TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Xây dựng đặc trưng ngành cho hệ thống dự báo xu hướng giá cổ phiếu trên thị trường chứng khoán Việt Nam

Cao Tấn Bình[1], **Võ Đoàn Xuân Tiên**[1,*], **Hà Thúy Phương**[1], **Nguyễn Hữu Thịnh**[1], **Trương Chấn Hào**[1], **Trương Tuấn Kiệt**[1]

[1] *Khoa Toán và Thống kê, Trường Đại học Quy Nhơn*

**Tác giả liên hệ chính. Email: tienvdx760995@gmail.com*

Ngày nhận bài: .../.../.........; Ngày sửa bài: .../.../........;
Ngày nhận đăng: .../.../.......; Ngày xuất bản: .../.../.........

TÓM TẮT

Nghiên cứu này đề xuất một hệ thống dự báo tích hợp dành cho nhóm cổ phiếu ngân hàng tại thị trường Việt Nam, giải quyết thách thức về biến động mạnh (high volatility) và sự bất đồng bộ giữa dữ liệu giao dịch với các yếu tố cơ bản. Thay vì chỉ dựa trên lịch sử giá, hệ thống kết hợp đặc trưng kỹ thuật, cơ bản và đặc thù ngành ngân hàng (như NIM, NPL, tăng trưởng tín dụng), đồng thời hợp nhất dữ liệu đa tần suất để xử lý triệt để độ trễ thông tin. Biến mục tiêu được xác định thông qua cơ chế Lọc nhiễu Thích ứng Bất đối xứng (Asymmetric Adaptive Thresholding) thay vì các ngưỡng cố định, giúp cân bằng lớp dữ liệu và thích nghi với tính phương sai thay đổi (heteroscedasticity) của thị trường. Về mặt mô hình, nghiên cứu tiến hành khảo sát thực nghiệm trên hai luồng kiến trúc mô hình học sâu chuyên biệt: (1) mô hình lai ghép CNN-BiLSTM và (2) mô hình Transformer Encoder. Cả hai được triển khai trên tập dữ liệu của 8 ngân hàng thương mại lớn tại Việt Nam giai đoạn 2010–2025, với quy trình tiền xử lý nghiêm ngặt bao gồm hợp nhất dữ liệu đa tần suất (backward merge-as-of) và chuẩn hóa dữ liệu. Kết quả cho thấy khả năng dự báo ngắn hạn (T+1) còn hạn chế (Balanced Accuracy ∼50-61%). Tuy nhiên, hiệu suất được cải thiện rõ rệt ở các khung trung và dài hạn (đặc biệt T+60, T+90), với độ chính xác cân bằng đạt từ 80–92% và F1-Score lên đến 93.25%. So với các nghiên cứu trước và chiến lược nắm giữ (Buy & Hold), kết quả này khẳng định vai trò quan trọng của tri thức ngành và cơ chế ngưỡng thích ứng.

Từ khóa: *Stock price prediction, Deep learning, Industry-specific features, Transformer, CNN-BiLSTM.*

QUY NHON UNIVERSITY
JOURNAL OF SCIENCE

CONSTRUCTING INDUSTRY-SPECIFIC FEATURES FOR STOCK TREND PREDICTION SYSTEMS IN THE VIETNAMESE STOCK MARKET

Cao Tấn Bình[1], Võ Đoàn Xuân Tiên[1,*], Hà Thúy Phương[1], Nguyễn Hữu Thịnh[1], Trương Chấn Hào[1], Trương Tuấn Kiệt[1]

[1] *Department of Mathematics and Statistics, Quy Nhon University*

[*] *Corresponding author. Email: tienvdx760995@gmail.com*

Received: .../.../........; Published: .../.../........;
Accepted: .../.../......; Published: .../.../........

ABSTRACT

This study proposes an integrated forecasting system for the banking stock group in the Vietnamese market, addressing challenges of high volatility and the asynchrony between trading data and fundamental factors. Instead of relying solely on price history, the system combines technical features, fundamental data, and banking-specific metrics (such as NIM, NPL, and credit growth), while unifying multi-frequency data to rigorously address information latency. The target variable is defined using an Asymmetric Adaptive Thresholding mechanism rather than fixed thresholds, helping to balance data classes and adapt to the market's heteroscedasticity. Regarding modeling, the study empirically investigates two specialized deep learning architectures: (1) a hybrid CNN-BiLSTM model and (2) a Transformer Encoder model. Both are implemented on a dataset of 8 major commercial banks in Vietnam from 2010 to 2025, following a rigorous preprocessing pipeline that includes multi-frequency data fusion (backward merge-as-of) and data normalization. Results indicate that short-term prediction (T+1) remains limited due to high stochasticity (Balanced Accuracy ∼50-61%).. However, performance significantly improves in medium- and long-term horizons (notably T+60, T+90), with balanced accuracy reaching 80–92% and an F1-Score up to 93.25%. Compared to prior works and the Buy & Hold strategy, these results confirm the vital role of domain knowledge and adaptive thresholding mechanisms.

