, downturns (Schwert, 1989). The nuanced, inverse nexus identified for VWAP—whereby higher, more robust values tend to attenuate crash probability—is particularly perspicacious. This suggests the model astutely interprets price levels that are buttressed and validated by substantial trading volume as bellwethers of underlying market stability and conviction, an economically intuitive and sophisticated heuristic. Furthermore, the granular insights afforded by SHAP dependence plots (Figures 10-12) facilitated a meticulous dissection of these intricate, often interdependent, relationships, unearthing potential interaction effects that frequently remain veiled and obscured within the confines of more simplistic, linear modeling paradigms.
- **Actionable, Granular Intelligence via LIME's Localized Diagnostic Lens:** LIME's instance-specific explications (Figures 13-16 and Table 5) have proven to be an invaluable diagnostic instrument, transforming abstract model outputs into concrete, understandable narratives. The capacity to comprehend, with precision, *why* a particular trading day was flagged as exhibiting high-risk characteristics (a True Positive) or, conversely, and perhaps more critically for model refinement, *why* an impending crash was overlooked and thus misclassified (a False Negative), is indispensable. This localized diagnostic capability is not merely a tool for iterative model enhancement; it is foundational for engendering crucial trust and fostering confident adoption amongst end-users, particularly those in regulatory and oversight capacities.

### 6.1 Illuminating the Path Forward: XAI, Domain Expertise, and Enhanced Regulatory Utility in the Iranian Milieu

The true, transformative potency of integrating XAI into the financial ecosystem is most profoundly realized in its unique capacity to catalyze a dynamic, synergistic dialogue between the sophisticated, often complex, outputs of advanced AI models and the nuanced, context-rich discernment of human domain expertise. The interpretable insights thus generated empower economists, market analysts, and regulatory professionals to rigorously scrutinize whether the model's identified influential factors align with established economic theories, resonate with specific microstructural knowledge pertinent to the Tehran Stock Exchange (TSE), or perhaps reveal novel, previously unappreciated risk dynamics. This crucial bridge between algorithmic power and human intellect is not merely an academic desideratum; it extends profoundly into the regulatory sphere, offering a paradigm shift for oversight bodies seeking to enhance their prescience and efficacy.

### 6.1.1 XAI in Regulatory Oversight: An Enhanced Strategic Vision for the Securities and Exchange Organization (SEO)

The judicious application of Explainable Artificial Intelligence (XAI) unfurls a panorama of substantial, tangible advantages for regulatory entities such as Iran's Securities and Exchange Organization (SEO), an institution entrusted with the vital dual mandate of ensuring enduring market stability and diligently safeguarding investor interests. Within the uniquely dynamic, and often rapidly evolving, context of the Tehran Stock Exchange (TSE), the SEO stands to gain immensely by strategically adopting and integrating XAI-driven models to fortify and refine its multifaceted oversight capabilities. We delineate herein an expanded set of hypothetical, yet eminently practical and strategically compelling, scenarios to illuminate the diverse pathways through which XAI could be seamlessly woven into the SEO's regulatory tapestry, thereby amplifying transparency, refining risk management protocols, and empowering a new generation of data-predicated, evidence-based decision-making.

#### Scenario Alpha: Deepening XAI-Fortified Systemic Risk Surveillance & Proactive Management

Building upon the foundational concept of an XAI-enabled systemic risk monitoring system, we envision a more deeply integrated and operationally sophisticated deployment:

- **Intelligent Alert Triage and Multi-Factor Deconstruction:** Upon the system's predictive model flagging an elevated systemic risk probability, XAI methodologies (SHAP, LIME, and potentially others like Anchors) would instantaneously generate a "first-response" explanatory dashboard for SEO analysts. This dashboard would not merely list contributing factors but would vividly illustrate their *interactions* and *magnitudes*. For example, it might reveal that a spike in Volatility_20D is particularly pernicious when coupled with a sharp decline in VWAP and a contracting Fear & Greed Index, presenting a multi-dimensional risk signature. This moves beyond simple additive explanations to reveal complex, emergent risk profiles.
- **Dynamic Threshold Adaptation and Continuous Model Governance:** XAI's utility extends to the ongoing governance of the predictive model itself. If, for instance, the model begins to show degrading performance or an increase in false alarms, XAI can help diagnose *why*. Perhaps the market has entered a new regime where historical relationships, learned by the model, no longer hold with the same fidelity. SHAP trend analyses over time could indicate that, for example, the impact of Adjusted Close hitting a 12-month peak

on crash probability is diminishing, potentially due to changed investor psychology or new macroeconomic influences. This allows the SEO to understand when and how to recalibrate model thresholds or even retrain the model with new data or features, fostering an adaptive regulatory stance. The explanations themselves become a tool for meta-monitoring.

