ISANE: An Interpretable Real-time News Event Evaluation Stock Prediction Model

Anonymous ACL submission

Abstract

Analyzing messages in the stock financial market helps investors to timely understand market sentiment, thereby reducing investment risks. Prior research typically employed pre-trained semantic analysis models to obtain embedded vectors of news events for predicting stock market trends. However, these stock prediction models lack real-time analysis of news events and encounter inaccuracies in text semantics. Moreover, they usually function as black-box models, making it challenging for investors to comprehend prediction logic and enhance it. Therefore, we propose a news event assessment model named ISANE. This model represents news events as graph-structured data and utilizes Graph Attention Networks (GAT) to acquire semantic structural information from news events. Simultaneously, it employs BiLSTM networks to capture dynamic changes in market information. Ultimately, the model combines the semantic structural information of news events with historical market data to evaluate their impact on the stock market and enhance the model's interpretability by visualizing the structural information of news text. Additionally, we use price fluctuations as labels for news events in model training, enabling the model to adapt more swiftly to market changes. Comparative experiments with existing state-of-the-art (SOTA) models demonstrate an 18% enhancement in our model's prediction accuracy. However, further empirical testing and validation are necessary to ensure its viability in real-world stock markets.

1 Introduction

Message analysis refers to the process of collecting, categorizing, interpreting, and evaluating messages disseminated through social media, news outlets, and publicly listed companies. Conducting a more comprehensive and effective analysis of news and public opinion aids investors in grasping market dynamics and trends, facilitating informed trading decisions with substantial economic value. Therefore, researching how to articulate news effectively and leveraging event information for stock prediction holds significant significance and practical application. Figure 1 illustrates the impact process of financial events on the stock market.

Figure 1: The Impact Process of Financial Events on the Stock Market

Presently, research focused on message analysis in the stock financial market commonly utilizes news events as information sources. It involves employing word-level language models (Selvin et al., 2017; Ding et al., 2015) or sentence-level language models (Asraf Roslan and Fazalul Rahiman, 2018; Lakshya et al., 2022) to obtain embedded vectors of news events, subsequently combined with daily market data to predict stock movement trends. However, word-level language models are excessively intricate, failing to capture long-term dependency information within news events. Sentence-level language models, on the other hand, are overly coarse, unable to emphasize crucial information and lacking in inherent logical analysis of news events, thereby encountering issues with inaccurate textual semantics.Moreover, current methodologies (Sridhar and Sanagavarapu, 2021; Mandalapu et al., 2019) predominantly use trading days as fundamental units, considering all news and stock technical features of the same day as input data for predicting the following day's stock movement trends. This approach lacks the capability to promptly handle news events, leading to information lag issues and rendering the model insufficient in responding to unforeseen events.Additionally, stock investment, being a high-risk high-return financial activity, cannot solely rely on model-predicted market fluctuations for investors' buy and sell decisions. Neural

network models, being typical 'black box' models, lack a means for humans to fully comprehend the model's decision-making process from a human perspective. Even if stock prediction models successfully forecast market rises or falls, individuals tend to assess the reliability of these predictions based on their own experiences and insights. In other words, regarding stock prediction, investors place a high demand on the interpretability (Shalev-Shwartz and Ben-David, 2013) of the model.

The Graph Attention Network(GAT) (Veličković et al., 2018) autonomously learns and captures complex dependencies in text through attention mechanisms, proving highly beneficial in comprehending intricate semantic structures within financial news. Its inherent ability for focal summarization also offers a means to enhance the post-hoc interpretability of deep learning models. Inspired by this, we propose an ISANE (Graph Attention-based News Event Evaluation) model utilizing GAT.This model integrates the elementary discourse units (EDUs) from Rhetorical Structure Theory (RST) (Taboada and Mann, 2006) as fundamental semantic information units to extract from news texts. It constructs a graph representation of the news text based on rhetorical relationships between each EDU, continually updating this representation using GAT. This graph representation, coupled with market information, facilitates stock movement classification prediction. Moreover, to emphasize the real-time nature of event-driven stock prediction models, we build data samples based on news events, automatically labeling news event texts according to historical stock data. This enables the establishment of associations between financial news and stock fluctuations, enabling real-time predictive modeling.

The primary contributions of this work can be summarized as follows:

1) We represent news events as graph-structured data and enhance the representation of news texts by incorporating structural information from financial news.

2) We have devised a financial news event assessment model based on Graph Attention Networks, denoted as ISANE. This model employs Graph Attention Networks (GAT) to update the graph representation of financial news and utilizes BiLSTM to gather stock market information. Ultimately, it integrates these pieces of information to achieve stock prediction.

