# Trend-Based Encoding of Exogenous Time-Series for Interpretable Financial Prediction

Khudran M. Alzhrani
Department of Computers-College of Engineering and Computing in Al-Qunfudhah
Umm Al-Qura University, Makkah, Saudi Arabia

***Abstract*—Integrating heterogeneous exogenous data into financial prediction models is challenging due to scale mismatches and semantic ambiguity. We propose a trend-encoding framework that transforms raw exogenous time-series into directional binary representations, improving predictive robustness while preserving interpretability. Using Saudi stock market data with COVID-19 indicators, we evaluate predictive models under baseline and trend-enhanced configurations. Results show that trend encoding consistently enhances predictive stability over raw inputs. Interpretable models benefit disproportionately, achieving performance comparable to black-box methods. Sectoral analysis reveals heterogeneous sensitivities: Banking responds strongly to case and mortality trends, Energy to recovery indicators, while Food & Beverages shows weaker alignment. These findings show that trend-based encoding of exogenous signals can improve cross-domain financial prediction, particularly for interpretable models.**

***Keywords*—Trend-based encoding; exogenous time-series; interpretable machine learning; financial prediction forecasting**

## I. Introduction

Financial markets are increasingly influenced by a wide range of exogenous signals, from public health trends and policy shifts to environmental and geopolitical events. These signals, when integrated into predictive models, have the potential to enhance forecasting accuracy by capturing real-world context that pure market metrics often miss [1], [2]. However, the raw form of such exogenous data — often noisy, scale-dependent, and temporally misaligned with financial time series — presents a significant barrier to effective and interpretable modeling. Representation of external signals thus emerges as a critical, yet underexplored, challenge in financial machine learning [3].

Traditional stock prediction models have predominantly relied on historical price and volume data. More recent approaches incorporate external datasets — such as sentiment from news and social media, government announcements, or macroeconomic indicators — to improve robustness, especially during volatile periods [4]. However, these methods often feed raw or minimally processed values directly into complex, non-interpretable models, obscuring the relationship between the external signal and the predicted outcome. This practice limits transparency and makes it difficult for financial decision-makers to understand why a model makes a particular prediction, which is crucial for risk assessment and regulatory compliance [5].

The core challenge lies in transforming heterogeneous, real-world exogenous data into a format that is both informative for prediction and amenable to interpretation. Raw indicators like daily case counts, policy severity indices, or emission levels have different scales, reporting frequencies, and noise profiles. When used directly, they can degrade model performance, introduce overfitting, and render the contributions of individual features inscrutable, particularly in inherently interpretable models like decision trees.

This study addresses this representation gap by proposing a trend-based encoding framework for exogenous time-series data. Instead of modeling raw values, we convert external indicators into directional, binary trend signals that capture whether the underlying metric is increasing or decreasing relative to the previous period. This transformation offers several advantages: it normalizes scale, reduces noise from magnitude fluctuations, provides a consistent semantic meaning (upward/downward pressure), and aligns naturally with the decision logic of interpretable models. We demonstrate this framework using COVID-19 epidemiological data — a prominent, publicly available, and high-frequency exogenous signal — as a case study to validate our approach. The pandemic serves not as the primary subject of analysis, but as an ideal real-world data source to test the encoding methodology under conditions of significant market disruption.

The contributions of this work are as follows:

- We propose a trend-based encoding approach that transforms raw exogenous time-series into directional features suitable for interpretable financial models.
- We design a causally valid data integration pipeline that aligns exogenous indicators with financial time-series while preventing look-ahead bias.
- Through comparative experiments, we show that interpretable models benefit more consistently from trend-encoded exogenous signals than non-interpretable models.
- We provide a sector-level analysis demonstrating that the impact of exogenous trends varies across Banking, Energy, Food & Beverages, and Health Care sectors.

The rest of the study is organized as follows: Section II reviews related work on exogenous data integration in financial prediction and highlights the gap in data representation strategies. Section III describes the datasets used and the construction of the trend-encoded financial-exogenous dataset. Section IV details the trend-encoding framework and the predictive modeling approach. Section V presents the experimental setup, results, and analysis. Finally, Section VI summarizes the findings and suggests directions for future work.

## II. Related Work

This section reviews literature relevant to our proposed trend-encoding framework, organized into three interconnected themes: the integration of exogenous data into financial prediction models, the representation and modeling of trends in time-series data, and the pursuit of interpretability in financial machine learning. A critical observation across these themes is the relative under-exploration of simple, transparent data transformations for aligning heterogeneous external signals with financial time-series.

### A. Exogenous Data in Financial Prediction

The use of non-market data to augment financial forecasting has grown substantially, particularly during periods of high market volatility driven by external shocks. Early works focused on macroeconomic indicators, news sentiment, and social media data [6]. The COVID-19 pandemic served as a catalyst, highlighting the predictive value of high-frequency, real-world exogenous signals. Several studies directly incorporated pandemic metrics: [7] used daily case and death counts with LSTM models to forecast stock prices; [8] integrated case counts and sentiment indices within a multi-task deep learning architecture; and [9] combined infection data with Twitter sentiment in a system using SVM and tree-based models. Other works employed related proxies, such as government mobility indices [10] or predicted mortality peaks [2]. A common thread is the treatment of these exogenous indicators as raw numerical inputs or as features for complex embedding layers within deep neural networks. While effective for predictive accuracy in specific contexts, this approach often obscures the direct, interpretable relationship between the external signal and the market response, especially when scale differences and noise are not explicitly addressed through the data representation itself.

### B. Trend and Directional Encoding in Time-Series

The concept of simplifying complex time-series into directional or trend-based representations has precedent in both finance and signal processing. In technical analysis, metrics like moving average convergence divergence (MACD) or the directional movement index (DMI) distill price action into momentum and trend signals [11]. In broader time-series analysis, change-point detection and regime-switching models aim to identify shifts in the underlying data-generating process [12]. However, the application of such concepts specifically to preprocess and encode *exogenous* signals for integration with a separate financial target series remains relatively limited. Existing studies in nowcasting and leading indicator analysis typically rely on scale-preserving transformations, including year-over-year changes, differencing, and Z-score normalization to align heterogeneous indicators [13], [14], [6]. While effective for capturing magnitude-driven dynamics, these approaches retain scale sensitivity and do not explicitly abstract directional information. The work of [15], which applied stationary wavelet transform to COVID-19 case data before feeding it into a BDLSTM, represents a form of multi-resolution filtering but remains within a complex, non-interpretable modeling framework. Our proposed method distinguishes itself by advocating for a minimal, scale-invariant binary encoding (up/down trend) as a generic preprocessing step, deliberately sacrificing magnitude information to gain robustness, alignment simplicity, and, crucially, interpretability.

### C. Interpretability in Financial Machine Learning

The demand for explainable AI in finance is driven by regulatory requirements, risk management needs, and the practical value of actionable insights [16]. Interpretable models, such as decision trees (CART, C5.0), rule-based systems, and linear models with constraints, offer transparency by design [17]. Post-hoc explanation techniques like SHAP [10] and LIME can be applied to black-box models but provide approximations rather than faithful representations of the underlying logic. A key challenge in employing inherently interpretable models is feature engineering: these models perform best when inputs are semantically clear, have manageable ranges, and exhibit logical relationships to the target [18]. Raw exogenous data—with erratic scales, missing values, and complex temporal dynamics—often violates these conditions, pushing practitioners toward black-box models that can learn complex embeddings automatically [19]. Consequently, while many studies compare model performance [20], [21], few systematically investigate how different *data representations* specifically affect the performance gap between interpretable and non-interpretable models. Our work posits that a tailored encoding strategy, such as trend-based transformation, can bridge this gap by creating features that are inherently compatible with the logical splitting mechanisms of interpretable models, thereby unlocking their potential without sacrificing predictive utility.

Prior research establishes the value of exogenous data and highlights the importance of interpretability but leaves a gap regarding systematic, transparent methods for converting raw external signals into features optimized for interpretable financial prediction. This study addresses that gap by proposing and evaluating a trend-based encoding framework.

