Heterogeneous Interactive Snapshot Network for Review-Enhanced Stock Profiling and Recommendation

Heyuan Wang1,3, Tengjiao Wang1,3, Shun Li2*, Shijie Guan1,3, Jiayi Zheng1,3 and Wei Chen1,3
1School of Computer Science, National Engineering Laboratory for Big Data Analysis and Applications, Peking University, China
2University of International Relations
3Institute of Computational Social Science, Peking University (Qingdao)
{wangheyuan, tjwang, guanshijie, jiayizheng, pekingchenwei}@pku.edu.cn, lishunmail@foxmail.com

Abstract

Stock recommendation plays a critical role in modern quantitative trading. The large volumes of social media information such as investment reviews that delegate emotion-driven factors, together with price technical indicators formulate a “snapshot” of the evolving stock market profile. However, previous studies usually model the temporal trajectories of price and media modalities separately while losing their interrelated influences. Moreover, they mainly extract review semantics via sequential or attentive models, whereas the rich text associated knowledge is largely neglected. In this paper, we propose a novel heterogeneous interactive snapshot network for stock profiling and recommendation. We model investment reviews in each snapshot as a heterogeneous document graph, and develop a flexible hierarchical attentive propagation framework to capture fine-grained proximity features. Further, to learn stock embedding for ranking, we introduce a novel twins-GRU method, which tightly couples the media and price parallel sequences in a cross-interactive fashion thereby catching dynamic dependencies between successive snapshots. Our model excels state-of-the-arts over 7.6% in terms of cumulative and risk-adjusted returns in trading simulations on both English and Chinese benchmarks.

1 Introduction

Accurately predicting stock movements is a prerequisite for making profitable investment decisions. Conventional literatures from the mathematical and deep learning communities [Nayak et al., 2015; Qin et al., 2017] hammer at decomposing stock’s volatility patterns from history technical indices. Nevertheless, unlike general time series, the dynamic of stock markets is highly stochastic and susceptible to non-stationary behaviors. With the growth of the Internet, massive online media reviews (such as financial news and tweets) provide great potential to reveal market environmental variations [Hu et al., 2018; Sawhney et al., 2021b]. However, previous studies on text-enhanced stock forecasting have limitations in terms of semantic representation and temporal dependency modeling. We elaborate our research insights from the following two perspectives.

First, existing stock prediction methods usually capture the semantic features of investment reviews via sequential or attentive models, whereas rich inter-text association knowledge is largely neglected. In fact, the consecutively delivered stock reviews are often context-relevant. Bridging them together through endogenous textual aspects such as topics and entities furnishes a broader scope of information for interpreting stock price oscillations. Fig. 1 illustrates the snapshots including daily technical signals and media reviews of $AAPL$’s evolving profile. On the $t_{th}$ trading day, the first and third texts both refer to “technology” and “product”, while the second and last texts are about “investment” opinions. These matched pairs portray complementary information toward specific topics. In addition, the first and last reviews mentioning the same entity transitively reveal the public’s expectation of stock benefits resulted from “product launches”. Moreover, the relatedness between different entities (e.g., “Apple” and “iPhone”) is also an important indication to expound the unified context, which can be discerned by referring to additional knowledge like Wikipedia. This inspires us building a heterogeneous information graph for media signals of each
stock snapshot, in order to make full use of semantic relevant clues to comprehensively assess underlying corporate status.

On another front, to synthesize the manifold market and media signals for stock profiling, prior studies generally follow two paradigms: (i) Early-stage concatenation, which treats the two modalities equally to constitute a compact input at each time point; and (ii) Last-stage interaction, which separately models each type of sequential data, then blends them together at the final layer. For instance, Xu and Cohen [2018] spliced every day’s price and text vectors without discriminating the traits of different feature spaces. Wang et al. [2020] implemented case-by-case bullish/bearish stance classification of substantial stock reviews to expand discrete price indicators, whereas the pipelined processing is inevitably labor-intensive. In [Sawhney et al., 2020], text and price sequences are independently modeled followed by a bilinear combination layer. However, fine-granular matching signals between the two-sided factors are lost since they cannot communicate until the final step of prediction. Indeed, real-life stock market is rapidly responsive to new messages, and conversely drastic price changes are apt to trigger extensive discussions on social media [Foucault et al., 2016]. In Fig. 1, the dynamics of stock trading volume and the amount of relevant investment reviews are great synchronous especially when a mutation occurs. This phenomenon tallies with the Efficient Market Hypothesis (EMH) [Malkiel, 2003] and Behavioral Finance (BF) [Shiller, 2003], which state that stock prices can reflect all available information yet actual irrational markets are greatly affected by participant psychological factors. Hence a finer interaction method is required to better learn the interdependencies between market and media information.