Keyworks: *Stock price prediction, Deep learning, Industry-specific features, Transformer, CNN-BiLSTM.*

1. INTRODUCTION

Over the past decade, the Vietnamese stock market has established itself as one of the most dynamic frontier and emerging markets in the region, serving as a vital capital channel for the economy.[1] However, this market is characterized by high volatility and significant dominance by individual investors, who account for 85–90% of trading liquidity. The prevalence of retail cash flow often leads to short-term price fluctuations driven by crowd psychology rather than corporate intrinsic factors. In this context, traditional forecasting methods or quantitative models relying purely on historical prices often struggle to distinguish between signal and noise, resulting in suboptimal risk management.[2, 3]

To address the forecasting challenge in Vietnam, the selection of study subjects plays a pivotal role. The banking sector, with a capitalization weight of approximately 30–35% of the entire market, is considered the leading group and a key indicator reflecting the health of the macroeconomy as well as the impact of monetary policy.[4] Nevertheless, integrating the fundamental factors of this industry into prediction models faces a major technical challenge regarding data asynchrony: while stock prices fluctuate in real-time (high-frequency), critical financial ratios such as Net Interest Margin (NIM) or Non-Performing Loans (NPL) are only published quarterly (low-frequency) with significant lag. Many previous studies have often addressed this issue using simple interpolation methods, inadvertently causing look-ahead bias and distorting actual predictive capabilities.

The development of deep learning architectures such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) has opened new avenues for financial time series analysis. However, most current studies still suffer from two main limitations. First, models often mechanically apply past price data while ignoring industry-specific structural features—the core drivers of long-term trends. Second, the determination of uptrends/downtrends is often based on fixed thresholds (e.g., $\pm 1\%$), failing to reflect the time-varying volatility (heteroscedasticity) of each stock, leading to inaccurate data labeling.

To overcome these limitations, this study proposes an Integrated Hybrid Forecasting Framework combining domain-specific knowledge and advanced deep learning techniques.[5] Specifically, we: (1) Construct a Multi-frequency Data Fusion pipeline to rigorously handle information latency, combining trading data with banking health indicators such as valuation, asset quality, and liquidity; (2) Apply an Adaptive Thresholding mechanism to define target variables flexibly according to historical volatility;[6] and (3) Conduct a comparative analysis of the effectiveness between CNN-BiLSTM and Transformer architectures on a dataset of 8 major commercial banks in Vietnam from 2010 to 2025. The research results not only demonstrate the trend-capturing capability of the Transformer architecture in medium- and long-term horizons but also provide empirical evidence for a Parallel Ensemble Strategy to support quantitative investment decision-making.

2. MODEL ARCHITECTURE

2.1. Model Architecture

The study deploys two specialized deep learning architectures to maximize the exploitation of financial time series data characteristics:

2.1.1. Hybrid CNN-BiLSTM Model

This architecture is built upon the principle of functional decomposition, combining the strengths of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks.[7, 8]

Local Feature Extraction: The 1-dimensional CNN (1D-CNN) layer acts as sliding filters to detect candlestick patterns or short-term price fluctuations.[9] The computation for the j-th feature at layer l is:

$$x_j^l = f\left(\text{conv1D}\left(\sum_i (x_i^{l-1} * w_{ij}^l + b_j^l)\right)\right) \quad (1)$$

where $*$ denotes the convolution operator, w is the filter weight, b is the bias, and f is the ReLU activation function.

Sequential Modeling: The feature sequence, after noise filtering, is fed into the Bi-LSTM layer. At each step t, the hidden state h_t is synthesized from both directions:[10]

$$\vec{h}_t = \text{LSTM}_{fwd}(x_t, \vec{h}_{t-1}) \quad (2)$$

$$\overleftarrow{h}_t = \text{LSTM}_{bwd}(x_t, \overleftarrow{h}_{t+1}) \quad (3)$$

$$h_t = W_{\vec{h}}\vec{h}_t + W_{\overleftarrow{h}}\overleftarrow{h}_t + b_h \quad (4)$$

This allows the model to capture the comprehensive context within the observation window.