## III. Data Description

This study integrates two heterogeneous, real-world time-series datasets to develop and validate a trend-encoding framework for cross-domain prediction. The core challenge is their inherent mismatch: the financial data (daily stock market activity) and exogenous public health data differ in temporal resolution, scale, reporting conventions, and semantic meaning. Specifically, the financial dataset consists of daily trading records from the Saudi Stock Exchange (Tadawul) [22], while the exogenous dataset comprises daily COVID-19 epidemiological reports published by the KAPSARC Data Portal [23]. This section details each dataset's characteristics and the preprocessing steps required to create a causally aligned, prediction-ready integrated dataset.

### A. Exogenous Public Health Signal Dataset

We utilize a publicly available epidemiological dataset as a representative high-frequency exogenous signal. The dataset consists of daily reports released by a national public health authority and is structured around four primary numerical indicators: Cases, Mortalities, Recoveries, and Tested. Each record includes a Report_Type attribute (Cumulative or Daily), a reporting Date, and geographic descriptors at the city and region levels.

TABLE I. COVID-19 DATASET STATISTICS

| Report Type | Indicator | Records |
|---|---|---|
| Cumulative | Cases | 193,121 |
| *Cumulative* | *Active* | *185,467 (excluded)* |
| *Cumulative* | *Critical* | *1,090 (excluded)* |
| Cumulative | Mortalities | 93,827 |
| Cumulative | Recoveries | 189,891 |
| Daily | Cases | **72,542 (retained)** |
| Daily | Recoveries | **70,444 (retained)** |
| Daily | Mortalities | **6,811 (retained)** |
| *Daily* | *Tested* | *1,151 (excluded)* |

The raw dataset exhibits several characteristics commonly observed in real-world exogenous signals, necessitating careful preprocessing and motivating the adoption of a trend-encoding strategy, as summarized in Table I. First, the reported values span multiple orders of magnitude, ranging from single-digit mortalities to thousands of daily cases. Direct incorporation of such absolute values would introduce substantial scale variance, which is particularly detrimental to model stability and interpretability in rule-based learning algorithms.

Second, the indicators are reported under both cumulative and daily formats. To capture the flow of newly released information most relevant to short-term market reactions, only Daily report types are retained for the Cases, Mortalities, and Recoveries indicators, while cumulative records are excluded from the analysis. Third, the Tested indicator exhibits substantial reporting gaps and inconsistent temporal coverage, with only 1,151 daily observations available, rendering it unsuitable for reliable daily modeling; it is therefore excluded.

In addition, spatial aggregation is applied to align the exogenous signals with national-level financial analysis. City- and region-level records are aggregated by date to produce a single national daily value for each retained indicator. Finally, administrative event annotations, such as lockdown announcements, are excluded to isolate the predictive contribution of structured numerical indicators and avoid conflating trend signals with discrete policy events.

Collectively, these properties—scale heterogeneity, reporting duality, incomplete coverage, and spatial fragmentation—illustrate the challenges of integrating raw exogenous data into financial prediction models. The trend-encoding framework described in Section IV is specifically designed to address these challenges through scale-invariant transformation and directional abstraction, yielding features that are both robust and semantically aligned with interpretable modeling approaches.

### B. Stock Market Dataset

The financial dataset is extracted from a national stock exchange and consists of daily trading records for listed companies. Each record includes standard market attributes such as Open, High, Low, and Close prices, along with Volume Traded, Value Traded, and the corresponding Industry Group classification.

To evaluate the proposed framework across sectors with theoretically grounded sensitivity to pandemic-related dynamics, four industries are selected based on established economic and financial literature. The Banking sector is included due to its close association with financial stability and systemic risk during periods of economic uncertainty. The Energy sector is considered because of its strong dependence on mobility patterns and industrial activity, both of which were significantly affected during the pandemic. The Food & Beverages sector is selected to capture the effects of supply chain disruptions and shifts in consumer demand, while the Health Care Equipment & Services sector is included given its direct exposure to changes in healthcare spending and public health conditions. This sectoral selection enables a robust assessment of whether the effectiveness of trend encoding varies across heterogeneous market responses.

The prediction target is defined as the next-day directional stock price movement, denoted by $Y_t$, and derived from the closing price. Specifically,

$$Y_t = \begin{cases} 1 & \text{if } Close_t > Close_{t-1}, \\ 0 & \text{otherwise.} \end{cases}$$

To prevent look-ahead bias, all financial features are lagged by one trading day, ensuring that the predictive models rely exclusively on information that would have been available at the time of the trading decision.

### C. Integrated Dataset for Trend-Based Analysis

The financial and exogenous datasets are merged on the Date field. To ensure causal validity, exogenous health indicators reported on day $d$ are paired with the stock market movement observed on day $d{+}1$, reflecting the fact that market participants respond to publicly available information from the preceding trading day.

The resulting pre-encoded dataset consists of lagged financial features together with the raw daily values of the three retained exogenous indicators—Cases, Mortalities, and Recoveries—which together form the baseline representation. The proposed framework subsequently applies the trend-encoding transformation $T(\cdot)$, formally defined in Eq. (1), to the exogenous indicators, converting them into binary, scale-invariant trend features (e.g., $CasesTrend_t$) suitable for integration with interpretable learning models.

TABLE II. SUMMARY OF INTEGRATED DATASET BY INDUSTRY

| Industry | Companies | Total Records | Training | Testing |
|---|---|---|---|---|
| Banks | 12 | 1,270 | 909 | 361 |
| Energy | 5 | 542 | 405 | 137 |
| Food & Beverages | 12 | 1,342 | 975 | 367 |
| Health Care Equipment & Services | 8 | 881 | 653 | 228 |

The integrated dataset is partitioned chronologically into 136 training days and 68 testing days, corresponding to approximately 71.5% and 28.5% of the available observations, respectively. This time-respecting split preserves the temporal structure of the data and ensures a realistic out-of-sample evaluation.

All analyses are conducted at the industry level to assess whether the benefits of trend encoding vary across sectors with different theoretical sensitivities to exogenous shocks.

The final structure of the integrated dataset across industries is summarized in Table II.

## IV. Methodology

This study proposes a formalized trend-based encoding framework designed to bridge the representational gap between heterogeneous exogenous time-series data and interpretable financial prediction models. The framework systematically addresses three core challenges in cross-domain financial modeling: scale and unit incompatibility, realistic causal temporal alignment, and semantic compatibility with interpretable learning algorithms. The methodology comprises two rigorously defined components: a multi-stage data engineering pipeline that transforms raw exogenous signals into aligned, semantically clear trend features, and a modeling strategy employing both intrinsically interpretable and non-interpretable machine learning algorithms to evaluate the effectiveness of the proposed encoding. The complete workflow is illustrated in Fig. 1, which outlines the sequential transformation, alignment, and integration processes. All models are evaluated under two feature configurations: a baseline setting using financial features only, and a trend-enhanced setting that augments financial inputs with the encoded exogenous trend vectors.

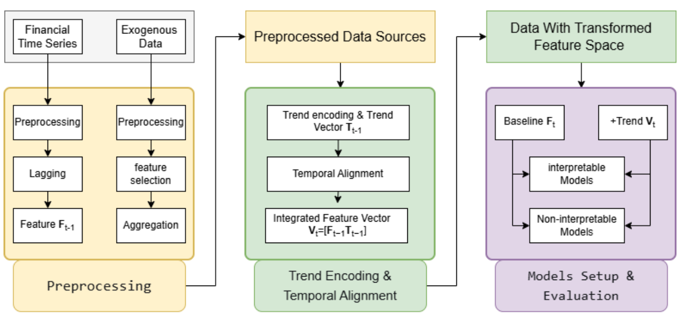


Fig. 1. Overview of the complete trend-encoding pipeline for cross-domain financial prediction. The pipeline processes financial and exogenous data streams independently before temporal alignment. Exogenous signals undergo trend encoding to produce binary trend vectors, while financial features are lagged to prevent look-ahead bias. The aligned features are then integrated and evaluated across both interpretable and non-interpretable models under baseline and trend-enhanced configurations.