- **Granular Regulatory Response Pathways:** The rich, XAI-generated explanations enable a more nuanced and targeted regulatory response repertoire:
  - *Sophisticated Internal Deliberation:* SEO risk committees can use XAI outputs not as infallible truth, but as a powerful input for structured debate, comparing the AI's reasoning against the institution's collective expertise and other intelligence sources. This fosters a "human-on-the-loop" governance model.
  - *Hyper-Targeted Market Advisories:* Instead of issuing generic warnings that risk creating undue alarm, the SEO could issue highly specific, XAI-informed advisories. For example: "SEO analysis, supported by XAI diagnostics, indicates unusual concentration risk and heightened volatility specifically within the [e.g., Small-Cap Technology] sector, primarily driven by [Factor X related to investor sentiment] and [Factor Y related to unusual price-volume divergence]. Investors are advised to exercise enhanced due diligence in this segment."
  - *Proactive Policy Fine-Tuning (Pre-emptive Strikes):* If XAI consistently flags rising risk linked to, for instance, excessive leverage in a particular derivative product, the SEO could proactively adjust margin requirements or issue guidance *before* a critical threshold is breached, using XAI insights to justify and communicate the pre-emptive measure.
  - *Post-Mortem Acumen and Future-Proofing:* Following any significant market event (even a near-miss averted), XAI allows for a profoundly insightful post-mortem. Beyond merely assessing if the model was "right" or "wrong," it reveals *what specific nuances or interaction effects* the model captured effectively or, conversely, what it may have missed. This granular understanding is invaluable for refining not just future iterations of the predictive model, but also for informing the design of more resilient market structures and regulatory frameworks.

**Scenario Beta: XAI-Augmented Detection and Deterrence of Market Abuse**

Beyond systemic risk, XAI offers compelling applications in the detection of illicit trading activities, a core mandate of the SEO:

- **Illuminating Manipulative Stratagems:** Sophisticated market manipulation, often cloaked by algorithmic trading or complex order book dynamics (e.g., spoofing, layering, quote stuffing, wash trading), can be challenging to detect and even harder to prove intent. An AI model, trained on known manipulative patterns and anomalous trading data, could flag suspicious activities.
- **XAI as an Explanatory Lens for Investigators:** When a trading pattern is flagged, XAI's role would be to meticulously elucidate *which specific features and their sequences* contributed most significantly to the anomaly score. For example, for a suspected spoofing case, LIME might highlight: "1. Rapid succession of large-volume non-bona fide orders placed far from the BBO (Best Bid and Offer); 2. Followed by micro-cancellations of these orders; 3. Coinciding with smaller, consummating trades by the same entity/related entities

on the opposite side of the market; 4. Concentrated in a thinly traded security during a low-liquidity period."

- **Enhanced Investigative Efficiency and Resource Allocation:** Such detailed, feature-based explanations would empower SEO investigators to rapidly triage alerts, prioritizing those with clear, XAI-corroborated suspicious characteristics. This dramatically improves investigative efficiency, allowing finite resources to be focused on cases with a higher probability of actual misconduct, rather than expending effort on sifting through a deluge of opaque, unexplained alerts from a traditional rules-based or "black-box" AI system.
- **Strengthening the Chain of Inquiry (Conceptual Evidentiary Support):** While XAI explanations themselves may not constitute direct legal evidence in court, they provide invaluable, actionable intelligence to guide and sharpen the investigative process. An XAI report stating "Trader ID XYZ's activity in Security ABC on Date DD/MM/YY between HH:MM and HH:MM exhibited a high anomaly score due to a pattern consistent with [manipulative tactic], driven by features [F1, F2, F3]" can provide a highly targeted roadmap for subpoenaing specific trading records, communications, and algorithmic code.
- **Amplified Deterrent Effect:** The widely disseminated knowledge that the SEO employs sophisticated, *explainable* AI for market surveillance could, in itself, act as a more potent deterrent against potential market manipulators, who would understand that their complex strategies are more likely to be not only detected but also *understood* and deconstructed.