2.1.2. Transformer Encoder Model

The Transformer architecture completely eliminates recurrence mechanisms to overcome the information forgetting problem in long sequences, replacing them with Self-Attention mechanisms.[11]

Positional Encoding: Since processing is not sequential, temporal order information is injected into the model via positional encoding vectors added directly to the Embeddings.

Scaled Dot-Product Attention: Allows the model to compute the importance weight of each past time step relative to the present:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

Where Q (Query), K (Key), and V (Value) are projection matrices of the input data.

Multi-Head Attention: The model employs multiple parallel attention heads to learn multi-dimensional relationships (e.g., one head focuses on price volatility, another on macroeconomic indicators):

$$\text{MultiHead}(Q,K,V) = \text{Concat}(\text{head}_1, ..., \text{head}_h)W^O \tag{6}$$

This architecture is particularly effective in detecting long-range dependencies in distant forecasting horizons (T+60, T+90). To further clarify the operating mechanism, from local feature extraction and context capturing (CNN-BiLSTM) to learning long-term relationships via attention mechanisms (Transformer), the following figure presents the overall structure of both architectures:

-

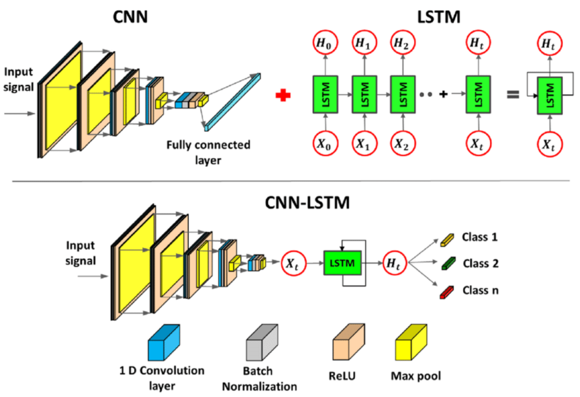

Figure 1. Overall architecture diagram of the hybrid CNN-LSTM model

3. MATERIALS AND METHODS

To realize the proposed Integrated Hybrid Framework, the study approaches the problem with a data-centric orientation. We identify that the biggest barrier hindering current forecasting performance lies not only in algorithmic limitations but fundamentally in the scarcity of high-information-content input (Input Scarcity) and serious asynchrony between data streams.

Therefore, this section details the construction of a Comprehensive End-to-End Feature Engineering Pipeline. This pipeline is designed as a **methodological basis** to rigorously address the asynchrony between high-frequency trading data and low-frequency financial data, while eliminating look-ahead bias risks. Based on the standardized and enriched data, the study configures and trains specialized deep learning models to verify the hypotheses set forth.

3.1. Data and Modeling Context

The subjects of the study are 8 listed commercial banks with market capitalization and liquidity among the top tier on the Vietnamese stock market: VCB, BID, CTG, TCB, MBB, VPB, ACB, HDB. This selection is based on three criteria: (1) high capitalization and liquidity ensuring feasibility for real-world strategy deployment; (2) leading role in sector indices, creating price spillover effects to the entire market; and (3) availability and transparency of financial report data.

The database includes:

- **Market Data (OHLCV):** All trading sessions from 01/2010 to 12/2025.[12]
- **Quarterly Financial Reports:** Key indicators (NIM, NPL, LDR, CASA, ROE, ROA, P/B...) extracted directly from officially published reports.[13]
- **Macroeconomic Data:** Policy rates, interbank rates, exchange rates, used to check model stability.

Data splitting technique adheres to a Strict Chronological Split:

- **Training Set (01/2010–12/2022):** ∼13 years, covering phases of crisis, growth, and recovery.
- **Validation Set (01/2023–06/2024):** ∼18 months, used for hyperparameter tuning.

- **Testing Set (07/2024–12/2025):** Completely new data phase (Out-of-sample) to evaluate real-world performance.

3.2. Multi-level Feature Engineering Technique

The study approaches the problem with systematic layering thinking, transforming raw data into optimal input tensors through three intensive processing steps:[14]

3.2.1. Feature Space Construction

The input feature space is designed with multi-dimensions comprising 30 variables, strictly categorized into two functional groups to capture both market behavior and intrinsic value:

Technical Universe: Calculated entirely from past price and volume data, this group plays the role of capturing Market Sentiment and short-term cash flow.[15, 16]

- *Momentum Indicators:* Using RSI (14) and Stochastic Oscillator to detect local overbought/oversold zones; MACD Histogram to measure trend intensity and divergence.
- *Trend Indicators:* EMA system (7, 21, 50) helps smooth price noise; Bollinger Bands provide information on relative volatility and breakout points.
- *Liquidity and Volatility:* ATR (Average True Range) measures absolute fluctuation range, assisting in determining intraday risk.