3) We conducted a series of comparative experiments to validate the effectiveness of our model. Compared to the existing state-of-the-art (SOTA) models, our model significantly outperforms baseline models across various metrics and exhibits efficient interpretability.

2 Related Work

2.1 Financial Event-Driven Analysis and Its Interpretability

Financial event-driven analysis, also known as fundamental analysis, primarily focuses on external event signals that influence financial markets, such as financial news, trending topics on Twitter and Weibo, national policies, and corporate behaviors—information that reflects market sentiment. In recent years, numerous attempts have been made to leverage online news information for stock trend prediction. Schumaker and Chen (2009) proposed a system named AZFinText, utilizing word-level natural language processing techniques to analyze financial news text and predict stock price trends based on the analysis results. Zou et al. (2022) introduced an approach called SRLP, utilizing Semantic Role Labeling (SRL) to create embedded representations for each news paragraph, ultimately using Transformers to process these embedded vectors for stock movement trend prediction.Furthermore, model interpretability (Wang and Ye, 2020) poses a significant challenge in financial event-driven analysis, enhancing stock investors' trust and comprehension of model predictions. This capability generally manifests in two aspects: focal information summarization and logical analysis of information.

In existing literature, attention mechanisms are often utilized to highlight key information. Chen et al. (2021) used attention to characterize the impact of individual news items within a trading day, predicting cumulative abnormal returns within a three-day window. Sawhney et al. (2020) employed hierarchical attention mechanisms to jointly learn models from historical prices, stock news, and internal stock relationships. However, traditional attention mechanisms only express the model's sensitivity to different data points, lacking an explanation for the logical relationships within the data.Furthermore, Deng et al. (2019) considered leveraging entities and relationships within a knowledge graph as an intuitive approach to understand and explain model predictions. Deng

et al. (2023) utilized Structural Equation Models (SEM) to establish causal relationships between stock prices and various factors. By inferring these causal relationships, they predicted future stock price trends. However, the complexity of structural relationships often impedes the emphasis on key information, leading to misleading interpretations for stock investors.Although numerous studies have explored post-hoc interpretability in financial event-driven analysis, an efficient and intuitive method to help stock investors understand why a model makes particular predictions remains lacking.

2.2 Graph Attention Networks

The Graph Attention Network (GAT) is a deep learning model utilized for graph-structured data. It autonomously learns and captures complex dependencies within text through attention mechanisms, proving highly useful in understanding intricate semantic structures within financial news. Presently, GAT has been employed in various NLP tasks such as text summarization and classification. Wang et al. (2022) proposed the EDU4FD model for fake news detection, utilizing Relation Graph Attention Network (RGAT) to acquire structural information from news text and combining it with semantic information captured by the GRU-GA-based model for fake news detection. GAT also offers a certain level of interpretability through attention mechanisms, aiding in comprehending the decision-making process of the model. Moreover, compared to some other graph neural networks, GAT can better utilize parallel computing resources as it does not explicitly require message-passing stages during computing node representations.However, due to the complexity of financial text data in stock event analysis, the application and research of Graph Attention Networks (GAT) in this domain are still in the early stages.

3 Methods

3.1 Overall Framework

The ISANE model mainly consists of three main components: the News Semantic Analysis module, the Market Information Analysis module, and the Fusion Classifier module. The News Semantic Analysis module is used to obtain semantic structural information from news events, while the Market Information Analysis module is designed to extract market temporal information. The Fusion Classifier module combines these two types of information to predict stock movements. Figure 2 illustrates the framework of our model, and we will provide detailed explanations of our approach in sections 3.3-3.5.

3.2 Real-time Labeling of News Events

To reflect the real-time nature of financial event-driven analysis, we choose to build the data samples based on financial events. For each news event D corresponding to a stock code c and the time of the news event t, we obtain the historical market data for that stock in the previous T days, denoted as $X_T = (x_1, x_2, \cdots, x_T) \in \mathbb{R}^{T \times F_1}$, where F_1 represents the number of feature factors. X_T is supplemented with the news event information and input into the BiLSTM model. Additionally, news event D is labeled with real-time label y based on the price change. In this task, for each news event D, the price change R of the stock's closing price on the day following the event's occurrence is calculated, and the specific calculation process is as follows:

$$R = \frac{Close_{t+1} - Close_t}{Close_{t+1}} \tag{1}$$

Where $Close_{t+1}$ represents the closing price of the stock on the day following the occurrence of news event D, and $Close_t$ represents the closing price of the stock on the day of the occurrence of news event D. The label y for news event D is determined based on the price change R:

$$y = \begin{cases} 1, & \text{if } R \geq \tau \\ 0, & \text{if } R < -\tau \end{cases} \tag{2}$$

Where τ represents the price change threshold used to filter out data with insignificant price changes. In this paper, τ is set to 0.04, and data with price changes below this threshold are discarded. Based on the labels, the event-driven stock prediction problem can be described as:

$$\hat{y} = f(D, X_T) \tag{3}$$

Here, $\hat{y}$ represents the output of the stock prediction model, indicating the probability of the corresponding stock going up after the occurrence of news event D. $f(\cdot)$ represents the stock prediction model. The model's input consists of the data sample composed of D and X_T.