### A. Trend Encoding and Temporal Alignment

The central methodological innovation is the transformation of raw exogenous numerical signals into binary, scale-invariant trend indicators, followed by a rigorous causal alignment with financial time-series. This process is formalized as a deterministic mapping from raw observational space to a semantically structured feature space suitable for interpretable learning.

*1) Trend encoding strategy:* For a raw exogenous time-series $X = \{x_1, x_2, \ldots, x_N\}$ of length $N$, we define a discrete-time trend-encoding function $T : \mathbb{R}^N \to \{0,1\}^N$ as:

$$T(x_t) = \begin{cases} 0 & \text{if } t = 1 \quad \text{(Neutral Baseline)} \\ 1 & \text{if } \Delta x_t > 0 \quad \text{(Upward Trend)} \\ 0 & \text{otherwise} \quad \text{(Non-Upward Trend)} \end{cases} \tag{1}$$

**Algorithm 1** Trend Encoding of Exogenous Time-Series

**Require:** Raw time-series $X = [x_1, x_2, \ldots, x_N]$
**Ensure:** Binary trend vector $\mathbf{T} = [T_1, T_2, \ldots, T_N]$

$T_1 \leftarrow 0$
**for** $t = 2$ to $N$ **do**
  **if** $x_t > x_{t-1}$ **then**
    $T_t \leftarrow 1$
  **else**
    $T_t \leftarrow 0$
  **end if**
**end for**
**return** $\mathbf{T}$

where, $\Delta x_t = x_t - x_{t-1}$ for $t \in \{2, 3, \ldots, N\}$. The initial observation $x_1$ is encoded as $T(x_1) = 0$, establishing a neutral baseline for the time series. This transformation can be implemented algorithmically as shown in Algorithm 1.

This encoding provides several mathematically grounded advantages for interpretable financial modeling. First, the transformation $T(\cdot)$ is invariant to positive affine transformations of the input, such that $T(\alpha x_t + \beta) = T(x_t)$ for any $\alpha > 0$ and $\beta \in \mathbb{R}$. This property eliminates scale dominance and unit-dependence across heterogeneous exogenous indicators, allowing features with different magnitudes and measurement units to be integrated without bias.

Second, by operating on the sign of the first-order difference $\Delta x_t$ rather than its absolute magnitude, the encoding reduces sensitivity to small-amplitude fluctuations and measurement noise. This improves robustness in the presence of reporting irregularities and short-term volatility commonly observed in real-world exogenous data sources.

Third, the resulting binary representation has clear semantic meaning, as each encoded value corresponds directly to the logical predicate $\Delta x_t > 0$. This alignment with rule-based reasoning makes the transformed features naturally compatible with interpretable learning algorithms, such as decision trees and rule-based classifiers, whose decision logic relies on explicit Boolean conditions.

Finally, the encoding enforces dimensional consistency across multiple heterogeneous indicators. A collection of exogenous time-series $X^{(1)}, X^{(2)}, \ldots, X^{(m)}$ is mapped into a common binary feature space $\{0,1\}^m$, enabling uniform integration with financial predictors and facilitating fair comparison across indicator combinations.

For the three exogenous indicators retained in this study—Cases, Mortalities, and Recoveries—the encoding yields a multivariate daily trend vector given by:

$$\mathbf{T}_t = \left[T(\text{Cases}_t),\ T(\text{Mortalities}_t),\ T(\text{Recoveries}_t)\right]^\top \in \{0,1\}^3.$$

*2) Independent preprocessing and causal alignment:* Ensuring strict temporal causality is essential for valid financial prediction, as any use of information unavailable at decision time results in look-ahead bias and inflated performance estimates. The proposed framework enforces causal validity by processing financial and exogenous data streams independently and applying a rigorous alignment strategy prior to feature integration.

Let $\mathbf{M}_t = [O_t, H_t, L_t, C_t, V_t]^\top$ denote the raw daily market observations at trading day $t$, corresponding to the open, high, low, close prices and trading volume. Following standard preprocessing steps, including the treatment of missing values and adjustments for corporate actions, the prediction target is defined as the next-day directional price movement:

$$Y_t = \mathbb{I}(C_t > C_{t-1}),$$

where, $\mathbb{I}(\cdot)$ is the indicator function. To prevent leakage of future information, all financial predictors are lagged by one trading period, yielding the feature vector $\mathbf{F}_{t-1} = [O_{t-1}, H_{t-1}, L_{t-1}, C_{t-1}, V_{t-1}]^\top$, which represents the complete set of market information available at prediction time.

Each exogenous indicator $X^{(i)}$ is processed independently of the financial stream to preserve semantic clarity and causal separation. The preprocessing pipeline consists of three sequential steps: temporal filtering to retain daily observations, spatial aggregation to produce national-level daily values, and application of the trend-encoding function $T(\cdot)$ defined in Eq. (1). This transformation produces a daily binary trend representation for each indicator, and the resulting exogenous trend vector is given by:

$$\mathbf{T}_t = \left[T(X_t^{(1)}),\ T(X_t^{(2)}),\ T(X_t^{(3)})\right]^\top.$$

Temporal alignment between the financial and exogenous streams is achieved using a bounded forward-fill mechanism. Let $\mathcal{D}_{\text{trade}}$ denote the set of trading days and $\mathcal{D}_{\text{exog}}$ the set of days on which exogenous indicators are reported. For each trading day $t \in \mathcal{D}_{\text{trade}}$, the aligned exogenous feature vector is defined as:

$$\tilde{\mathbf{T}}_{t-1} = \mathbf{T}_{\tau(t-1)},$$

where,

$$\tau(t-1) = \max\left\{d \in \mathcal{D}_{\text{exog}} : d \leq t-1 \text{ and } t-1-d \leq \delta_{\max}\right\},$$

with $\delta_{\max} = 1$ day. This bounded alignment ensures that only information publicly available prior to the trading decision is incorporated, while avoiding excessive temporal gaps that could distort trend continuity.

Together, independent preprocessing and bounded causal alignment guarantee that the integrated feature representation respects the chronological order of information flow, providing a realistic and methodologically sound foundation for subsequent predictive modeling.

*3) Integrated feature construction:* The final feature vector for predicting $Y_t$ is the concatenation of lagged financial features and aligned exogenous trends [see Eq. (2)]:

$$\mathbf{V}_t = \begin{bmatrix} \mathbf{F}_{t-1} \\ \tilde{\mathbf{T}}_{t-1} \end{bmatrix} \quad \text{where} \quad \mathbf{F}_{t-1} \in \mathbb{R}^5,\ \tilde{\mathbf{T}}_{t-1} \in \{0,1\}^3. \quad (2)$$

This integrated representation projects raw exogenous magnitudes into a semantically structured binary subspace, directly resolving scale incompatibility while preserving economically meaningful directional information.

### B. Model Selection for Interpretability Analysis

To rigorously evaluate the impact of trend encoding and test whether it preferentially benefits interpretable models, we employ a curated suite of machine learning models spanning the interpretability spectrum. Each model is chosen based on its representational transparency and learning characteristics, with particular attention to how they process the integrated feature space of financial indicators and exogenous trend signals.

*1) Interpretable models:* We select three tree-based algorithms that offer intrinsic interpretability through explicit rule structures, making them particularly suitable for analyzing how exogenous trend features interact with financial predictors.

*a) C5.0:* An advanced decision tree algorithm that optimizes information gain ratio for splitting. In our framework, C5.0 evaluates whether binary trend features like $T(\text{Cases})$ or combinations such as $T(\text{Cases}) = 1 \wedge T(\text{Mortality}) = 0$ provide predictive power for next-day stock direction. It generates both tree structures and simplified rule sets, with optional boosting and pruning. The model's interpretability stems from its ability to produce human-readable if-then rules directly relating market conditions and exogenous trends, for example: IF $\text{Volume}_{t-1} > \theta_1$ AND $T(\text{Cases}_{t-1}) = 1$ THEN $\hat{Y}_t = 1$.