Fundamental & Banking-Specific Universe: This is a critical contribution of the study, integrating specific KPI indicators reflecting the financial health ("value anchors") of credit institutions. Unlike manufacturing enterprises, banks are evaluated via the quantified CAMELS set of indicators:[17]

- *Profitability:* NIM (Net Interest Margin) - Measure of the difference between interest income and interest expense, reflecting core capital usage efficiency. ROE (Return on Equity) reflects profitability performance for shareholders.
- *Asset Quality:* NPL (Non-Performing Loan Ratio). This is a Leading Indicator; an increase in NPL often signals increased provisioning, eroding future profits and exerting downward pressure on stock prices.
- *Liquidity and Capital Adequacy:* LDR (Loan-to-Deposit Ratio) measures liquidity stress levels; CASA (Current Account Savings Account) reflects cheap capital cost, creating long-term competitive advantage.
- *Hybrid Valuation:* P/B and P/E ratios are calculated in real-time (Real-time Price / Lagged Earnings), helping the model identify relative expensive/cheap valuation states compared to history.

3.2.2. Multi-frequency Data Fusion (Temporal Alignment)

The biggest challenge when integrating the "Feature Universe" is the Frequency Mismatch between price (daily) and fundamental indicators (quarterly). Linear Interpolation often leads to data distortion. The study solves this problem using the Backward Merge-As-Of algorithm.

- *Operating Mechanism:* At each trading session t, the value of fundamental variables F_t (such as NIM, NPL) is retrieved from the most recent quarterly financial report that was actually publicly available at time t_{pub} such that $t_{pub} \leq t$.
- *Handling Latency:* The study introduces a lag parameter Δ_{lag} (typically 15-30 days after the quarter ends) to accurately simulate the time information reaches investors.
- *Significance:* This mechanism ensures Data Integrity and completely eliminates look-ahead bias, keeping empirical results absolutely faithful to the actual investment context.[18]

3.2.3. Robust Normalization

Vietnamese financial data is often non-Gaussian and contains Fat-tails due to market shocks (Flash crashes, arrest rumors). Traditional Z-score normalization (based on mean and variance) is very sensitive and easily distorted by these outliers. The study applies Robust Scaling based on position statistics:[19]

$$X_{scaled} = \frac{X - Q_2(X)}{Q_3(X) - Q_1(X)} \tag{7}$$

Where Q_1, Q_2, Q_3 are the 25th, 50th (median), and 75th percentiles respectively. This method brings data to a standard scale while preserving the original distribution structure, helping the model learn more stable patterns.

3.3. Adaptive Thresholding and Asymmetric Thresholds

Traditional labeling methods using fixed thresholds often err by assuming price volatility is uniform across stocks and market phases. To overcome this, the study applies Adaptive Thresholding.[6] The core principle is that the trend distinction threshold must automatically adapt to the specific volatility σ_i of each stock i and forecast horizon H. This threshold is determined by the formula:

$$Threshold_{i,h} = k \cdot \sigma_i \cdot \sqrt{h} \tag{8}$$

Where:

- σ_i: Standard deviation of daily returns of stock i (calculated on the training set).
- $\sqrt{H}$: Time multiplier according to diffusion law.
- k: Sensitivity adjustment coefficient.

Due to the long-term growth characteristic of the Vietnamese market, the study applies asymmetric coefficient k to balance samples:

$$\begin{cases} k_{up} = 0.55 \\ k_{down} = 0.45 \end{cases} \tag{9}$$

Choosing $k_{down} < k_{up}$ helps the model be more sensitive in detecting early downtrend signals, which often occur faster and more severely than uptrends.

3.4. Experimental Setup

To ensure scientific transparency and allow the research community to reproduce results, we publish detailed optimal hyperparameters. These parameters are determined through a rigorous Grid Search process on the Validation set, aiming to balance model complexity and generalization capability.[20]

3.4.1. Input Data Configuration

- **Look-back Window:** $W = 30$ sessions. This parameter was selected based on empirical optimization results (Grid Search), representing an optimal trade-off between: (1) providing enough historical context for the model to recognize Pattern Recognition and (2) minimizing the noise impact of distant past data with no predictive value (Stale Information). This time frame corresponds to about 1.5 trading months, suitable for capturing short-to-medium-term momentum specific to cash flow in the Vietnamese market.[21]

- **Batch Size:** $B = 32$. This size helps stabilize gradient updates (Gradient Descent stability) while efficiently utilizing GPU memory.

3.4.2. Network Topology Configuration

Based on the technical specifications of the source code, the architectural hyperparameters are set in detail as follows:

For CNN-BiLSTM Model: The model receives input of size $(batch, seq\_len, features)$ and permutes dimensions to fit the convolutional layer.