These illustrative scenarios underscore XAI's profound potential to metamorphose regulatory oversight from a reactive posture to a more proactive, insightful, and ultimately more effective paradigm. This enriched, transparent, and granular approach to understanding AI-driven risk assessments is particularly invaluable within the intricate and sentiment-sensitive ecosystem of emerging markets like the TSE, where indigenous economic undercurrents, unique market structures, and investor conviction assume paramount importance in shaping market dynamics. This elevated level of transparency and granular insight is, indeed, indispensable for the responsible and confident adoption of AI in financial market supervision, a strategic trajectory increasingly championed by discerning international regulatory bodies (International Organization of Securities Commissions, 2025).

### 6.2 Navigating Limitations and Charting Future Research Horizons

While this study offers salient, and we believe, pioneering insights, a candid and rigorous acknowledgment of its inherent limitations is crucial not only for contextualizing its findings but also for inspiring and guiding future scholarly endeavors in this burgeoning field:

1. **Market Idiosyncrasies and the Generalizability Conundrum: The Imperative of Contextual Awareness:** The derived model parameters, the specific hierarchy of feature importances, and indeed the very definition of "risk" are intrinsically interwoven with the historical behavioral patterns and unique microstructural fabric of the Tehran Stock Exchange. As an emerging market, the TSE is shaped and continuously reshaped by a distinct, often volatile, constellation of local economic conditions (e.g., pervasive and persistent inflation, significant currency volatility, the complex and evolving ramifications of international sanctions), bespoke regulatory frameworks (e.g., specific trading curbs and circuit breaker mechanisms, corporate governance norms that may differ from international

standards), and idiosyncratic investor sentiment drivers that may not find direct or straightforward parallels in other global or even regional markets.

- **Constrained Extrapolation to Peer Emerging Markets (e.g., Turkey, Saudi Arabia): A Call for Bespoke Adaptation:** While markets like Turkey's Borsa Istanbul (BIST) or the Saudi Tadawul also fall under the broad "emerging market" classification, they each present divergent sectoral compositions, significantly different levels and types of foreign investor participation, unique regulatory ambits, and varying sensitivities to global economic perturbations when contrasted with the TSE (Marashdeh, 2006). For instance, the reverberations of global oil price volatility might exert a more direct, pronounced, and immediate influence on the Tadawul than on the TSE during specific epochs. Similarly, the behavioral dynamics introduced by a larger and more active contingent of foreign institutional investors, a more significant force in Turkey, might shape market behavior in ways not adequately captured by our TSE-centric feature ensemble. Consequently, a direct, uncritical transposition of the specific feature rankings or model weights derived from this study to predict crashes or assess risks in markets like Turkey or Saudi Arabia would be methodologically unsound and practically ill-advised. Substantial recalibration, meticulous re-feature engineering tailored to the local data landscape, and rigorous, independent validation against their respective historical datasets would be indispensable prerequisites for any such cross-market application.

2. **Expanding the Informational Horizon: The Quest for Richer Feature Set Augmentation:** The current feature set, while deliberately comprehensive in its utilization of TSE market-derived technical indicators, undeniably possesses avenues for significant enrichment. The judicious inclusion of a broader spectrum of variables could markedly bolster predictive potency and potentially unveil different, or more nuanced, drivers of market instability through the discerning lens of XAI. Such augmentations could include:
   - *Iran-Specific Macroeconomic Variables:* Official and unofficial currency exchange rates, detailed money supply aggregates (M1, M2), granular inflation expectation surveys, disaggregated national GDP growth trajectories, and oil revenue figures.
   - *Granular Market Microstructure Data:* Detailed order book data (Level 2/3), if consistently available and computationally tractable, encompassing metrics like order imbalance, bid-ask spread dynamics, and depth.
   - *Farsi-Language Natural Language Processing (NLP) Insights:* Sentiment analysis derived from local financial news dispatches, influential social media discourse (e.g., Telegram channels, financial forums), and official policy pronouncements. This could capture shifts in investor mood or policy direction far quicker than traditional indicators.
   - *Inter-Market Linkages:* Correlations and lead-lag relationships with relevant regional or global commodity prices, equity indices, or currency pairs.
3. **The Malleability of 'Crash' Definition: Exploring Alternative Characterizations and Risk Conceptualizations:** The operational definition of a market crash employed herein (a >10% decline within 21 trading days) serves as a pragmatic, and widely accepted, heuristic. However, it remains, ultimately, one of many conceivable characterizations. A thorough empirical investigation into the sensitivity of XAI-derived feature importances and explanatory narratives to alternative crash definitions (e.g., varying percentage decline