Along these lines, in this paper, we formulate stock forecast as a learning to rank task and propose a Heterogeneous Interactive Snapshot Network (HISN) for review-enhanced investment recommendation. Our main contributions include:

- We represent stock media reviews for each trading day as a heterogeneous document graph incorporating additional knowledge. Besides, a hierarchical attentive propagation (HAP) framework with flexible node- and view-level aggregators is developed to assimilate important proximity features for more accurate semantic modeling.

- We propose twins-GRU, which tightly couples market and media parallel sequences in a cross-interactive fashion to catch time-evolving mutual impacts between different modalities. Being aware of time attenuation, significant hidden states of stock dynamic snapshots are integrated to learn stock profile embeddings for expected profit ranking.

- We conduct extensive experiments on English and Chinese datasets spanning various stock exchange markets to verify HISN’s applicability. The cumulative and risk-adjusted returns outperform state-of-the-arts by over 7.6% and 10.2%.

2 Related Work

Technical Analysis. Traditional stock prediction studies are based on analyses of price quote data. Many machine learning algorithms such as HMM, SVM and Random Forest were early used to capture volatility patterns [Nayak et al., 2015; Khaidem et al., 2016], whereas the hypothetical stochastic process may become stranded in handling highly non-stationary oscillations. Recently, deep neural networks have shown great prospects by formulating the stock forecast as a time series modeling problem, especially RNNs are mainstream technologies [Qin et al., 2017; Zhang et al., 2017; Sawhney et al., 2021a]. For instance, Qin et al. [2017] applied an attentive recurrent network to identify driving input and hidden states of stock dynamics. Zhang et al. [2017] injected a state frequency memory in LSTM to discover multi-frequency trading patterns based on Discrete Fourier Transform. Modern researches resorted to other neural methods to embed stock time series, such as adopting adversarial training to improve model generalization [Feng et al., 2019a], mining compositive volatilities based on Gaussian Transformer encoder [Ding et al., 2020] and gating causal convolutions with the idea of WaveNet [Wang et al., 2021].

Text Assisted Analysis. Except for numeric-based methods, another line of studies probes into exploring stock-related social media content such as financial news, tweets and earnings reports to learn stock environmental variables [Hu et al., 2018; Xu and Cohen, 2018; Sawhney et al., 2021b]. These works usually align the bullishness/bearishness of stock price movement with the sentiment of media messages. For instance, Hu et al. [2018] proposed bidirectional GRUs with hybrid attention on news and days for stock prediction. Wang et al. [2020] developed a stance detection and expert mining procedure to identify high-quality investment opinions from online discussion boards. Sawhney et al. [2020] adopted deep attentive learning for multipronged stock-related information. Sawhney et al. [2021b] applied BERT [Devlin et al., 2019] to encode intraday reviews and then used a time-aware LSTM to aggregate them considering temporal irregularities. Despite their success, the dynamic interdependencies between media and market information have not been handled delicately. Besides, in these works stock reviews are generally represented as the set or sequence of texts, which may be not optimal since inter-text relatedness cannot be effectively captured. As the chaotic online content usually involves various information and is characterized with high degrees of obscurity, the intersection of Finance and heterogeneous semantic information embedding presents a promising research avenue.

3 Methodology

3.1 Problem Statement

The stock recommendation can be formulated as a learning to rank problem. Let $\mathcal{S} = \{s_1, \ldots, s_N\}$ denotes a set of $N$ candidate stocks, where for stock $s_i$ on trading day $t$, there is an associated closing price $p_i^t$ and a 1-day return ratio $r_i^t = \frac{y_i^t - y_{i-1}^t}{y_{i-1}^t}$. To learn an optimal investment ranking $\mathcal{Y}^* = \{y_1^* > y_2^* \ldots > y_N^*\}$ of all stocks on $t$, each $s_i$ entails a series of dynamic snapshots with lag size of $\Delta T$ as the input tensor $\{\mathcal{F}_i^\tau, \ldots, \mathcal{F}_i^{\tau-1}\}$, where $\mathcal{F}_i^\tau = \{P_i^\tau, D_i^\tau\}$ represents its price and media review signals at historical day $\tau$. Briefly, the top-ranked stocks are expected to earn higher investment revenues in daily trading with such an ordering framework.
Fig. 2 presents an overview of our proposed HISN. In the following subsections, we first show how the heterogeneous document graph is built for investment reviews of each stock snapshot, and present the hierarchical attentive propagation method to coordinate multiple neighbor features (§3.2). We then describe twins-GRU, which couples duplex media and market signals across successive snapshots in the lookback period (§3.3). Lastly, we combine significant hidden states to learn stock embeddings for profit ranking in trading (§3.4).