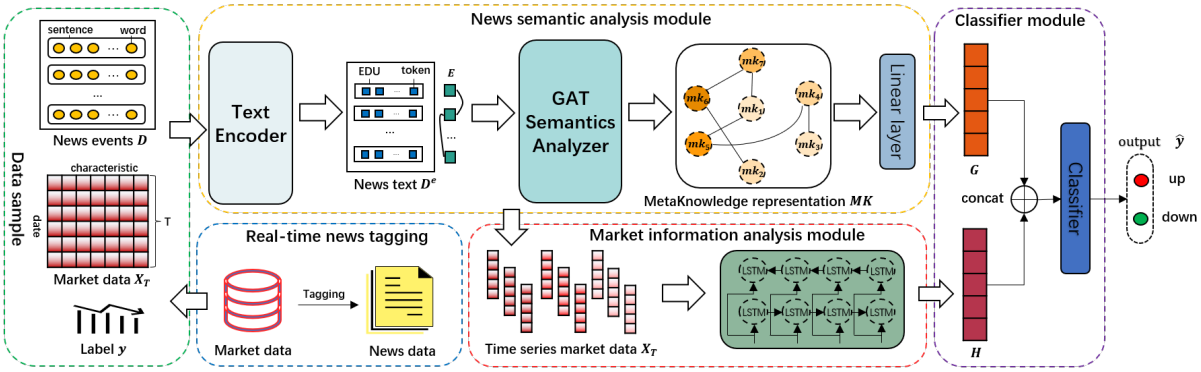


Figure 2: The overall framework of ISANE.

3.3 News Semantic Analysis module

The News Semantic Analysis Module used to process news text data with the aim of extracting meta-knowledge from news events to obtain a compact representation. This module consists of a Text Encoder and a GAT Semantic Analyzer. The Text Encoder first segments the news text into independent EDU speech units and then constructs a graph representation of the news text based on EDU with their feature representations and the rhetorical relationships between them. The GAT Semantic Analyzer utilizes a GAT network to continuously aggregate the features of EDU nodes into the meta-knowledge of the news event, thereby obtaining a compact representation of the news event. Figure 3 illustrates the main framework of the News Semantic Analysis Module.

3.3.1 Text Encoder

The Text Encoder first represents news events into $D = (w_1, w_2, \cdots, w_N) \in \mathbb{R}^{N \times F_2}$ using the Stanford CoreNLP[17] tokenization tool, where w_i represents the i-th token, N represents the total number of tokens in the text, and F_2 is the text feature dimension. Unlike previous research, we then use Elementary Discourse Units (EDUs) from rhetorical theory as the basic units for extracting news semantics. To better segment the text into EDUs, we employ the Discourse Parse and Label Propagation (DPLP) model proposed by Ji and Eisenstein (2014). The news text segmented by the DPLP model can be represented as $D^e = DPLP(D) = (EDU_1, EDU_2, \cdots, EDU_n) \in \mathbb{R}^{n \times S \times F_2}$, where n represents the number of EDUs, S represents the maximum field length of EDUs, and the i-th EDU is represented as $EDU_i = (w_{i1}, w_{i2}, \cdots, w_{i\delta_i}) \in \mathbb{R}^{\delta_i \times F_2}$, where w_{ij} represents the j-th token of the i-th EDU, and δ_i represents the number of tokens in the i-th EDU. Then, all EDU units are parsed using the Rhetorical Structure Theory (RST) parser to obtain the rhetorical relations between EDUs, $E = \{(EDU_i, EDU_j, r) \mid EDU \in D^e, r \in \varepsilon\}$, where ε is the set of rhetorical relations. In this paper, we adopt the 19 rhetorical relations proposed by Li et al. (2014) to describe the complex discourse structure in news text.

3.3.2 GAT Semantic Analyzer

The GAT Semantic Analyzer is utilized for handling the graph-structured data of news text.Its main tasks are divided into two steps: Graph Representation and Graph Learning.