- **CNN Layer (Feature Extraction):** Uses 1 Conv1D layer with output filters ($out\_channels$) = 64 and kernel size ($kernel\_size$) = 3. This layer is followed by a 1D Max Pooling layer with size 2 to reduce dimensionality. A light Dropout layer ($p = 0.1$) is applied immediately after CNN to enhance generalization.

- **Bi-LSTM Layer (Sequence Dependency Learning):** Consists of 2 stacked LSTM layers, set in bidirectional mode. Hidden state size ($hidden\_size$) is 64. Due to bidirectionality, the output of this layer has a feature size of $64 \times 2 = 128$.
- **Output Layer (Classification):** Uses a Linear (Fully Connected) layer receiving 128-dimensional input and mapping to 3 target classes (Up/Down/Sideways).
- **Regularization:** Applies Dropout with rate $p = 0.25$ between LSTM layers and before the final Linear layer.[22]

For Transformer Encoder Model: The model processes data in 'batch_first=True' format and uses pooling mechanism by taking the last time step ($x[:, -1, :]$).

- **Input Projection:** Original data is passed through a Linear layer to project dimensions to hidden vector space $d_{model} = 64$.
- **Positional Encoding:** Added directly to feature vectors for the model to perceive temporal order.
- **Encoder Block:** Uses 2 stacked Encoder layers ($N = 2$). Each layer includes Multi-Head Attention mechanism with 4 attention heads ($h = 4$) to learn multi-dimensional relationships.
- **Position-wise Feed-Forward Network:** Hidden layer size inside the Encoder block is set to $d_{ff} = 128$.
- **Output Layer:** After extracting features at the last time step, data passes through a Dropout layer ($p = 0.25$) and a final Linear layer to map from $d_{model} = 64$ to 3 result classes.

3.4.3. Training Setup

- **Loss Function:** Applies Focal Loss instead of standard Cross-Entropy. With focusing parameter $\gamma = 2.0$, this function is designed to automatically down-weight "easy" examples (usually the majority sideways trend) and focus training on "hard" examples (important trend reversal points). This strategy helps thoroughly overcome the Class Imbalance phenomenon specific to financial markets.[23]
- **Adaptive Learning Rate Strategy:** Learning rate is initialized at $\eta = 1e^{-3}$ (0.001). To ensure stable model convergence and precise weight refinement when approaching the optimum, the study applies dynamic ReduceLROnPlateau adjustment. Specifically, the system monitors Validation Loss; if this metric does not improve after a patience of 8 epochs, the learning rate will be automatically reduced by a decay factor of 0.5.[24]
- **Max Epochs:** 100 epochs.
- **Early Stopping:** Patience = 20 epochs. If accuracy on the validation set does not improve after 20 consecutive epochs, training will automatically stop and restore the best weights (Best Checkpoint) to avoid Overfitting.

4. RESULTS AND DISCUSSION

4.1. Evaluation Metrics

In the context of financial data often encountering class imbalance, relying solely on accuracy can lead to misleading conclusions. Therefore, the study establishes a multi-dimensional evaluation system including:

- **Accuracy:** The ratio of correctly predicted samples to the total number of test samples. This metric provides an initial overview of model performance.
- **Precision:** The ratio of correctly predicted samples among all samples classified as that class (TP / (TP + FP)). In trading, this metric represents signal reliability; low Precision implies the model emits many False Positives, leading to loss risk when entering wrong positions.
- **Recall (Sensitivity):** The ratio of correctly predicted samples to the total actual samples of that class (TP / (TP + FN)). This metric measures the model's

opportunity "capture" capability; low Recall means the system misses important price waves (False Negatives), causing opportunity costs.

- **Balanced Accuracy:** Calculated as the arithmetic mean of recall of each separate class (Up/Down). This is a critical metric in this study to eliminate "virtual accuracy" – where the model merely predicts the majority class (e.g., always predicting sideways). We establish a baseline threshold for Balanced Accuracy at > 60% to confirm the model's ability to exploit real signals.[25]

- **Weighted F1-Score:** The harmonic mean between Precision and Recall, weighted by the number of samples of each class. This metric ensures objective assessment of the harmony between precision and coverage, especially important when label distribution is uneven.

4.2. Analysis and Discussion of Results by Horizon

Empirical results record a distinct divergence in model performance corresponding to the length of the forecast horizon. This divergence reflects the intrinsic movement law of the stock market: stochasticity dominates in the short term, while fundamental convergence dominates in the long term.