### 3.2 Intra-Day Semantic Encoder

#### Heterogeneous Document Graph Construction

Given the set of media reviews $D^T_c$ discussing news events or opinions from $F^T_c$, we construct a heterogeneous graph which contains three kinds of nodes including texts $\mathcal{D}$, topics $\mathcal{O}$ and entities $\mathcal{E}$ to learn the unified semantic context. As illustrated in Fig. 1, topics are important indicators for revealing the nub of text content and public concerns. Therefore, we leverage LDA [Blei et al., 2003] to mine latent topics of the entire daily review document as graph nodes, and connect each text to $[\hat{r}]$ topic nodes with the largest relevance probabilities. To be consistent with previous research [Sawhney et al., 2021b], we generate the initial states of text nodes by averaging token level outputs from the final layer of BERT [Devlin et al., 2019]. Each topic node is represented by weighing word embeddings based on its probability distribution over the vocabulary. Moreover, stock reviews delivered consecutively usually concentrate on a few specific entities which complementarily portray the dynamic corporate status. In this regard, we utilize an entity linking tool TAGME\footnote{https://sobigdata.d4science.org/group/tagme/} to recognize the meaningful spots in each text, and map them to pertinent entities in Wikipedia. Whereas, the textual relationships formed only based on explicitly included entities may be sparse. Motivated by [Hu et al., 2019], we further attach edges between similar entity nodes to promote information diffusion in semantic learning. Specifically, we empower entity embeddings using word2vec trained on Wikipedia annotation corpus. Then a pair of entities are connected if their cosine similarity is above a threshold $\delta$ to finalize the graph structure $G^l_l$.

#### Hierarchical Attentive Propagation (HAP)

Once the graph topology is anchored, we perform node interactions to push forward semantic diffusion taking initial node states $[e^l_0]_{v \in \mathcal{V}^l} \in \mathbb{R}^{d_c \times |\mathcal{V}^l|}$ as input. Specifically, for target node $v$ at $l_{th}$ propagation layer, HAP convolves over its neighborhood following a hierarchical attention structure: $Attention_{node} \rightarrow Attention_{type} \rightarrow Attention_{node}$. Features of each type of $v$’s adjacent nodes are first transformed and fused together, then fed to the inter-type polymerization to enrich the representation of $v$.

#### Node-level Aggregator

We implement $Attention_{node}$ by a multi-head self-attention layer [Vaswani et al., 2017]. The primary motivation is that multiple node neighbors from the same relation type may complement each other in characterizing specific corporate profiles (e.g., implicit semantic correlation between entities of “Apple” and “iPhone”, and topics of “technological innovation” and “launching new products”). Specifically, by concatenating the $r$-type of $v$’s neighborhood including itself (denoted as $N_r(v)$) into a context matrix $E^l_{v,r} \in \mathbb{R}^{d_e \times d_r}$, where $d_r$ is the dimension of node states), we transform the Key-Value tuples as $K^l_{v,r,h} = E^l_{v,r} W^Q_h \in \mathbb{R}^{d_e \times d_r}$, and take the current state of target node $v$ as a Query tuple $Q^l_{v,r,h} = e^l_v W^Q_h \in \mathbb{R}^{1 \times d_r}$, where $h = 1, \ldots, M$ are indices of different attention heads, $W^Q_h \in \mathbb{R}^{d_e \times d_r}$ is a linear projection parameter. Here, we set $d_f = \frac{d_r}{3f}$ in order to maintain the dimension of the node states. After projection, the self-attention operator exploits element-wise dependencies to enhance the representation of $v$ (the query) by attending to the context nodes (keys):

$$
q^l_{v,r} = \text{Multihead}(Q^l_{v,r,h}, K^l_{v,r,h}, V^l_{v,r,h}) = \sum_{h=1}^{M} \text{Softmax}(Q^l_{v,r,h} K^l_{v,r,h}^T / \sqrt{d_f}) V^l_{v,r,h},
$$