Graph Representation: The GAT semantic analyzer first uses the BERT model to transform each token into word embeddings, i.e.,

$$h_{ij} = bert(w_{ij}) \tag{4}$$

Here, $h_{ij} \in \mathbb{R}^{F_3}$ represents the word encoding of the j-th token in the i-th EDU, where F_3 is the feature dimension. Next, we employ the self-attentive span extractor (SpanExt) proposed by Lee et al. (2017) to learn the textual features of each EDU. The specific calculation is as follows:

$$\alpha_{ij} = W_2 \cdot ReLU(W_1 h_{ij} + b_1) + b_2 \tag{5}$$

$$a_{ij} = \frac{exp(\alpha_{ij})}{\sum_{k=1}^{\delta_i} exp(\alpha_{ik})} \tag{6}$$

$$h_i^e = \sum_{j=1}^{\delta_i} a_{ij} \cdot h_{ij} \tag{7}$$

Here,α_{ij} represents the attention score for the j-th token in the i-th EDU, W_1, W_2, and b are the parameters to be learned, and a_{ij} is the attention weight for the j-th token in the i-th EDU. $h_i^e \in \mathbb{R}^{F_3}$ represents the feature representation of the i-th EDU. At this point, we can obtain the EDU representation of the news text, denoted as

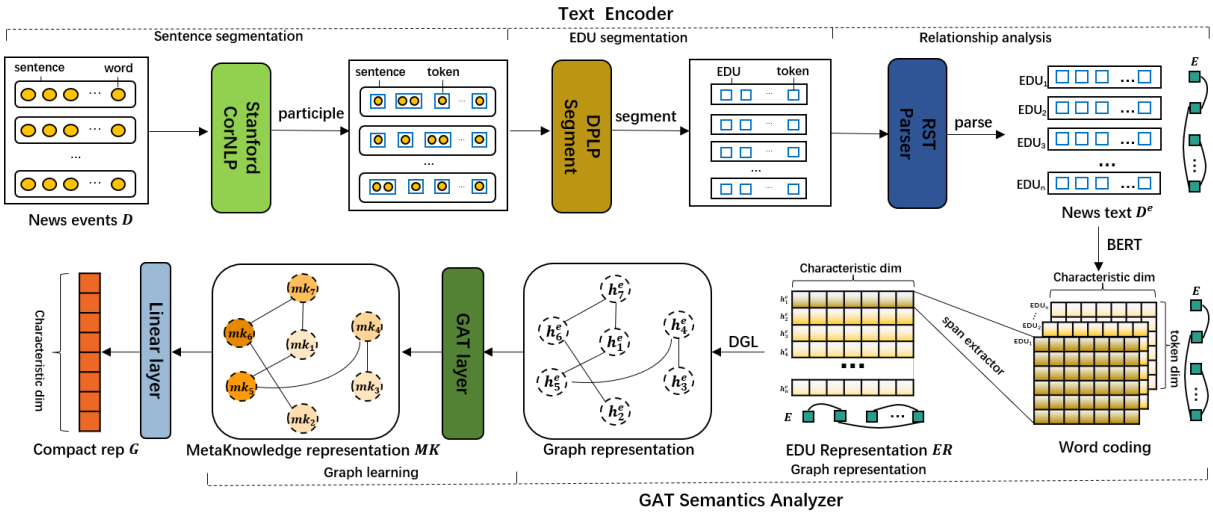


Figure 3: News Semantic Analysis module

$ER = \{h_1^e, h_2^e, \cdots, h_n^e\} \in \mathbb{R}^{n \times F_3}$.Next, we use the feature representations of each EDU as node information and consider each discourse dependency between EDUs E as undirected edges to construct the graph representation of the news event. The visual result is shown in Figure 4.

Graph Learning: During the graph learning process, the Graph Attention Network (GAT) is used to handle the graph representation of the news event, thereby obtaining the meta-knowledge of the news text. It can aggregate neighboring information based on dependencies and highlight key adjacent information through attention mechanisms to capture internal information between nodes effectively. The GAT network primarily consists of Graph Attention Layers, and the attention coefficient calculation for node pairs (i, j) in each attention layer is as follows:

$$\beta_{ij} = \frac{exp(\sigma(\mathcal{Q}^{\mathcal{T}}[W_3 h_i^e \parallel W_3 h_j^e]))}{\sum_{k=1}^{k \in \mathcal{N}_i} exp(\sigma(\mathcal{Q}^{\mathcal{T}}[W_3 h_i^e \parallel W_3 h_k^e]))} \quad (8)$$