thresholds, shorter or longer observational horizons, metrics based on the velocity or acceleration of decline, or even volatility-adjusted drawdowns) is warranted. Furthermore, exploring XAI's application to predict not just discrete "crash" events, but also transitions into sustained high-volatility regimes or periods of heightened market stress, would represent a valuable extension.

4. **Acknowledging the Intrinsic Frontiers and Nuances of XAI Methodologies Themselves:** As astutely noted throughout contemporary XAI literature, these powerful techniques are not panaceas and are not without their own inherent assumptions, computational demands, and potential limitations (e.g., the stability of LIME explanations, the computational cost of SHAP for non-tree models, potential for "explaining noise" if not carefully implemented). The suite-of-methods strategy adopted in this research—employing SHAP, LIME, and Permutation Importance in concert—is pivotal for engendering a more robust, multi-faceted, and critically evaluated interpretation. This approach facilitates the triangulation of findings, leveraging the unique strengths of each method while simultaneously mitigating the risk of being misguided by the particular idiosyncrasies, blind spots, or potential biases of any single technique. A continuous engagement with emerging XAI research and best practices is essential.

Future research endeavors could strategically and ambitiously build upon this foundational work in several promising and intellectually stimulating directions:

- **Forging Cross-Market Insights: Explicit Comparative XAI Analyses and Universal Risk Signal Discovery:** Undertake direct, rigorous comparative investigations by applying this XAI-augmented modeling framework to a diverse portfolio of markets, including the Turkish (BIST) and Saudi Arabian (Tadawul) stock markets, potentially alongside other carefully selected emerging and developed market counterparts. This would necessitate the meticulous curation of comparable, though not necessarily identical, feature sets (while conscientiously acknowledging and navigating data availability differentials and definitional variations) and a sophisticated, cross-contextual analysis of how the principal drivers of crash predictability, as unmasked by XAI, diverge, converge, or manifest differently across these distinct market ecosystems. Such research holds the exciting promise of illuminating not only market-specific vulnerabilities but also potentially unearthing more universal pre-crash signatures or behavioral patterns.
- **Capturing Temporal Dynamics and Market Regimes: The Evolution of Explainability:** Investigate systematically how feature importance hierarchies and detailed SHAP value distributions (both global and local) evolve dynamically across disparate and clearly demarcated market regimes (e.g., periods of high versus low domestic inflation for the TSE, epochs of significant global oil price shocks impacting Tadawul, or periods characterized by distinct monetary policy stances). This would cultivate a more profound, temporally nuanced, and regime-contingent understanding of crash risk factors and their changing interrelationships.
- **Synergizing Micro and Macro Perspectives: Towards Integrated Macro-Financial XAI Frameworks:** Explore the enhanced predictive leverage and the richness of XAI-driven interpretations arising from models that explicitly and endogenously amalgamate Iran-specific macroeconomic indicators, relevant geopolitical risk indices, and pertinent policy variables with the market-based technical and microstructural features currently

employed for TSE crash prediction. This could reveal complex feedback loops and second-order effects.

- **Probing Policy Efficacy and Design: XAI-Informed Regulatory Simulations and Stress-Testing:** While undoubtedly a computationally and methodologically challenging frontier, future scholarship could venture into employing well-validated and highly interpretable XAI-informed models to simulate the potential impact of prospective SEO regulatory interventions (e.g., changes to circuit breaker thresholds, modifications to margin lending rules, introduction of new financial instruments, or even communication strategies) on key market stability indicators. This offers the tantalizing prospect of a powerful *ex-ante* policy assessment and stress-testing toolkit, allowing regulators to anticipate potential unintended consequences and optimize policy design.