*b) CHAID (Chi-squared automatic interaction detection):* Employs Chi-squared tests of the form $\chi^2 = \sum \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$ to determine optimal multi-way splits, making it effective for detecting statistical dependencies between categorical trend features and stock movements. It merges statistically indistinguishable predictor categories, producing shallow, wide trees where each split—such as separating days with rising cases from days with stable or falling cases—has a clear statistical justification directly tied to market outcomes.

*c) CRT (Classification and regression tree):* Uses recursive binary partitioning based on Gini impurity minimization, where impurity at node $S$ is calculated as $\text{Gini}(S) = 1 - \sum_{c=1}^{C} p_c^2$, with $p_c$ representing the proportion of upward or non-upward movement classes. Feature importance is directly inferable from split proximity to the root, allowing us to identify whether exogenous trends like $T(\text{Recoveries})$ or traditional financial features like lagged closing price provide the primary discriminative power for sector-specific predictions.

These models are *intrinsically interpretable* because their decision logic can be fully articulated as Boolean expressions over input features—including both financial indicators and exogenous trend signals—enabling exact traceability of how pandemic developments influence trading predictions.

*2) Non-interpretable models:* For comparative analysis, we employ two models with strong predictive capacity but opaque internal representations, testing whether complex mappings can outperform interpretable rules when integrating heterogeneous data sources.

*a) Support Vector Machine (SVM) with RBF kernel:* Learns a separating hyperplane in a high-dimensional feature space induced by the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, where $\mathbf{x}_i$ represents combined financial-exogenous feature vectors. The decision function $f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b$ lacks direct correspondence to original features like $T$(Mortality) or lagged volume, requiring post-hoc techniques like SHAP for approximate explanations of how exogenous trends contribute to predictions.

*b) k-Nearest Neighbors (KNN):* Predicts stock direction based on the majority class among the $k$ most similar historical instances in Euclidean space $\mathcal{L}_2$, where similarity is computed across both financial and trend-encoded features. While locally explainable by examining neighboring days with comparable market conditions and pandemic trends, global interpretability is limited due to the absence of an explicit decision function and the curse of dimensionality when integrating multiple exogenous indicators.

These models are classified as *non-interpretable* in our context because providing faithful, globally valid explanations requires approximation methods that may not reflect true model mechanics—particularly problematic in financial applications where understanding the influence of specific exogenous signals like case trends or recovery rates is crucial for risk assessment and regulatory compliance.

## V. Experiments and Results

This section evaluates the effectiveness of the proposed trend-based encoding framework through a controlled experimental study designed to answer three research questions:

- RQ1: Does trend encoding of exogenous signals improve predictive performance over financial-only baselines?
- RQ2: Do inherently interpretable models benefit more from trend-encoded signals than non-interpretable models?
- RQ3: Which exogenous indicators are most relevant across different market sectors?

### A. Experimental Design and Setup

*1) Temporal partitioning and feature configurations:* The temporally aligned dataset is partitioned chronologically using a derived date attribute (Date_Derive) to prevent information leakage and ensure a realistic out-of-sample evaluation. Following standard practice in time-series forecasting, approximately 71.5% of company-day observations (136 trading days) are allocated for model training, while the remaining 28.5% (68 trading days) are reserved for testing, as summarized in Table II. This time-respecting split preserves the natural flow of information and prevents models from exploiting future data when predicting past outcomes.

As illustrated in Fig. 1, each learning model is trained and evaluated under two feature configurations designed to isolate the effect of incorporating exogenous information. In the baseline configuration, only lagged financial predictors $\mathbf{F}_{t-1}$ are used, reflecting conventional financial modeling practice without external signals. In contrast, the trend-enhanced configuration augments the financial features with aligned exogenous trend indicators, using the integrated feature vector $\mathbf{V}_t = [\mathbf{F}_{t-1} \quad \mathbf{T}_{t-1}]$. Comparing model performance across these two configurations enables a controlled assessment of the incremental predictive value introduced by the proposed trend-encoding framework.

TABLE III. Best Model–Indicator Combinations (Banking)

| Indicator | Learning Model | Metric | Value |
|---|---|---|---|
| Cases & Mortality | CHAID | F1-Score | 0.5882 |
| Mortality & Recoveries | KNN | Precision | 0.5273 |
| Cases & Mortality | CHAID | Recall | 0.7965 |
| Cases | CRT | Accuracy | 0.5880 |
| Baseline (No Indicators) | CRT | AUC | 0.5771 |

*2) Model implementation and evaluation protocol:* Experiments are conducted across four sectors—Banking, Energy, Food & Beverages, and Health Care—using five predictive models evaluated under two feature configurations: a financial-only baseline and a trend-enhanced setting incorporating exogenous indicators. All experiments follow a unified and reproducible pipeline with model training and evaluation carried out using IBM SPSS Modeler to ensure consistent preprocessing, model configuration, and evaluation procedures across all learning algorithms.

The proposed trend encoding framework was implemented as an IBM SPSS Modeler stream comprising data integration, feature derivation, and modeling nodes. Trend encoding was realized using conditional derive nodes applying Eq. (1), while temporal alignment was enforced through record linking and lag operations. All machine learning algorithms were instantiated using their respective SPSS Modeler nodes with default configurations maintained across sectoral analyses to ensure fair comparison.

Model performance is assessed using Accuracy, Precision, Recall, AUC, and F1-score. Given the class imbalance inherent in next-day directional stock movement prediction, F1-score is adopted as the primary evaluation metric throughout the analysis, as it provides a balanced assessment of Precision and Recall. For trend-enhanced models, multiple indicator configurations are evaluated to ensure a fair and conservative comparison against the single baseline configuration.

*3) Industry-level analysis:* Consistent with the sector selection in Section III-B, all experiments are conducted separately for each industry (Banks, Energy, Food & Beverages, Health Care Equipment & Services). This enables assessment of whether the benefits of trend encoding vary across market segments with different theoretical sensitivity to the exogenous shocks, providing actionable insights for sector-specific modeling strategies.

### B. Overall Impact of Trend Encoding Across Sectors

Fig. 2 compares average F1-scores between Baseline and Trend-Enhanced configurations across sectors. Trend encoding yields consistent performance gains in three of the four sectors, with particularly pronounced improvements in Banking and Energy.

TABLE IV. BEST MODEL–INDICATOR COMBINATIONS (ENERGY)

| Indicator | Learning Model | Metric | Value |
|---|---|---|---|
| Mortality & Recoveries | CHAID | F1-Score | 0.5487 |
| Cases | C5 | Precision | 0.4359 |
| Mortality & Recoveries | CHAID | Recall | 0.8378 |
| Cases & Recoveries | C5 | Accuracy | 0.5859 |
| Baseline (No Indicators) | SVM | AUC | 0.5663 |

TABLE V. BEST MODEL–INDICATOR COMBINATIONS (FOOD)

| Indicator | Learning Model | Metric | Value |
|---|---|---|---|
| Recoveries | C5 | F1-Score | 0.5714 |
| Cases | CRT | Precision | 0.4762 |
| Baseline (No Indicators) | C5 | Recall | 0.9444 |
| Cases | CRT | Accuracy | 0.6574 |
| Recoveries | C5 | AUC | 0.6528 |

TABLE VI. BEST MODEL–INDICATOR COMBINATIONS (HEALTH)

| Indicator | Learning Model | Metric | Value |
|---|---|---|---|
| Mortality | CHAID | F1-Score | 0.5667 |
| Mortality & Recoveries | KNN | Precision | 0.5254 |
| Cases | CHAID | Recall | 1.0000 |
| Mortality & Recoveries | KNN | Accuracy | 0.6420 |
| Mortality & Recoveries | KNN | AUC | 0.6155 |

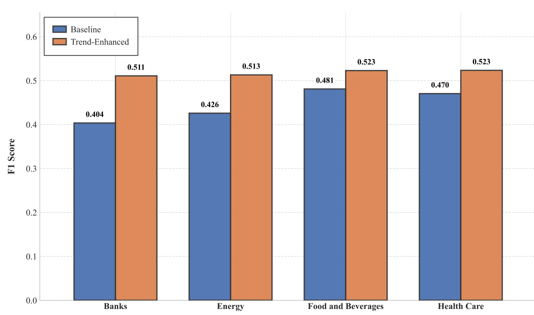


Fig. 2. Average F1-score comparison between Baseline and Trend-Enhanced configurations across sectors.