Where β_{ij} represents the importance of node j to node i, $h_i^e \in \mathbb{R}^{F_3}$ are the input features of the nodes in this layer, $\mathcal{Q}$ and W_3 are shared learnable parameters, σ is the activation function, $\mathcal{N}_i$ denotes the neighborhood of node i, and $\parallel$ represents the concatenation operation. Next, update node features according to the contribution of the features. The specific calculation process is as follows:

$$mk_i = \sigma(\sum_{j=1}^{j \in \mathcal{N}_i} \beta_{ij} W_3 h_j^e) \quad (9)$$

Where $mk_i \in \mathbb{R}^{F_3}$ represents the text meta-knowledge obtained by aggregating node i. At this point, we can obtain the meta-knowledge representation of the news text, $MK = (mk_1, mk_2, \cdots, mk_n) \in \mathbb{R}^{n \times F_3}$, as visualized in Figure 4. Next, we use a linear layer to aggregate these meta-knowledges to obtain the compact representation of the news text. The specific calculation is as follows:

$$G = W_4 \cdot MK + b_3 \quad (10)$$

Where $G \in \mathbb{R}^{F_4}$ represents the compact representation of the news text D, W_4 and b_3 are the weight matrix and bias vector, and F_4 is the feature dimension. It's worth noting that we can analyze the importance of different information in the text by visualizing the meta-knowledge of the news text, thus providing post hoc interpretability to the model. Figure 4 illustrates a brief process of a news event being processed by the news semantic analysis module.

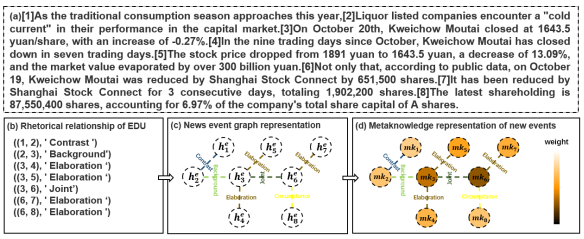


Figure 4: The process of the news semantic analysis module.Each [*] in Subfigure (a) represents an Elementary Discourse Unit (EDU).

3.4 Market Information Analysis Module

The Market Information Analysis Module is utilized to capture the historical market information

corresponding to the stock. During this process, the temporal data X_T is inputted into the BiLSTM (Graves and Schmidhuber, 2005) model to extract sequential characteristics of the market. The specific calculation process is as follows:

$$\overrightarrow{h} = \overrightarrow{LSTM}(X_T) \tag{11}$$

$$\overleftarrow{h} = \overleftarrow{LSTM}(X_T) \tag{12}$$

$$H = \overrightarrow{h} \parallel \overleftarrow{h} \tag{13}$$

Where $\overrightarrow{h} \in \mathbb{R}^{F_4}$ represents the forward hidden state of the BiLSTM model at the final time step, $\overleftarrow{h} \in \mathbb{R}^{F_4}$ represents the backward hidden state of the BiLSTM model at the final time step, $X_T \in \mathbb{R}^{T \times F_1}$ is the time series feature data of the stock market for the n days before the occurrence of the news event, and $H \in \mathbb{R}^{2F_4}$ represents the market information representation, encapsulating bidirectional temporal information from the stock market.

3.5 The Fusion Classifier Module

The Fusion Classifier Module concatenates the compact representation G of the news text with the market information H to predict stock movements, as calculated below:

$$\hat{y} = \sigma(W_5(G \parallel H) + b_4) \tag{14}$$

$\hat{y}$ represents the model's output, denoting the probability of a stock price increase after a news event. $W_5 \in \mathbb{R}^{3F_4}$ is the weight matrix, and b_4 is the bias vector. During the backpropagation process, cross-entropy loss is employed as the loss function to compute prediction errors and optimize network parameters. The loss function is defined as follows:

$$L = -\sum_{i=1}^{k} y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i) \tag{15}$$

Where K is the number of samples, y_i is the true label of the i-th sample (taking values of 0 or 1), and $\hat{y}_i$ is the model's predicted probability of the i-th sample belonging to the positive class.

4 Experimental Evaluation

This section first introduces the experimental data, parameter settings, evaluation metrics, and comparison methods. Then, it evaluates the performance of the proposed method through comprehensive experiments. Finally, it provides a comprehensive analysis of the effectiveness of each module.

4.1 Experimental Data

We selected the constituents of the Shanghai and Shenzhen 300 Index (HS 300) as the research objects and used a dataset consisting of 169,357 news records and daily technical indicator data from June 2021 to January 2023[1]. We split the data into three sets: the first 14 months for training, the middle 4 months for validation, and the last 2 months for testing. Figure 5 illustrates the data split for the experiments.