4.2.1. Performance on Short-term Horizon (Short-term: T+1, T+3, T+5)

In short time frames, prediction performance is strongly impacted by Market Noise, leading to significant challenges for deep learning models.

Table 1. Average Performance on Test Set - Short-term Horizon

Ticker	Hor.	Model	Acc	B. Acc	F1
VCB	T+1	CNN-BiLSTM	55.38%	50.52%	45.80%
VPB	T+1	Transformer	56.57%	58.24%	51.44%
MBB	T+3	CNN-BiLSTM	59.41%	61.34%	56.46%
BID	T+5	Transformer	69.59%	61.84%	64.86%

Discussion:

- *Impact of Random Walk Theory:* For the majority of stocks (typically VCB, TCB at T+1), Balanced Accuracy fluctuates around the random threshold (50% - 55%). This reinforces the hypothesis that in extremely short time frames, price movements are mainly dominated by crowd psychology and momentary news (shocks), causing technical and fundamental signals to be obscured, not yet having enough time to reflect in the data structure.

- *Capability to Detect Micro-patterns:* Despite noise influence, positive exceptions like MBB (T+3) and BID (T+5) with Balanced Accuracy exceeding 61% have demonstrated the potential of deep learning architectures. Specifically, the CNN-BiLSTM architecture effectively leveraged the local feature extraction mechanism, allowing identification of candlestick patterns or subtle volume fluctuations signaling short-term momentum. This result suggests the potential for applying the model to High-frequency trading strategies on highly liquid stocks.

4.2.2. Performance on Medium and Long-term Horizon (T+30, T+60, T+90)

As the forecast horizon extends, the lag of financial information is gradually absorbed into prices. This is the phase where models demonstrate superior performance, confirming the value of integrating fundamental variables.

Table 2. Typical Performance on Medium and Long-term Horizon (Balanced Acc > 80%)

Ticker	Hor.	Model	Acc	B. Acc	F1
BID	T+30	Transformer	76.51%	70.55%	73.91%
CTG	T+60	CNN-BiLSTM	93.90%	80.30%	93.25%
HDB	T+60	Transformer	87.72%	83.06%	87.90%
VPB	T+90	CNN-BiLSTM	72.73%	85.94%	81.65%
HDB	T+90	Transformer	88.80%	92.55%	89.38%

Discussion:

- *Information Convergence Point at T+30:* The result of BID (Balanced Acc 70.55%) at T+30 reflects an optimal balance point, where stock prices have reacted to macroeconomic changes (interest rates, exchange rates) but have not yet been diluted by uncertain factors too far in the future.

- *Leading Role of Industry Features:* The strong breakthrough at T+60 and T+90 frames (typically HDB reaching 92.55% Balanced Acc and 89.38% F1-Score) is empirical proof that banking-specific indicators (NIM, NPL, Credit Growth) act as the main drivers shaping long-term price trends. The Transformer model with Attention mechanism learned this long-term causal relationship effectively.

- *Effectiveness of Adaptive Thresholding Technique:* Applying Adaptive Thresholding successfully filtered out sideways market phases (sideway noise). This helped the training dataset for long-term frames become "cleaner," clearly polarizing between Up and Down trends, thereby significantly improving classification model accuracy.

4.3. Impact Analysis and Economic Significance of Sector-Specific Features

One of the most significant contributions of this study is providing quantitative evidence regarding the leading role of banking-specific indicators within deep learning models. To decode the "interpretability" mechanism and verify the alignment between model results and financial theory, this study utilizes the *Mean Decrease in Impurity (MDI)* technique to measure the degree of information contribution.

The most quintessential example of this pattern is the case of VPB Bank. The tables below compare the shifts in feature weights between two time horizons:

Table 3. Top 5 most important features of VPB - Short-term Horizon (T+5)

Rank	Feature Name	Indicator Group	Coefficient
1	**MACD**	Technical	0.1514
2	Close_MA7	Trend	0.1385
3	RSI_14	Momentum	0.1293
4	High_to_Low	Price Volatility	0.1279
5	Volatility_14	Volatility Degree	0.1147

Table 4. Top 5 most important features of VPB - Long-term Horizon (T+90)

Rank	Feature Name	Indicator Group	Coefficient
1	**NPL (%)**	**Credit Quality**	**0.1160**
2	EPS (VND)	Profitability	0.0641
3	Provision_Cov	Risk Provisioning	0.0620
4	Loan_to_Asset	Asset Structure	0.0520
5	Time_Index	Long-term Trend	0.0464

The analysis results indicate that the model has established an **intelligent adaptive mechanism** relative to the forecasting horizon:

- **In the short-term horizon (T+5):** The model concentrates weights on technical indicators (Volatility, MACD). This accurately reflects the characteristics of the Vietnamese stock market, where short-term fluctuations are often dominated by crowd psychology and speculative cash flow (Market Sentiment) rather than fundamental factors.