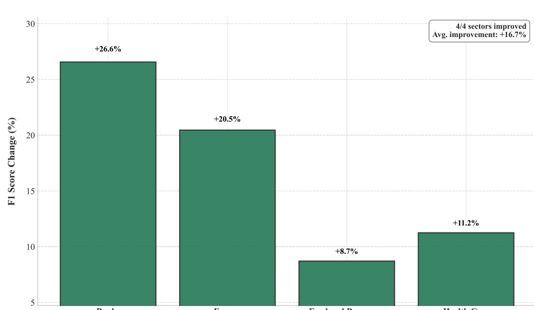


Fig. 3. Percentage change in F1-score induced by trend encoding across sectors. Error bars indicate variability across models.

To quantify the relative effect of trend encoding, Fig. 3 presents the percentage change in F1-score from baseline, together with variability across learning models.

The Banking sector exhibits the largest relative gain, indicating strong sensitivity to exogenous directional signals. Energy shows moderate but consistent improvement, while Health Care demonstrates smaller yet positive gains. In contrast, Food & Beverages displays weaker and less consistent improvements, suggesting limited alignment between short-term epidemiological trends and sector dynamics. These results directly address RQ1, demonstrating that trend encoding improves predictive performance in sectors where exogenous signals carry economically meaningful information.

To further contextualize these aggregate trends, Table III, Table IV, Table V, and Table VI summarize the top-performing combinations of learning models and COVID-19 indicators for each sector across multiple evaluation metrics, including F1-score, Precision, Recall, Accuracy, and AUC. Rather than reflecting a single dominant model, these results highlight how sector characteristics influence the effectiveness of specific model–indicator pairings.

In the *Banking* sector, models incorporating COVID-19 indicators generally outperform their baseline counterparts across most metrics. The CHAID model using combined *Cases & Mortality* achieves the highest F1-score, indicating an improved balance between Precision and Recall. While indicator-enhanced models dominate in sensitivity-related metrics, the baseline CRT model retains the highest AUC, suggesting that baseline configurations may still offer competitive discrimination capability in certain settings.

For the *Energy* sector, trend-encoded indicators yield consistent gains across evaluation criteria. The highest F1-score is achieved by the CHAID model using *Mortality & Recoveries*, while Precision and Accuracy peak under different indicator–model combinations. Similar to Banking, baseline models occasionally remain competitive in AUC, reflecting their robustness in capturing overall class separability despite lower pointwise predictive gains.

In the *Food & Beverages* sector, performance improvements from trend encoding are more heterogeneous. Although several indicator-enhanced models surpass baseline performance in F1-score, Precision, and Accuracy, the baseline configuration achieves the highest Recall, indicating strong sensitivity even without exogenous inputs. This mixed behavior aligns with the weaker aggregate gains observed in Fig. 3 and suggests sector-specific limitations in the predictive relevance of short-term epidemiological trends.

The *Health Care* sector exhibits consistent benefits from trend encoding across most metrics. Indicator-enhanced CHAID and KNN models achieve the highest F1-score, Accuracy, and AUC, with some configurations reaching perfect Recall. These results suggest strong alignment between epidemiological dynamics and sector behavior, reinforcing the effectiveness of context-aware modeling in this domain.

Overall, this sector-wise analysis complements the aggregate results by showing that while baseline models may occasionally achieve competitive discrimination performance, integrating trend-encoded exogenous indicators consistently improves balance-oriented metrics such as F1-score, Precision, and Recall across most sectors. This further supports the conclusion that trend encoding enhances predictive modeling when exogenous signals align with underlying sector dynamics.

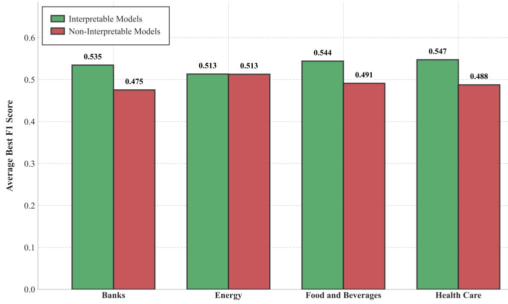


Fig. 4. Comparison of interpretable and non-interpretable models using trend-enhanced features. Values correspond to best configuration per model.

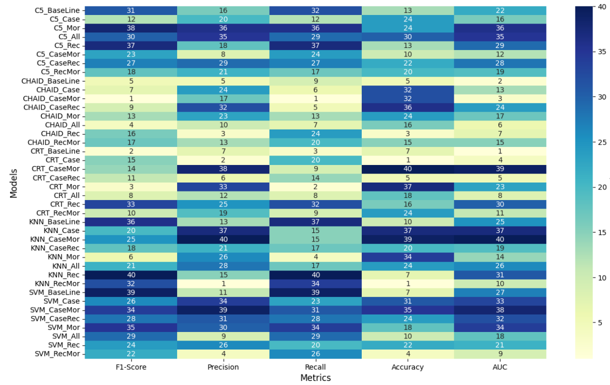


Fig. 5. Bank models rank.

### C. Interpretable vs. Non-Interpretable Models

To assess whether trend encoding preferentially benefits interpretable models, we group models into interpretable (C5, CHAID, CRT) and non-interpretable (KNN, SVM) categories and compare their best trend-enhanced F1-scores. Fig. 4 summarizes this comparison.

Across Banking, Food & Beverages, and Health Care, interpretable models consistently achieve higher F1-scores than their non-interpretable counterparts. In Energy, performance between the two groups is comparable, indicating that the benefit of interpretability is sector-dependent rather than universal.

This pattern provides strong empirical support for RQ2, indicating that trend-encoded signals align particularly well with the logical splitting mechanisms of interpretable models and enable them to close, or in some cases reverse, the traditional performance gap with black-box approaches.

*1) Model ranking across evaluation metrics:* To further examine relative model behavior beyond best-case configurations, we analyze model rankings aggregated across evaluation metrics. Fig. 5, Fig. 6, Fig. 7, and Fig. 8 present heatmap-based rankings for the Banking, Energy, Food & Beverages, and Health Care sectors, respectively.

Across all sectors, models incorporating trend-encoded COVID-19 indicators consistently occupy higher ranks than their baseline counterparts for most metrics, including F1-score, Precision, Recall, Accuracy, and AUC. This pattern indicates that the observed gains are not confined to isolated configurations but reflect a systematic improvement across evaluation criteria.

In the Banking and Energy sectors, indicator-enhanced interpretable models, particularly CHAID-based configurations, frequently achieve top-ranked positions for balance-oriented metrics such as F1-score and Recall. While baseline models occasionally remain competitive in AUC, their overall rankings are generally lower, suggesting reduced robustness across metrics. A similar trend is observed in the Food & Beverages and Health Care sectors, where indicator-enriched models dominate the upper ranks, whereas baseline configurations consistently appear in lower positions.

Taken together, these rankings reinforce the findings in Fig. 4 by demonstrating that interpretable models not only achieve strong peak performance but also maintain stable and competitive behavior across multiple evaluation dimensions when combined with trend-encoded exogenous signals.

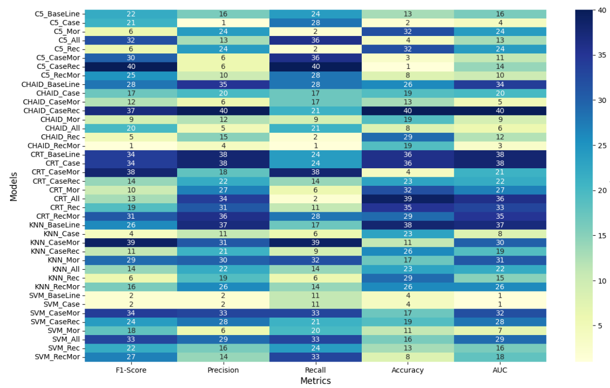


Fig. 6. Energy models rank.