where $\sum$ denotes the concatenation of output vectors from all $M$ heads, the scaled dot product measures the informativeness of different neighbors for contextual summarization.
Type-level Aggregator. With the aggregated semantic-specific node messages, we engage Attention$_{type}$, a type-level attention mechanism to capture the importance of different information aspects for updating a target node. Specifically, we compute the weighting coefficient of type $r$ to $v$ combining two terms—the matching similarity between $(g^l_{v,r}, e^l_v)$ determined by SUBMULT+$\mathcal{NN}$ function [Mou et al., 2016], and an adaptive type-level attention tensor $T_{1:|\mathcal{R}|} \in \mathbb{R}^{1:|\mathcal{R}| \times 4L}$ which serves to indicate the general semantic significance:

$$
\mu^l_{v,r} = T_r [g^l_{v,r} \parallel e^l_v \parallel (g^l_{v,r} - e^l_v) \parallel (g^l_{v,r} \odot e^l_v)]^\top,
$$

$$
\beta^l_{b,v,r} = \frac{\exp(\mu^l_{v,r})}{\sum_{r'=1}^{|\mathcal{R}|} \exp(\mu^l_{v,r'})},
$$

where $-$ and $\odot$ mean the difference and element-wise product.

The information of all types is fused as $\tilde{g}^l_v = \sum_{r=1}^{|\mathcal{R}|} \beta^l_{b,v,r} g^l_{v,r}$. Additionally, we place each HAP layer inside a residual connection block for updating current states of all graph nodes, which helps retain low-order features and facilitate gradient back-propagation. By this means, we obtain $e^{l+1}_v = \tilde{g}^l_v + e^l_v$.

After substantiating the graph correlations for $L$ rounds, we apply a soft gating transformation based on two multilayer perceptions (denoted as $f_1$, $f_2$) to integrate the features of all text nodes. Since each review plays a role in the day’s public voice while the salient ones contribute more explicitly, we jointly apply max-pooling and average functions to produce the final sentiment embedding of $\mathcal{F}^l$, defined as:

$$
\bar{e}_v = \sigma(f_1(e^l_v)) \odot \tanh(f_2(e^l_v)) \quad \forall v \in \mathcal{D},
$$

$$
e^{d,\tau}_v = \frac{1}{|\mathcal{D}|} \sum_{v \in \mathcal{D}} \bar{e}_v + \text{Maxpooling}(\bar{e}_1 \ldots \bar{e}_|\mathcal{D}|).
$$

3.3 Inter-day Temporal Encoder

Temporal Dependency Mining

How to effectively integrate the manifold information across dynamic snapshots in a lookback period is another key issue to discover the stock evolution regularities for future movement prediction. In light of EMH and BF literatures [Malkiel, 2003; Shiller, 2003], stock prices can timely react to the emotional signals from social media, and in turn can drive public’s expectation of new market situations. It is essential to learn evolving stock profiles by synthetically modeling such highly correlated dual-pronged factors. To this end, unlike most methods separately modeling or simply splicing the price and media information sequences of all snapshots (denoted as $P^t_\tau = [p^t_\tau; \Delta T^t_{\tau-1}], E^t_\tau = [e^t_\tau; \Delta T^t_{\tau-1}]$), we align them more finely along with the temporal trajectory of stock evolution in a cross-interactive fashion. Specifically, we design twins-GRU, which takes the advantage of two parallel GRU networks that not only focus on the sequence of each modality but interact with each other progressively.

As shown in Fig. 2, we represent the neural unit in twins-GRU cell as $\Psi$. Take price embedding for example, at every time-step $\tau$, the new coming input vector $p^t_\tau$, and two predecessor hidden states from both price and media sequences $e^c_{\tau-1}, e^c_{\tau-1}$ are fed into $\Psi$, such that the duplex information is simultaneously incorporated to characterize the dynamic context. Concerning different impacts of semantic and price factors, the $\oplus$ operator in $\Psi$ engages an attentive summation:

$$
u^p_{\tau-1} = \phi c^p_{\tau-1} + (1 - \phi)c^c_{\tau-1},
$$

where $\phi$ is a coefficient learned to rectify the fusing proportion of each kind of hidden signals, which is calculated by:

$$
\phi = \frac{\exp(\eta^p_{\tau-1})}{\exp(\eta^p_{\tau-1}) + \exp(\eta^c_{\tau-1})},
$$