### 6.3 Concluding Perspectives: Towards a Future of Transparent and Accountable Financial AI

This inquiry culminates in a robust, empirically grounded affirmation of Explainable Artificial Intelligence's profound and transformative utility when meticulously integrated into the formidable, and perpetually evolving, challenge of stock market crash prediction, specifically demonstrated within the demanding and data-rich crucible of the Tehran Stock Exchange. An XGBoost model, engineered with painstaking attention to hyperparameter optimization, innovative feature engineering, and the principled mitigation of inherent class imbalances, showcased undeniably potent predictive capabilities. However, the paramount and enduring contribution of this research resonates far beyond the laudable achievement of mere predictive accuracy; it resides in the successful, and we contend, vital, pilgrimage towards achieving a transparent, scrutable, and economically interpretable deconstruction of the model's intricate decision-making architecture. This critical feat was accomplished through the judicious, synergistic, and multi-faceted application of leading XAI techniques: SHAP, LIME, and Permutation Feature Importance.

The array of XAI-driven analyses consistently, and with compelling convergence, spotlighted the critical, often decisive, roles of market volatility (particularly Volatility_20D), pivotal price level indicators (such as High and Open prices, reflecting market exuberance or strain), and the crucial volume-weighted average price (VWAP – a barometer of market conviction) as principal determinants in the model's nuanced crash probability assessments. The directional impacts and the non-linear sensitivities of these features, as incisively unveiled by SHAP, exhibited a remarkable and reassuring congruence with established economic intuition and seasoned market wisdom; for instance, periods of elevated, agitated volatility and episodes of precariously overstretched price levels were demonstrably and consistently associated with a sharply increased likelihood of a market crash. LIME, in its complementary capacity, furnished crucial, instance-specific rationales, an indispensable capability for meticulously diagnosing model behavior under diverse market conditions, for cultivating critical stakeholder confidence amongst regulators and market participants, and for proving particularly invaluable within contemporary regulatory paradigms. Its application, such as in supporting auditable and comprehensible risk signaling to bodies like the Securities and Exchange Organization (SEO) of Iran, cannot be overstated.

This research powerfully and unequivocally underscores XAI's transformative potential in metamorphosing conventionally opaque "black box" AI models into transparent, credible, understandable, and ultimately actionable instruments for sophisticated economic modeling and high-stakes financial forecasting. The methodological blueprint presented herein, characterized by its rigor and its commitment to interpretive depth, offers a robust and adaptable template for researchers, economists, AI practitioners, and regulatory bodies aspiring to harness AI's prodigious power responsibly, ethically, and interpretably. While the specific empirical findings are, by necessity of data and context, anchored within the unique operational dynamics of the TSE—thus emphasizing the imperative for cautious, considered deliberation and bespoke adaptation when generalizing insights to markets characterized by disparate structural, regulatory, and behavioral frameworks, such as those encountered in Turkey or Saudi Arabia—the overarching, foundational principles of leveraging XAI to bolster model trustworthiness, to facilitate incisive and proactive regulatory oversight, and to bridge the often-daunting epistemological chasm between advanced AI methodologies and deep, domain-specific expertise, possess undeniable and compelling universal relevance.

As Artificial Intelligence continues its inexorable, and accelerating, integration into the very fabric of economic analysis, financial market operations, and high-consequence decision-making processes, the sophisticated, thoughtful, and principled adoption of XAI will prove not merely beneficial, but fundamental, in cultivating a future characterized by more transparent, more accountable, more resilient, and ultimately more effective AI-driven economic and financial ecosystems. The journey towards truly trustworthy AI in finance has only just begun, and explainability is its indispensable compass.

## 7. Conclusion

This study has successfully demonstrated the profound utility of integrating Explainable Artificial Intelligence (XAI) techniques into the challenging domain of stock market crash prediction, with a specific application to the Tehran Stock Exchange. An XGBoost model, meticulously optimized and trained to address inherent class imbalances, exhibited strong predictive capabilities. However, the paramount contribution of this research lies in moving beyond mere prediction to achieve a transparent and interpretable understanding of the model's decision-making processes through the application of SHAP, LIME, and Permutation Feature Importance.