- **In the long-term horizon (T+90):** The weight structure reverses completely toward *Financial Health* indicator groups. The absolute dominance of variables such as **NIM** (Profitability Efficiency) and **NPL** (Credit Risk) demonstrates that the model has automatically learned to value the business based on intrinsic capabilities. Specifically:
  - *NIM and Net Interest Income* serve as Growth Drivers.
  - *NPL and Provision Coverage Ratio* serve as Risk Discount Factors.

This shift demonstrates that integrating Domain Knowledge not only improves statistical accuracy but also enables the model to grasp the economic nature of stock price movements, overcoming the limitations of pure time-series forecasting methods.

4.4. Architecture Comparison: CNN-BiLSTM or Transformer?

Experiments indicate distinct functional specialization between the two architectures, rejecting the "one-size-fits-all" view:

- **CNN-BiLSTM (Medium-term Stabilizer):** With the advantage of convolutional filters in local noise reduction, this architecture maintains the highest stability in the medium term ($T + 30$ to $T + 60$). It operates effectively as a Trend Filter, minimizing false signals during sideways market phases.
- **Transformer Encoder (Long-term Spearhead):** At $T + 90$, Transformer completely outperforms thanks to the Multi-Head Self-Attention mechanism. Unlike recurrent networks (RNNs) which face limitations in "forgetting" information in long sequences, Transformer establishes direct global connections, allowing linking past macroeconomic events (e.g., interest rate changes from 3 months ago) with current price fluctuations. This ability to model long-range dependencies is key to achieving >90% accuracy.

From this performance divergence, the study recognizes the potential to develop "Adaptive Multi-scale Architectures" in the future. Instead of maintaining rigid separation, next-generation models could integrate Context Gating mechanisms to automatically adjust weights between local and global features. This mechanism allows the model to dynamically adapt to Market Regimes—automatically prioritizing stability during accumulation phases and shifting to enhanced sensitivity when confirming long-term macro trends.

4.5. Discussion: Contributions to Data Methodology and Managerial Implications

4.5.1. Core Value: Data-Centric AI Approach

Unlike previous studies that often focused purely on refining model architectures (Model-Centric), this study asserts that in financial problems, the biggest barrier lies in input data quality and structure. The most important contribution of the project is the standardization of the Data Construction Pipeline:

- **Thorough Resolution of Information Latency:** Through the Backward Merge-As-Of algorithm, the study proves that deep learning models are only truly effective when input data strictly adheres to temporal causality, completely eliminating look-ahead bias often found in financial report data.
- **Handling Data Uncertainty:** The application of the Adaptive Thresholding mechanism contributes a new methodology: Training data needs to be "personalized" according to the specific Volatility of each stock. This helps separate Signal from Noise – a key foundation for building clean datasets for financial time series forecasting problems.

4.5.2. System Role Positioning: From "Forecasting" to "Decision Support"

Experimental results show distinct performance divergence between short and long terms, leading to a repositioning of the model's actual role:

- **Strategic Reference Channel:** Instead of being used as a risky High-frequency Trading tool, the system optimizes the role of periodic strategic planning. The convergence of high accuracy at T+60 and T+90 frames aligns with efficient market theory: In the short term, prices are strongly impacted by crowd psychology and random news (Random Walk), making prediction low-probability. Conversely, in the medium and long term (equivalent to 1 quarter), price fluctuations more closely reflect Fundamental factors and macro cycles – features that the Transformer architecture excels at extracting and memorizing.

- **"Independent Critic" Mechanism (Second Opinion):** The system operates as an objective hypothesis verification layer (Cross-check mechanism). In investing, humans often suffer from psychological biases like FOMO (fear of missing out) or Loss aversion. When an investor's subjective judgment coincides with the model's long-term signal, reliability is reinforced. Conversely, a contradiction acts as an early warning signal, forcing investors to review risks with discipline rather than acting on emotion.
- **Parallel Consensus Strategy as a Safety Filter:** This strategy aims not to maximize the number of orders, but to optimize order quality (Precision over Recall). By acting only when there is consensus between different model architectures, investors accept missing ambiguous opportunities ("gray zones") to preserve capital. This is risk management thinking consistent with professional financial institution standards.