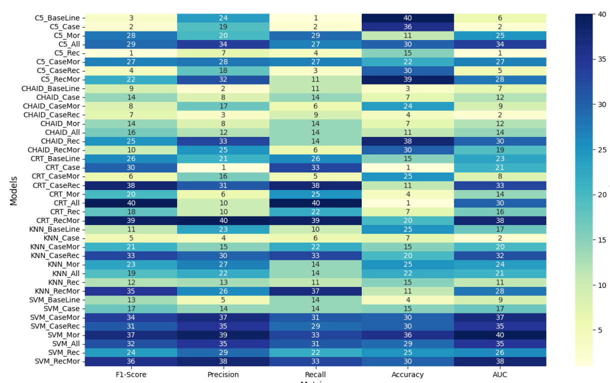


Fig. 7. Food models rank.

TABLE VII. PERFORMANCE METRICS OF BANKING-SECTOR MODELS

| Metric | Mean | Std Dev | Min | Median | Max |
|---|---|---|---|---|---|
| F1-Score | 0.426 | 0.084 | 0.228 | 0.429 | 0.588 |
| Precision | 0.462 | 0.029 | 0.394 | 0.460 | 0.527 |
| Recall | 0.420 | 0.148 | 0.150 | 0.398 | 0.796 |
| Accuracy | 0.546 | 0.027 | 0.472 | 0.547 | 0.588 |
| AUC | 0.529 | 0.024 | 0.471 | 0.526 | 0.577 |

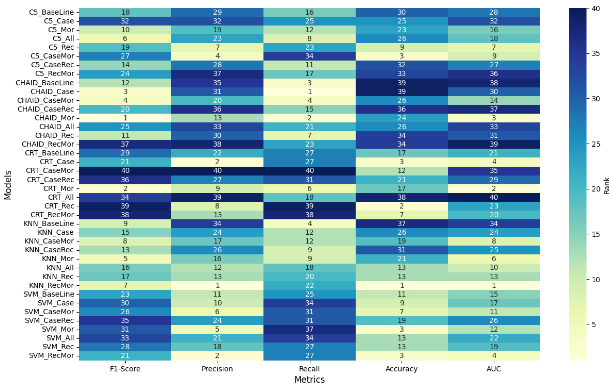


Fig. 8. Health models rank.

TABLE VIII. Performance Metrics of Energy-Sector Models

| Metric | Mean | Std Dev | Min | Median | Max |
|---|---|---|---|---|---|
| F1-Score | 0.438 | 0.059 | 0.281 | 0.450 | 0.549 |
| Precision | 0.374 | 0.033 | 0.288 | 0.379 | 0.436 |
| Recall | 0.554 | 0.145 | 0.216 | 0.527 | 0.838 |
| Accuracy | 0.483 | 0.057 | 0.343 | 0.485 | 0.586 |
| AUC | 0.498 | 0.041 | 0.378 | 0.507 | 0.566 |

TABLE IX. Performance Metrics of Food & Beverages-Sector Models

| Metric | Mean | Std Dev | Min | Median | Max |
|---|---|---|---|---|---|
| F1-Score | 0.416 | 0.100 | 0.178 | 0.434 | 0.571 |
| Precision | 0.402 | 0.049 | 0.294 | 0.413 | 0.476 |
| Recall | 0.465 | 0.193 | 0.111 | 0.472 | 0.944 |
| Accuracy | 0.594 | 0.038 | 0.509 | 0.602 | 0.657 |
| AUC | 0.562 | 0.047 | 0.479 | 0.566 | 0.653 |

TABLE X. Performance Metrics of Health-Sector Models

| Metric | Mean | Std Dev | Min | Median | Max |
|---|---|---|---|---|---|
| F1-Score | 0.454 | 0.073 | 0.186 | 0.454 | 0.567 |
| Precision | 0.408 | 0.041 | 0.320 | 0.410 | 0.525 |
| Recall | 0.552 | 0.178 | 0.131 | 0.533 | 1.000 |
| Accuracy | 0.523 | 0.064 | 0.389 | 0.534 | 0.642 |
| AUC | 0.529 | 0.038 | 0.427 | 0.528 | 0.615 |

### D. Statistical Summary Across Performance Metrics

Given the large number of evaluated models, indicator configurations, and industry sectors, a detailed presentation of all individual results would be impractical. We therefore provide a statistical summary of model performance to assess robustness and consistency across evaluation metrics. This analysis complements the aggregate comparisons presented earlier by consolidating model behavior across F1-score, Precision, Recall, Accuracy, and AUC.

Table VII, Table VIII, Table IX, and Table X report statistical summaries for the Banking, Energy, Food & Beverages, and Health Care Equipment & Services sectors, respectively. For each metric, descriptive statistics including mean, standard deviation, minimum, maximum, median, and 25th percentile are presented. The summaries include both baseline configurations and models incorporating trend-encoded COVID-19 indicators, enabling an assessment of overall performance stability with and without exogenous inputs.

TABLE XI. Performance Metrics of Banking-Sector Learning Models by Indicator.

| Indicator | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| All Indicators | 0.4422 | 0.4630 | 0.4372 | 0.5491 | 0.5342 |
| Baseline (No Indicators) | 0.4037 | 0.4822 | 0.3788 | **0.5678** | **0.5426** |
| Cases | **0.4506** | 0.4493 | 0.4678 | 0.5308 | 0.5225 |
| Mortality | 0.4339 | 0.4450 | 0.4566 | 0.5311 | 0.5212 |
| Recoveries | 0.3888 | 0.4778 | 0.3434 | 0.5618 | 0.5327 |
| Cases & Mortality | 0.4472 | 0.4315 | **0.4920** | 0.5124 | 0.5097 |
| Cases & Recoveries | 0.4452 | 0.4583 | 0.4425 | 0.5423 | 0.5290 |
| Mortality & Recoveries | 0.4285 | **0.4848** | 0.3982 | 0.5618 | 0.5400 |

In the *Banking* sector, models exhibit moderate average performance across both indicator-enhanced and baseline configurations. The mean F1-score reflects a balance between Precision and Recall, while the relatively high standard deviation in Recall indicates variability in sensitivity to upward stock movements across indicator combinations. Average Accuracy and AUC values suggest performance slightly above random classification, with several configurations performing comparably to the baseline.

For the *Energy* sector, performance statistics reveal a more challenging prediction environment. While average Recall values indicate sensitivity to positive movements under certain indicator configurations, lower Precision and AUC values reflect a higher incidence of false positives and reduced discriminative capability. Overall, the statistical summaries indicate that improvements over the baseline are less consistent in this sector.

Models applied to the *Food & Beverages* sector demonstrate comparatively stable and consistent behavior. Average F1-score, Precision, and Recall values indicate balanced predictions, while high maximum Recall values suggest that some configurations effectively capture strong demand-driven movements. The corresponding Accuracy and AUC statistics further indicate improved separation between upward and non-upward movements relative to baseline performance.

In the *Health Care* sector, statistical summaries show moderate average performance with notable variability across configurations. Mean F1-score values reflect a trade-off between Precision and Recall, while the occurrence of perfect Recall in certain configurations indicates strong alignment between epidemiological indicators and sector behavior. However, the observed variability suggests that these gains are configuration-dependent rather than uniform across all models.

Overall, the statistical summaries indicate that incorporating trend-encoded exogenous indicators generally enhances balance-oriented metrics such as F1-score and Recall, while baseline models may remain competitive in certain discrimination-oriented measures. These findings reinforce the importance of examining performance distributions alongside aggregate results when evaluating the robustness of trend-aware predictive models.