$$
\eta^c_{\tau-1} = \nu c^c_{\tau-1} \parallel p^t_{\tau-1}, \quad \eta^p_{\tau-1} = \nu c^p_{\tau-1} \parallel p^t_{\tau-1},
$$

where $\nu c$ is the attention vector. The module $J$ in $\Psi$ further transforms the unified hidden state by a dense layer:

$$
\tilde{c}^p_{\tau-1} = \text{ReLU}(\tilde{W}c^p_{\tau-1} + b_c),
$$

where $W_c$, $b_c$ are weight and bias parameters. Then $\tilde{c}^p_{\tau-1}$ and $p^t_\tau$ are fed into a canonical GRU unit to update the price memory $c^p_{\tau}$. Analogously we can acquire new memory $c^c_{\tau}$.

Temporal Aggregation

The last part of the framework fuses the entire information stream to learn stock’s compact representation for profit ranking. Intuitively, historical transaction days have varied predictive influences, where days with profound reviews and drastic price changes usually excite investors’ behaviors more intensively. Meanwhile, the study of Hawkes process in finance shows that the impact of emerging market variables dwindles over time [Bacry et al., 2015]. Thus we conduct day-level aggregation by taking both feature salience and time attenuation into account. The weight of the day $\tau$ is specified as:

$$
\theta_\tau = \frac{\exp(e^p_{\tau}, Wc^p_{\tau-1})}{\sum_k \exp(e^p_{k}, Wc^p_{k-1})} \times (1 + \epsilon \exp(-\gamma \Delta o_\tau)),
$$

where $W$ is a transformation matrix, $\epsilon$ is excitation coefficient, $\gamma$ is a decay scaler and $\Delta o_\tau$ is the lag of $\tau$ to latest time. The overall price sequence is then summed as $p^t_\tau = \sum_\tau \theta_{\tau}c^p_{\tau}$. The condensed semantic embedding can be distilled in a very similar way. Finally, we tailor the profile of stock $s_i$ as the concatenation of abstracts $z_{s_i} = [\tilde{p}_{i;\tau} \parallel \bar{e}_{i;\tau}]$, and all candidate stocks form the tensor $Z_t \in \mathbb{R}^{N \times d}$.

3.4 Model Training

Following [Sawhney et al., 2021b], we pass $Z_t$ to a dense layer with the activation of LeakyReLU to acquire stock expected return ratios $r^t_{[1:N]}$ on day $t$. Based on joint optimization of point-wise regression and pairwise ranking loss, we minimize the discrepancy between $r^t_{[1:N]}$ and ground-truth $r^t_{[1:N]}$ meanwhile maintaining the relative order of stocks:

$$
\mathcal{L}_r = \sum_{i=1}^N ||r^i_t - r^j_t||^2 + \lambda \sum_{i=1}^N \max(0, - (r^i_t - r^j_t) (r^i_t - r^j_t)),
$$

where $\lambda$ is a weighting coefficient. Stocks at the top of the forecast order are selected for investment recommendations.
Interactive modeling

4 Experiments

4.1 Dataset and Experimental Setting

We validate HISN on two real-world stock forecast datasets: US S&P 500 [Xu and Cohen, 2018] contains 109,915 English Twitter data between Jan. 2014 and Jan. 2016, related to 88 high-capital-size stocks from popular S&P-500 Composite Index spanning 9 industries in NASDAQ and NYSE markets. Tweets and stocks are associated using the regex query |. Samples of the first 19 months are split for training, those of the last 3 months for testing and the others for validation. 

Ashare&HK [Huang et al., 2018] collects 90,361 news headlines from major financial websites in Chinese during Jan. and Dec. 2015. It targets at 99 top-traded A-share and HK stocks spanning Shanghai, Shenzhen and Hong Kong Exchange markets. Samples are chronologically divided, leaving us with a date range of the first 8 months for training, the last 3 months for testing, and the rest for validation in chronological order.