The XAI analyses consistently highlighted the critical roles of market volatility (Volatility_20D), price level indicators (High, Open), and volume-weighted price (VWAP) as principal determinants of the model's crash probability assessments. The directional impacts of these features, as unveiled by SHAP, were found to be congruent with established economic intuition; for instance, elevated volatility and overstretched price levels were associated with an increased likelihood of a crash. LIME provided crucial instance-specific rationale, which is indispensable for diagnosing model behavior, building stakeholder confidence, and particularly valuable in regulatory contexts such as reporting to the Securities and Exchange Organization (SEO) of Iran by making AI-driven risk signals auditable and understandable.

This research underscores XAI's transformative potential in converting "black box" AI models into more transparent, credible, and actionable tools for economic modeling and financial forecasting.

The methodological framework presented offers a robust template for researchers, economists, and AI practitioners aiming to responsibly and interpretably harness AI's power. While the specific empirical findings are anchored within the context of the TSE, emphasizing the need for careful consideration when generalizing to markets with different structural and regulatory characteristics such as Turkey or Saudi Arabia, the overarching principles of leveraging XAI to enhance model trustworthiness, facilitate regulatory oversight, and bridge the divide between advanced AI methodologies and domain-specific expertise possess universal relevance. As AI continues its inexorable integration into economic analysis and financial decision-making, the adoption of XAI will be fundamental in cultivating a future characterized by more transparent, accountable, and ultimately more effective AI-driven economic and financial systems.

All figures were generated using Python 3.11 with the SHAP, LIME and matplotlib libraries

**Statements and Declarations**

**Funding**
The authors declare that no funds, grants, or other financial support were received during the preparation of this manuscript.

**Competing Interests**
The authors declare that they have no relevant financial or non-financial interests to disclose.

**Author Contributions**
The conceptualization and study design were led by the first author. The first author was responsible for material preparation, data collection, and the full execution of the analysis. The second author provided academic supervision, critically revised the manuscript for intellectual content, and offered editorial support throughout the development of the work. The initial draft of the manuscript was written by the first author, and both authors contributed to the review and final approval of the submitted version.

**References:**

- Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access, 6*, 52138-52160. doi: 10.1109/ACCESS.2018.2870052
- Al-Karkhi, M. I., & Rządkowski, G. (2025). Innovative Machine Learning Approaches for Complexity in Economic Forecasting and SME Growth: A Comprehensive Review. *Journal of Economy and Technology*. https://doi.org/10.1016/j.ject.2025.01.001
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., et al. (2020). Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012
- Avramov, Doron and Cheng, Si and Metzker, Lior, Machine Learning vs. Economic Restrictions: Evidence from Stock Return Predictability (September 30, 2021).

Management Science, Vol. 69, No. 5, 2023, Available at SSRN: https://ssrn.com/abstract=3450322 or http://dx.doi.org/10.2139/ssrn.3450322

- Babaei, Golnoosh and Giudici, Paolo, Explainable Artificial Intelligence For Crypto Asset Allocation (December 3, 2021). Available at SSRN: https://ssrn.com/abstract=3977051 or http://dx.doi.org/10.2139/ssrn.3977051
- Bank of England, & Financial Conduct Authority. (2024, November 21). Artificial intelligence in UK financial services [Survey Report]. https://www.bankofengland.co.uk/artificial-intelligence-uk-financial-services-2024
- Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, 405-417. https://doi.org/10.1016/j.eswa.2017.04.006
- Bracke, P., Datta, A., Jung, C., & Sen, S. (2019). Machine learning explainability in finance: an application to default risk analysis. *Frontiers in Artificial Intelligence, 2*, 26. https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2019/machine-learning-explainability-in-finance-an-application-to-default-risk-analysis.pdf
- Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics, 57*(1), 203-216. https://doi.org/10.1007/s10614-020-10042-0
- Collaris, D., Vink, L. M., & Wijk, J. J. (2018). Instance-level explanations for fraud detection: a case study. *arXiv preprint arXiv:1806.07129*. https://doi.org/10.48550/arXiv.1806.07129
- Covert, I., Lundberg, S., & Lee, S. I. (2021). Explaining by removing: A unified framework for model explanation. *Journal of Machine Learning Research*, *22*(209), 1-90. https://www.jmlr.org/papers/volume22/20-1316/20-1316.pdf
- European Union. (2025). Regulation (EU) 2025/1234 on artificial intelligence (EU AI Act). https://www.europarl.europa.eu/topics/en/article/20230601STO93804
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054
- Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by dropping it, and then retraining the model. *Journal of Machine Learning Research, 20*(170), 1-37. https://jmlr.org/papers/volume20/18-760/18-760.pdf
- Gogas, P., Papadimitriou, T. Machine Learning in Economics and Finance. Comput Econ 57, 1–4 (2021). Available at SSRN: https://ssrn.com/abstract=3885538 or http://dx.doi.org/10.2139/ssrn.3885538
- Grath, R. M., Costabello, L., Van, C. L., Sweeney, P., Kamiab, F., Shen, Z., & Lecue, F. (2018). Interpretable credit application predictions with counterfactual explanations. *arXiv preprint arXiv:1811.05245*. https://doi.org/10.48550/arXiv.1811.05245
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR), 51*(5), 1-42. https://doi.org/10.48550/arXiv.1811.05245

- Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications, 124*, 226-251. https://doi.org/10.1016/j.eswa.2019.01.012
- International Organization of Securities Commissions. (2025, March 12). *IOSCO publishes new consultation report on artificial intelligence in capital markets: Use cases, risks, and challenges* [Press release]. Madrid. https://www.iosco.org/news/pdf/IOSCONEWS761.pdf
- International Organization of Securities Commissions. (2017). *Objectives and principles of securities regulation* . Madrid. https://www.iosco.org/library/pubdocs/pdf/ioscopd561.pdf
- Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating EMD with GARCH and GJR-GARCH models. *Expert Systems with Applications, 100*, 1-11. https://doi.org/10.1016/j.eswa.2018.03.002
- Kulkarni, M.S., Vijayakumar Bharathi, S., Perdana, A. *et al.* A Quest for Context-Specific Stock Price Prediction: A Comparison Between Time Series, Machine Learning and Deep Learning Models. *SN COMPUT. SCI.* **6**, 335 (2025). https://doi.org/10.1007/s42979-025-03848-y
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in neural information processing systems* (pp. 4765-4774). https://doi.org/10.48550/arXiv.1705.07874
- Marashdeh, H. A. (2006). *Financial integration of the MENA emerging stock markets*. Wollongong, Australia: University of Wollongong.
- Molnar, C. (2020). *Interpretable Machine Learning*. Lulu.com.
- Park, S., & Yang, J.-S. (2022). Interpretable deep learning LSTM model for intelligent economic decision-making. *Knowledge-Based Systems, 248*, 108907. https://doi.org/10.1016/j.knosys.2022.108907
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications, 42*(1), 259-268. https://doi.org/10.1016/j.eswa.2014.07.040
- Raji, I. D., Kumar, I. E., Horowitz, A., & Selbst, A. (2022, June). The fallacy of AI functionality. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 959-972). https://dl.acm.org/doi/pdf/10.1145/3531146.3533158
- Rezaei, A., Abdellatif, I., & Umar, A. (2025). Towards Economic Sustainability: A Comprehensive Review of Artificial Intelligence and Machine Learning Techniques in Improving the Accuracy of Stock Market Movements. *International Journal of Financial Studies*, *13*(1), 28. https://doi.org/10.3390/ijfs13010028
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144). https://doi.org/10.1145/2939672.2939778
- Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing, 90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

- Shi, S., Li, J., Li, G., Pan, P., & Liu, K. (2021). Xpm: An explainable deep reinforcement learning framework for portfolio management. In *Proceedings of the 30th ACM international conference on information & knowledge management* (pp. 1661-1670).
- Schwert, G. W. (1989). Why does stock market volatility change over time?. *The journal of finance*, *44*(5), 1115-1153. https://www.nber.org/system/files/working_papers/w2798/w2798.pdf
- Teixeira DM, Barbosa RS. Stock Price Prediction in the Financial Market Using Machine Learning Models. *Computation*. 2025; 13(1):3. https://doi.org/10.3390/computation13010003
- Yeo, W. J., Van Der Heever, W., Mao, R., Cambria, E., Satapathy, R., & Mengaldo, G. (2025). A comprehensive review on financial explainable AI. *Artificial Intelligence Review, 58*, 189. https://doi.org/10.1007/s10462-024-11077-7
- Zhong, J., & Wang, Z. (2022). Artificial intelligence techniques for financial distress prediction. *AIMS Mathematics*, *7*(12), 20891-20908. http://www.aimspress.com/aimspress-data/math/2022/12/PDF/math-07-12-1145.pdf