### E. Exogenous Indicator Relevance and Sector-Specific Effects

We next analyze which exogenous indicators contribute most to performance improvements. Fig. 9 shows the average

TABLE XII. PERFORMANCE METRICS OF ENERGY-SECTOR LEARNING MODELS BY INDICATOR.

| Indicator | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| All Indicators | 0.4319 | 0.3746 | 0.5351 | 0.4869 | 0.4966 |
| Baseline (No Indicators) | 0.4259 | 0.3560 | 0.5351 | 0.4606 | 0.4756 |
| Cases | 0.4193 | 0.3756 | 0.5009 | 0.4976 | 0.4983 |
| Mortality | **0.4672** | 0.3779 | **0.6270** | 0.4747 | 0.5054 |
| Recoveries | 0.4614 | 0.3765 | 0.6162 | 0.4758 | 0.5041 |
| Cases & Mortality | 0.3893 | 0.3768 | 0.4108 | **0.5273** | 0.5038 |
| Cases & Recoveries | 0.4063 | 0.3618 | 0.5081 | 0.4707 | 0.4782 |
| Mortality & Recoveries | 0.4439 | **0.3784** | 0.5568 | 0.4949 | **0.5074** |

TABLE XIII. PERFORMANCE METRICS OF FOOD & BEVERAGES-SECTOR LEARNING MODELS BY INDICATOR.

| Indicator | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| All Indicators | 0.3474 | 0.3921 | 0.3444 | 0.5981 | 0.5347 |
| Baseline (No Indicators) | **0.4810** | **0.4300** | **0.5833** | 0.5963 | **0.5931** |
| Cases | 0.4336 | 0.4097 | 0.4981 | 0.5975 | 0.5727 |
| Mortality | 0.3957 | 0.4010 | 0.3944 | **0.6019** | 0.5500 |
| Recoveries | 0.4010 | 0.3816 | 0.4500 | 0.5824 | 0.5493 |
| Cases & Mortality | 0.4309 | 0.3929 | 0.4944 | 0.5833 | 0.5611 |
| Cases & Recoveries | 0.3982 | 0.3914 | 0.4500 | 0.5944 | 0.5583 |
| Mortality & Recoveries | 0.3408 | 0.3506 | 0.3667 | 0.5722 | 0.5208 |

TABLE XIV. PERFORMANCE METRICS OF HEALTH-SECTOR LEARNING MODELS BY INDICATOR.

| Indicator | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| All Indicators | 0.4522 | 0.3888 | 0.5541 | 0.4988 | 0.5097 |
| Baseline (No Indicators) | 0.4705 | 0.3933 | 0.6230 | 0.4840 | 0.5115 |
| Cases | 0.4464 | 0.4020 | 0.5454 | 0.5210 | 0.5258 |
| Mortality | **0.5147** | **0.4326** | **0.6623** | 0.5432 | **0.5668** |
| Recoveries | 0.4296 | 0.4226 | 0.4689 | 0.5475 | 0.5320 |
| Cases & Mortality | 0.4206 | 0.4114 | 0.4787 | **0.5556** | 0.5403 |
| Cases & Recoveries | 0.4529 | 0.3846 | 0.5672 | 0.4938 | 0.5084 |
| Mortality & Recoveries | 0.4308 | 0.4219 | 0.4590 | 0.5457 | 0.5285 |

TABLE XV. OPTIMAL MODEL AND INDICATOR CONFIGURATION PER SECTOR BASED ON HIGHEST F1-SCORE.

| Sector | Best Model | Optimal Indicator(s) | F1-Score |
|---|---|---|---|
| Banks | CHAID | Cases + Mortality | 0.588 |
| Energy | CHAID | Recoveries + Mortality | 0.549 |
| Food & Beverages | C5 | Recoveries | 0.571 |
| Health Care | CHAID | Mortality | 0.567 |

F1-score achieved when each indicator (or indicator combination) is included, aggregated by sector.

Distinct sectoral patterns emerge. Banking benefits most from combined severity signals (Cases and Mortality), reflecting sensitivity to systemic risk conditions. Energy responds more strongly to outcome-oriented indicators such as Recoveries and Mortality, consistent with demand normalization effects. Food & Beverages exhibits weaker overall sensitivity to exogenous indicators, while Health Care performance is primarily influenced by Mortality trends (see Table XI to Table XIV).

Table XV summarizes the optimal model–indicator combinations per sector based on the highest achieved F1-score.

To further assess indicator relevance independently of individual learning algorithms, we examined performance distributions aggregated across models for each indicator configuration.

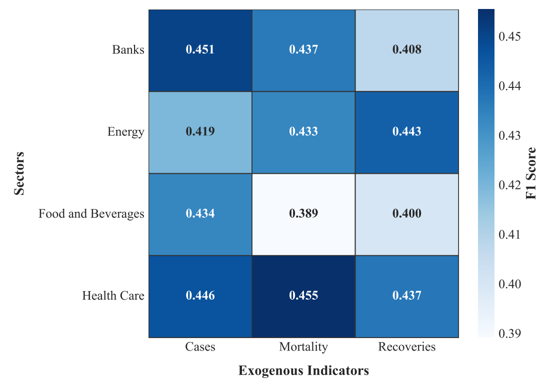


Fig. 9. Indicator relevance across sectors based on average F1-score performance.

TABLE XVI. PERFORMANCE METRICS OF BANKING-SECTOR LEARNING MODELS.

| Learning Model | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| C5 | 0.3851 | 0.4557 | 0.3418 | 0.5492 | 0.5216 |
| CHAID | **0.5023** | **0.4723** | **0.5542** | 0.5463 | **0.5474** |
| CRT | 0.4825 | 0.4687 | 0.5166 | 0.5435 | 0.5400 |
| KNN | 0.3901 | 0.4561 | 0.3706 | 0.5389 | 0.5165 |
| SVM | 0.3705 | 0.4558 | 0.3186 | **0.5496** | 0.5189 |

This analysis confirms that indicators related to reported *Cases* and *Mortality* tend to produce higher balance-oriented metrics such as F1-score and Recall across most sectors, while *Recoveries*-based indicators exhibit more variable effects.

In several sectors, baseline configurations without exogenous indicators remain competitive for specific metrics, particularly Precision or Recall. This behavior is most evident in the Food & Beverages sector, where models achieve strong predictive performance even in the absence of COVID-19 indicators, suggesting weaker short-term coupling between epidemiological trends and market dynamics. In contrast, Banking, Energy, and Health Care sectors demonstrate more consistent benefits from indicator-enhanced configurations, although the magnitude of improvement varies across indicators and metrics.

Overall, these findings indicate that the predictive value of exogenous indicators is neither uniform across sectors nor independent of evaluation criteria. While certain indicators provide informative signals under specific conditions, indiscriminate inclusion of all external variables may dilute performance. These results directly address RQ3, demonstrating that exogenous indicator relevance is inherently sector-specific and must be carefully aligned with both market characteristics and modeling objectives.

To complement the indicator-centric analysis, we examine how different learning models respond to variations in COVID-19 indicator configurations. This model-wise perspective evaluates whether performance gains from exogenous signals are consistent across interpretable and non-interpretable algorithms, and whether specific model families exhibit stronger sensitivity to indicator-driven trends.