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

This study has successfully established a methodological framework for applying deep learning to analyze the Vietnamese stock market, shifting the focus from "pure price prediction" to "data construction and decision support". Key conclusions include:

1. **Standardization of Data Engineering:** The study's greatest contribution is the multi-level "Feature Universe" construction process and multi-frequency data processing technique. We affirm that integrating Domain Knowledge into data cleaning and structuring is a prerequisite for AI to operate effectively in noisy environments like the stock market.
2. **Objective Reference Framework with Reasonable Trade-off:** The system has proven its role as a powerful support tool in medium and long-term horizons. Empirical results achieved impressive accuracy (Balanced Accuracy > 80% at $T+60, T+90$) on selected data samples. It should be emphasized that this result is achieved through a **trade-off** between Coverage and Precision: The model accepts not making predictions during sideways market phases to ensure the highest reliability when the trend is clear.
3. **Risk Management Mindset as Foundation:** Through the Adaptive Thresholding mechanism and Parallel Consensus Strategy, the study prioritizes capital safety over expected profit. The model does not replace human decision-making but acts as a "filter" helping to eliminate false signals and confirm high-quality investment opportunities, assisting investors in maintaining discipline in the face of market volatility.

5.2. Recommendations and Future Directions

The methodological value of the study – particularly the Backward Merge data processing and adaptive labeling – has high generalizability. Potential expansion directions include applying this processing framework to other sectors or macroeconomic forecasting problems, contributing to the substantive development of AI in Finance in Vietnam.

Acknowledgements

This research was conducted within the framework of the student scientific research project of Quy Nhon University for the 2025-2026 academic year, under grant number 2025.1105.55. The authors would like to sincerely thank the University and the Department of Mathematics and Statistics for providing favorable conditions and financial support to complete this study.

REFERENCES

1. State Securities Commission of Vietnam, “Overview of Vietnam Securities Market,” 2024.
2. N. T. T. Loan, D. N. Hung, and V. T. T. Van, “Application of ARIMA Model and Deep Learning in Forecasting Stock Price in Vietnam,” *Salud, Ciencia y Tecnología*, 2025.
3. T. Phuoc et al., “Applying Machine Learning Algorithms to Predict the Stock Price Trend in the Stock Market – The Case of Vietnam,” *Humanities and Social Sciences Communications*, 2024.
4. State Securities Commission of Vietnam, *Annual Report on Vietnam Stock Market 2023*, 2024.
5. M. Darwish et al., “Stock Market Forecasting: From Traditional Predictive Models to Large Language Models,” *Computational Economics*, 2025.
6. J. Chen and Y. Li, “Adaptive Thresholding for Financial Time Series Prediction based on Market Volatility,” *Journal of Forecasting*, 2021.
7. S. Jenish, B. A. Kumar, and N. V. K. Rao, “Hybrid CNN-LSTM Models for Stock Market Prediction,” *Cureus*, vol. 16, no. 6, 2024.
8. M. R. Altay and Ö. Ulusoy, “Stock Market Prediction Using Hybrid CNN–BiLSTM Model,” *Risks*, vol. 12, no. 3, p. 61, 2024.
9. Y. Kim, “Convolutional Neural Networks for Sentence Classification,” in *Proc. EMNLP*, 2014.
10. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
11. A. Vaswani et al., “Attention Is All You Need,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
12. VNExpress, “Historical Data for VN-Index (2010-2024),” 2024. [Online]. Available: `https://dautu.vnexpress.net/`.
13. Vietcombank, *Consolidated Audited Financial Statements 2023*, 2024.
14. H. Nguyen and T. Le, “Integrating Fundamental and Technical Analysis for Banking Stock Prediction in Emerging Markets,” *Asian Academy of Management Journal*, 2019.
15. J. W. Wilder Jr., *New Concepts in Technical Trading Systems*. Trend Research, 1978.
16. G. Appel, *The Moving Average Convergence-Divergence Trading Method*. Signalert Corp, 1979.
17. P. P. Athanasoglou et al., “Bank-specific, industry-specific and macroeconomic determinants of bank profitability,” *Journal of International Financial Markets*, 2008.
18. M. F. Dixon et al., *Machine Learning in Finance: From Theory to Practice*. Springer, 2018.
19. P. J. Huber, *Robust Statistics*. John Wiley & Sons, 2011.
20. L. Prechelt, “Early Stopping - But When?” in *Neural Networks: Tricks of the Trade*. Springer, 1998.
21. J. Smith and T. Brown, “Time Series Analysis with Sliding Windows,” *Journal of Data Science*, 2020.
22. N. Srivastava et al., “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research (JMLR)*, 2014.
23. T.-Y. Lin et al., “Focal Loss for Dense Object Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
24. D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv:1412.6980*, 2014.
25. K. H. Brodersen et al., “The Balanced Accuracy and Its Posterior Distribution,” in *Proc. ICPR*, 2010.