Figure 5: Dataset Split

4.2 Evaluation Metrics

In this study, we selected precision, recall, F1 score, direction accuracy (DA), and Matthews correlation coefficient (MCC) as quantitative evaluation metrics for the stock prediction experiments. The calculation methods for each metric are as follows:

$$Precision = \frac{t_p}{t_p + f_p} \tag{16}$$

$$Recall = \frac{t_p}{t_p + f_n} \tag{17}$$

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{18}$$

$$DA = \frac{n_{correct}}{n_{total}} \tag{19}$$

$$MCC = \frac{t_n \times t_p - f_n \times f_p}{\sqrt{(t_p + f_n)(t_p + f_p)(t_n + f_n)(t_n + f_p)}} \tag{20}$$

Where, $n_{correct}$ represents the number of days correctly predicted, n_{total} is the total number of days in the test set, $t_p(t_n)$ is the number of correctly predicted positive (negative) labels, and $f_p(f_n)$ is the number of incorrectly predicted positive (negative) labels.

4.3 Baseline Models

To validate the effectiveness of the ISANE model, we compared it with the following state-of-the-art (SOTA) models as baselines:

FinText (Rahimikia et al., 2021): Utilizing the FinText machine learning model for predicting stock movements and leveraging SHAP to extract

[1]You can obtain news data from https://www.cnrds.com and indicator data from https://tushare.pro.

Parameter	FinText	SRLP	FinBERT	ISANE
BSize	128	64	64	128
HLayer	500	300	100-25	64-32
FtDim	300	300	512	48,300,512,32
ActFunc	ReLU	ReLU	ReLU	ReLU/Sigmoid
Optmzr	Adam	Adam	Adam	Adam
LR	0.025	0.001	0.001	0.001
Epochs	5	300	300	500
EStop	/	/	50	10
Loss	MSE	CE,SSL	MSE	BCE

Table 1: Comparison of major parameters between ISANE and baseline models.

key phrases within news events, thereby enhancing the model's interpretability.

SRLP (Zou et al., 2022): Constructing representations of stock news events based on semantic role labels within sentence structures.

FinBERT (Jiang and Zeng, 2023): FinBERT conducts semantic analysis on financial news data and employs an LSTM model to predict financial market trends.

Additionally, we created several variants of ISANE to analyze its major modules in detail:

ISANE-E: Constructing graph representations of news events using sentences as basic units.

ISANE-M: This variant of ISANE uses only market information for stock prediction.

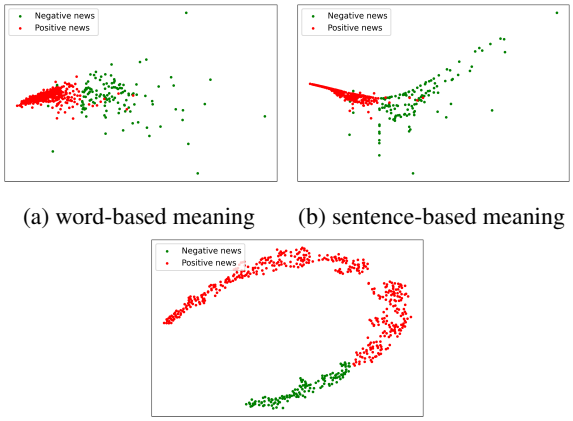
ISANE-N: This variant of ISANE uses only financial news information for stock prediction.

Table 1 shows a major parameter comparison between the ISANE model and the baseline models, with the baseline models using the best hyperparameters from their respective papers.

4.4 Experimental Evaluation

4.4.1 Text Semantics Analysis

To demonstrate the semantic representation of graph-structured data in news text, we conducted visualization experiments on the validation dataset. Specifically, for each text, we initially vectorized the news text using various semantic approaches, including word-based textual semantics, sentence-based textual semantics, and graph-structured textual semantics. Subsequently, utilizing K-means (Macqueen, 1967) clustering based on labels, we clustered the data samples and finally employed t-SNE (Van Der Maaten and Hinton, 2008) for visualization. The results, as depicted in Figure 6, illustrate each point representing a news text, where negative and positive news are respectively displayed in green and red. Through data observation, we observed the following: Sample clustering in Figure 6a exhibits poor clustering and unclear boundaries. This suggests that using word-based textual semantics fails to precisely capture text information, resulting in inaccuracies in semantic representation. Meanwhile, samples in Figure 6b demonstrate some clustering capabilities, yet the boundaries remain unclear. This is primarily because using sentence-based textual semantics, although capable of capturing long-term dependency information in text, struggles to highlight the key information due to the complexity of semantic information, leading to ambiguous semantic representation. Clearly, the samples in Figure 6c display better coherence and clearer boundaries. This is attributed to representing news as graph structures, which not only considers textual semantics but also incorporates structural information between sentences. Compared to conventional semantic approaches, graph structures exhibit superior capabilities in producing high-quality semantic representations.