Table XVI, Table XVII, Table XVIII, and Table XIX sum-

TABLE XVII. PERFORMANCE METRICS OF ENERGY-SECTOR LEARNING MODELS.

| Learning Model | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| C5 | 0.4193 | **0.3937** | 0.4865 | **0.5227** | 0.5154 |
| CHAID | **0.4633** | 0.3742 | **0.6182** | 0.4710 | 0.5007 |
| CRT | 0.4243 | 0.3464 | 0.5676 | 0.4369 | 0.4632 |
| KNN | 0.4470 | 0.3663 | 0.5912 | 0.4672 | 0.4922 |
| SVM | 0.4371 | 0.3870 | 0.5068 | 0.5189 | **0.5165** |

TABLE XVIII. PERFORMANCE METRICS OF FOOD & BEVERAGES-SECTOR LEARNING MODELS.

| Learning Model | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| C5 | 0.4718 | 0.3976 | **0.6285** | 0.5648 | 0.5807 |
| CHAID | **0.4752** | **0.4303** | 0.5382 | 0.6053 | **0.5885** |
| CRT | 0.3477 | 0.4142 | 0.3368 | **0.6238** | 0.5521 |
| KNN | 0.4254 | 0.4087 | 0.4583 | 0.6019 | 0.5660 |
| SVM | 0.3592 | 0.3589 | 0.3611 | 0.5752 | 0.5217 |

TABLE XIX. PERFORMANCE METRICS OF HEALTH-SECTOR MODELS

| Learning Model | F1-Score | Precision | Recall | Accuracy | AUC |
|---|---|---|---|---|---|
| C5 | 0.4666 | 0.4038 | 0.5697 | 0.5154 | 0.5262 |
| CHAID | 0.4943 | 0.3799 | **0.7254** | 0.4537 | 0.5075 |
| CRT | 0.3821 | 0.4017 | 0.4057 | 0.5471 | 0.5191 |
| KNN | **0.5010** | 0.4239 | 0.6311 | 0.5270 | **0.5476** |
| SVM | 0.4285 | **0.4317** | 0.4262 | **0.5718** | 0.5429 |

marize performance statistics across learning models for the Banking, Energy, Food & Beverages, and Health Care sectors, respectively. The evaluated models include interpretable approaches (C5, CHAID, CRT) and non-interpretable approaches (KNN, SVM), assessed using F1-score, Precision, Recall, Accuracy, and AUC.

Across sectors, interpretable models—particularly CHAID—frequently achieve strong balance-oriented performance, reflected in higher F1-scores and Recall values under indicator-enhanced configurations. This behavior is most pronounced in relatively stable sectors such as Banking and Food & Beverages, where rule-based splitting mechanisms appear effective in leveraging trend-encoded exogenous signals.

In contrast, non-interpretable models exhibit comparatively stronger generalization behavior in more volatile sectors. In Energy and Health Care, distance- and margin-based models such as KNN and SVM often achieve competitive Accuracy and AUC values, suggesting improved discrimination capability when market dynamics are less structured. However, these gains are sometimes accompanied by reduced Recall, indicating more conservative prediction behavior.

Overall, this analysis indicates that model sensitivity to exogenous indicators is closely tied to both sector characteristics and model architecture. Interpretable models tend to benefit more consistently from trend encoding in sectors with stable structural patterns, while non-interpretable models demonstrate resilience in complex or highly variable environments. These findings reinforce the importance of aligning indicator selection with both model transparency objectives and sector-specific dynamics.

## VI. CONCLUSION

This study introduced a trend-encoding framework for integrating heterogeneous exogenous signals into financial prediction models, addressing challenges related to scale mismatch, semantic ambiguity, and temporal alignment. By transforming raw epidemiological time series into directional trend representations, the proposed approach improves predictive robustness while preserving interpretability. Experimental results across multiple market sectors demonstrate that interpretable models benefit substantially from trend-encoded features, often achieving performance comparable to more complex black-box methods. The findings further highlight that representation design can be as influential as model complexity in cross-domain financial prediction and that the relevance of exogenous signals is inherently sector-dependent. Overall, the proposed framework offers a practical and transparent strategy for incorporating external information into financial forecasting systems where explainability and robustness are essential.

## REFERENCES


[1] S. A. Al-Thaqeb and B. G. Algharabali, "Economic policy uncertainty: A literature review," *The Journal of Economic Asymmetries*, vol. 20, p. e00133, 2019.

[2] K. Arias-Calluari, F. Alonso-Marroquin, M. N. Najafi, and M. Harré, "Methods for forecasting the effect of exogenous risks on stock markets," *Physica A: Statistical Mechanics and its Applications*, vol. 568, p. 125587, 2021.

[3] Y. Jiang, K. Ning, Z. Pan, X. Shen, J. Ni, W. Yu, A. Schneider, H. Chen, Y. Nevmyvaka, and D. Song, "Multi-modal time series analysis: A tutorial and survey," in *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, 2025, pp. 6043–6053.

[4] M. Darwish, E. E. Hassanien, and A. H. Eissa, "Stock market forecasting: From traditional predictive models to large language models: M. darwish et al." *Computational Economics*, pp. 1–45, 2025.

[5] P.-D. Arsenault, S. Wang, and J.-M. Patenaude, "A survey of explainable artificial intelligence (xai) in financial time series forecasting," *ACM Computing Surveys*, vol. 57, no. 10, pp. 1–37, 2025.

[6] D. Zhang, M. Hu, and Q. Ji, "Financial markets under the global pandemic of covid-19," *Finance research letters*, vol. 36, p. 101528, 2020.

[7] D. Štifanić, J. Musulin, A. Miočević, S. Baressi Šegota, R. Šubić, and Z. Car, "Impact of covid-19 on forecasting stock prices: An integration of stationary wavelet transform and bidirectional long short-term memory," *Complexity*, vol. 2020, no. 1, p. 1846926, 2020.

[8] C. Yuan, X. Ma, H. Wang, C. Zhang, and X. Li, "Covid19-mlsf: A multi-task learning-based stock market forecasting framework during the covid-19 pandemic," *Expert Systems with Applications*, vol. 217, 5 2023.

[9] A. Almehmadi, "Covid-19 pandemic data predict the stock market," *Computer Systems Science and Engineering*, vol. 36, pp. 451–460, 2021.

[10] C. Yang, M. Z. Abedin, H. Zhang, F. Weng, and P. Hajek, "An interpretable system for predicting the impact of covid-19 government interventions on stock market sectors," *Annals of Operations Research*, pp. 1–28, 2023.

[11] D. Wu, X. Wang, J. Su, B. Tang, and S. Wu, "A labeling method for financial time series prediction based on trends," *Entropy*, vol. 22, no. 10, p. 1162, 2020.

[12] A. Fraikin, A. Bennetot, and S. Allassonnière, "T-rep: Representation learning for time series using time-embeddings," *arXiv preprint arXiv:2310.04486*, 2023.

[13] M. Bańbura, D. Giannone, M. Modugno, and L. Reichlin, "Now-casting and the real-time data flow," in *Handbook of economic forecasting*. Elsevier, 2013, vol. 2, pp. 195–237.

[14] D. Giannone, L. Reichlin, and D. Small, "Nowcasting: The real-time informational content of macroeconomic data," *Journal of monetary economics*, vol. 55, no. 4, pp. 665–676, 2008.

[15] P. R. Jena and R. Majhi, "Are twitter sentiments during covid-19 pandemic a critical determinant to predict stock market movements? a machine learning approach," *Scientific African*, vol. 19, 3 2023.

[16] N. Bussmann, P. Giudici, D. Marinelli, and J. Papenbrock, "Explainable ai in fintech risk management," *Frontiers in Artificial Intelligence*, vol. 3, p. 26, 2020.

[17] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature machine intelligence*, vol. 1, no. 5, pp. 206–215, 2019.

[18] A. A. Freitas, "Comprehensible classification models: a position paper," *ACM SIGKDD explorations newsletter*, vol. 15, no. 1, pp. 1–10, 2014.

[19] S. Gu, B. Kelly, and D. Xiu, "Empirical asset pricing via machine learning," *The Review of Financial Studies*, vol. 33, no. 5, pp. 2223–2273, 2020.

[20] H. Goel and B. K. Som, "Stock market prediction, covid-19 pandemic and neural networks: an scg algorithm application," *EconomiA*, vol. 24, pp. 134–146, 7 2023.

[21] R. Chandra and Y. He, "Bayesian neural networks for stock price forecasting before and during covid-19 pandemic," *PLoS ONE*, vol. 16, 7 2021.

[22] salwaalzahrani, "Saudi Stock Exchange (Tadawul) Dataset," https://www.kaggle.com/datasets/salwaalzahrani/saudi-stock-exchange-tadawul, 2023, kaggle dataset.

[23] King Abdullah Petroleum Studies and Research Center (KAPSARC), "Saudi Arabia Coronavirus Disease (COVID-19) Situation," https://datasource.kapsarc.org/explore/dataset/saudi-arabia-coronavirus-disease-covid-19-situation/information/, 2020, dataset.