(a) word-based meaning (b) sentence-based meaning

(c) graph-structured meaning

Figure 6: Comparison of News Event Embeddings

4.4.2 Accuracy Comparison Analysis

To compare the predictive performance of the models, we evaluated the baseline models on the entire dataset for stock prediction. Each method was run 12 times with model parameters set to the best parameter configurations provided in their respective papers, The entire training process was conducted on a server equipped with a GTX 1060 graphics card. From Table 2, we can observe the FinText model, which focuses solely on word-level granularity for semantic representation of news events, lacks semantic analysis of news event texts, resulting in relatively poorer performance compared to other models. While the FinBERT model in-

Models	Index				
	Precision	Recall	F1	DA	MCC
FinText	0.541	0.624	0.580	0.721	0.0795
SRLP	0.558	0.649	0.600	0.734	0.0804
FinBERT	0.549	0.697	0.614	0.707	0.0884
ISANE-E	0.517	0.614	0.561	0.724	0.0658
ISANE-M	0.505	0.683	0.581	0.716	0.0601
ISANE-N	0.512	0.695	0.590	0.701	0.0652
ISANE	0.593	0.717	0.649	0.766	0.0983

Table 2: Comparison of Accuracy between ISANE and Baseline Models.

corporates long dependency information between sentences, enabling better semantic representation of news events and enhancing model performance, it still lacks significant superiority compared to the SRLP model that considers triplet structures within sentences. Our proposed ISANE model captures long dependency information more accurately by employing more precise discourse units, enhancing the semantic representation of text with structural information. In experimental validation, the ISANE model achieved overall optimal results. To demonstrate the effectiveness of various modules within our model, we conducted three experiments based on variants of the ISANE model, as presented in Table 2.The ISANE-E model, utilizing sentences as the fundamental discourse unit, exhibited the poorest performance, validating that EDUs are more suitable than sentences for constructing textual graph-structured data. Although the ISANE-M and ISANE-N models performed reasonably well, their effectiveness was notably lower than that of the ISANE model, indicating the efficacy of our designed model modules. The comparison between the results of the ISANE-N and ISANE-M models demonstrates the complementary nature of our approach compared to traditional stock prediction models.

4.4.3 Interpretability Analysis

We selected a news event that was predicted to be a rise in stock price from the test data: "SF Express Holdings repurchases 2 billion shares for self-rescue" and visualized the news meta-knowledge obtained through the ISANE model for post hoc interpretability analysis, as shown in Figure 7. From the figure, we can see that among all the EDU nodes in the news, the second and sixth EDUs have the highest weights, indicating that they contain crucial information in the news text. The corresponding news content for these EDUs includes

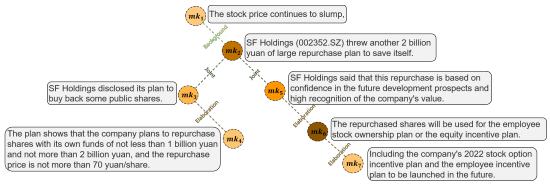


Figure 7: Interpretable analysis of news events.

"SF Express Holdings repurchases 2 billion shares for self-rescue" and "The repurchased shares will be used for employee stock ownership plans or equity incentive plans." This suggests that the ISANE model believes that SF Express Holdings' significant repurchase and equity incentive actions will drive an increase in its stock price. Overall, by analyzing the meta-knowledge of news events, ISANE provides investors with enhanced post hoc interpretability.

5 Conclusion

This paper addresses the significance of financial news analysis in stock market prediction and introduces a novel news event evaluation model, ISANE, based on Graph Attention Networks (GAT). ISANE represents news text as graph-structured data, utilizing GAT networks to extract concise representations of news text. It combines this information with market data for stock prediction. ISANE employs the price fluctuation as labels for news events, enabling real-time analysis. Moreover, ISANE enhances model interpretability by visualizing the metaknowledge within news text.The effectiveness and superiority of the ISANE model are validated through comparative experiments across multiple dimensions with existing state-of-the-art (SOTA) models.

Limitations

Despite having sound data sources, the concept of constructing data samples based on news events poses a highly time-consuming challenge in handling financial news data, thereby posing a difficulty for model training. Additionally, this method restricts the inclusion of other valuable information for stock prediction, such as interrelated data between news events and fundamental corporate information. Furthermore, simple concatenation of two message sources is not the optimal choice; hence, we aim to explore more efficient methods for information fusion.

References

Muhammad Aiman Asraf Roslan and Mohd Hezri Fazalul Rahiman. 2018. Stock prediction using sentiment analysis in twitter for day trader. In *2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pages 177–182.

Wei Chen, Manrui Jiang, Wei-Guo Zhang, and Zhensong Chen. 2021. A novel graph convolutional feature based convolutional neural network for stock trend prediction. *Information Sciences*, 556:67–94.

Shumin Deng, Jiaoyan Chen, Ningyu Zhang, Jeff Z. Pan, Wen Zhang, and Huajun Chen. 2019. Knowledge-driven stock trend prediction and explanation via temporal convolutional network. In *2019 World Wide Web Conference, WWW 2019, May 13, 2019 - May 17, 2019*, The Web Conference 2019 - Companion of the World Wide Web Conference, WWW 2019, pages 678–685. Association for Computing Machinery.

Yiqi Deng, Yuzhi Liang, and S. M. Yiu. 2023. Towards interpretable stock trend prediction through causal inference.

Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2015. Deep learning for event-driven stock prediction. In *24th International Joint Conference on Artificial Intelligence, IJCAI 2015, July 25, 2015 - July 31, 2015*, volume 2015-January of *IJCAI International Joint Conference on Artificial Intelligence*, pages 2327–2333. International Joint Conferences on Artificial Intelligence.

Alex Graves and Jürgen Schmidhuber. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5):602–610.

Yangfeng Ji and Jacob Eisenstein. 2014. Representation learning for text-level discourse parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13–24. Association for Computational Linguistics.

Tingsong Jiang and Andy Zeng. 2023. Financial sentiment analysis using FinBERT with application in predicting stock movement. Version: 1.

Lakshya, Prateek, and Divyasikha Sethia. 2022. Stock price prediction using news sentiment analysis. In *2022 2nd International Conference on Intelligent Technologies (CONIT)*, pages 1–6.

Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end neural coreference resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 188–197. Association for Computational Linguistics.

Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 25–35. Association for Computational Linguistics.

J Macqueen. 1967. Some methods for classification and analysis of multivariate observations. *Multivariate Observations*.

Arun Chaitanya Mandalapu, Saranya Gunabalan, Avinash Sadineni, Taotao Cai, Nur Al Hasan Haldar, and Jianxin Li. 2019. Correlate influential news article events to stock quote movement. In *15th International Conference on Advanced Data Mining and Applications, ADMA 2019, November 21, 2019 - November 23, 2019*, volume 11888 LNAI of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pages 331–342. Springer Science and Business Media Deutschland GmbH.

Eghbal Rahimikia, Stefan Zohren, and Ser-Huang Poon. 2021. Realised volatility forecasting: Machine learning via financial word embedding. *SSRN Journal*.

Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Ratn Shah. 2020. Deep attentive learning for stock movement prediction from social media text and company correlations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8415–8426. Association for Computational Linguistics.

Robert P. Schumaker and Hsinchun Chen. 2009. Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Trans. Inf. Syst.*, 27(2):1–19.

Sreelekshmy Selvin, R Vinayakumar, E. A Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. 2017. Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1643–1647.

Shai Shalev-Shwartz and Shai Ben-David. 2013. *Understanding machine learning: From theory to algorithms*, volume 9781107057135 of *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press. Pages: 397.

Sashank Sridhar and Sowmya Sanagavarapu. 2021. Analysis of the effect of news sentiment on stock market prices through event embedding. In *16th Conference on Computer Science and Intelligence Systems, FedCSIS 2021, September 2, 2021 - September 5, 2021*, Proceedings of the 16th Conference on Computer Science and Intelligence Systems, FedCSIS 2021, pages 147–150. Institute of Electrical and Electronics Engineers Inc.

M. Taboada and W.C. Mann. 2006. Rhetorical structure theory: Looking back and moving ahead. *Discourse Studies*, 8(3):423–459.

Laurens Van Der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2625.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph attention networks.

Lewen Wang and Zuoxian Ye. 2020. An interpretable framework for stock trend forecasting. *J. Phys.: Conf. Ser.*, 1634(1):012026. Publisher: IOP Publishing.

Yuhang Wang, Li Wang, Yanjie Yang, and Yilin Zhang. 2022. Detecting fake news by enhanced text representation with multi-EDU-structure awareness. *Expert Syst. Appl.*, 206:117781.

Jinan Zou, Haiyao Cao, Lingqiao Liu, Yuhao Lin, Ehsan Abbasnejad, and Javen Qinfeng Shi. 2022. Astock: A new dataset and automated stock trading based on stock-specific news analyzing model.