We collect split-adjusted open, high, low, close and volume indicators from professional Wind-Financial Terminal to constitute daily price vectors of all stocks. We keep the same setting as in previous works [Sawhney et al., 2021b] and leverage a consecutive 5-day lookback trading window (i.e., 5 daily snapshots) to generate each sample. To build heterogeneous document graphs, we set the number of topic nodes $O$, the per-text related topics $\hat{d}$ and the similarity threshold between entities $\delta$ to $\{15, 2, 0.5\}$. Hence, 15,448 and 6,696 entity nodes are exploited for S&P 500 and Ashare&HK datasets, respectively. The word embeddings to initialize graph nodes are 300-dimensional. We set $L = 2$ graph layers and $M = 4$ attention heads in HAP. The hidden state size of twins-GRU is 64. The loss weighing factor $\lambda = 4$. We apply dropout [Srivastava et al., 2014] with the ratio of 0.3 at the end of each layer to mitigate overfitting. Parameters are tuned using Adam optimizer [Kingma and Ba, 2015] on a GeForce RTX 3090 GPU for 50 epochs, the batch size is 16 and learning rate is 1e-3. Each experiment is repeated 5 times. We report average MRR, nDCG@5 to assess the model’s ranking ability. To compare the practical revenue of stock recommendation, we follow [Feng et al., 2019b; Sawhney et al., 2021b] and calculate the Sharpe ratio (SR) and cumulative investment return ratio (CIR) by simulating a daily buy-hold-sell trading strategy. That is, when the market closes on day $t-1$, the trader buys $\kappa$ stocks with the highest expected returns, then sells the bought shares on next day’s close market. Specifically, $\text{CIR} = \sum_{t \in S^I} \frac{p_t^{\text{up}} - p_t^{\text{down}}}{p_t^{\text{down}}}$, where $S^I$ denotes selected stocks in portfolio. SR characterizes how well the earned return $R_c$ compensates a trader for the borne risk, i.e., $SR = \frac{R_c[i]}{\text{std}[R_f]}$, where $R_f$ is the risk-free return.

4.2 Performance Evaluation

We compare HISN’s profitability with different lines of stock prediction methods. The evaluation results are shown in Table 1, which reveal several findings: 1) Overall, RL and ranking approaches excel most traditional price regression and classification methods, illustrating the efficacy of formulating stock forecast as a learning-to-rank paradigm that directly hammers at optimizing investment profits. 2) Exploiting web
information beyond numeric prices (e.g., MFN, MAN-SF, FAST) can usually impose positive effect on judging the stock tendency. It indicates the desirability of mining dependencies of multifaceted affecting signals confronting the high non-stationarity of financial markets. 3) The test periods of S&P 500 and Ashare&HK datasets cover standard and turbulence (bearish) market conditions respectively. Facing the diverse scenarios, our proposed HISN can consistently outperform compared baselines in terms of both cumulative and risk-adjusted returns. We attribute the performance gains to two major reasons. First, HISN learns the joint effect of market and media signals in a cross-interactive way being aware of their temporal interdependencies. Further, by enriching the representation of chaotic reviews with additional knowledge including topics, entities and Wikipedia descriptions, latent semantic clues can be absorbed for more comprehensive stock profiling. Next, we move to explore how different components and hyper-parameters affect the capability of HISN.

4.3 Ablation Study

Effect of Absorbed Information. We first evaluate the effectiveness of different information sources on stock profiling. From Table 2, significant improvements are noted by blending the textual signals beyond unimodal price features. In addition, we find that both topic and entity nodes contribute to the performance by facilitating the semantic learning of online media texts that usually have high degrees of obscurity and insufficiency. This is mainly because topics can depict the nub of review content, while entities and their neighbors portray more detailed meanings of key elements in each text.

Effect of Model Variants. We then investigate ablation impacts of different components in HISN. Fig. 3a shows the results of applying plain GCN [Kipf and Welling, 2017] to convolve entire review document graph by concatenating feature spaces of all types of nodes (w/o HetG), and discarding node- and semantic-level attentions (w/o NodAtt, SemAtt). We find that ignoring the heterogeneity of node relationships seriously hurts the effect of graph embedding. Meanwhile, the hierarchical attention mechanism is conducive to the model’s performance. This proves the advantage of HAP in learning the significance of different nodes and semantics to a specific target thereby reducing the propagation of noise information.

We further conduct ablation research on the inter-day temporal encoder, including decoupling parallel GRUs to model price and semantic sequences in isolation (w/o Cross), discarding the attention on modalities in Ψ (w/o ModAtt) and directly feeding the last-day’s hidden state for ranking (w/o TimeAtt). From Fig. 3b, interactively modeling the transition of media and market signals clearly benefits stock profiling and the estimation of expected investment revenues. Besides, weighing different kinds of information and historical time points is helpful to the modeling and integration of stock dynamic trajectory, which may be because it highlights salient influential variables and reduces false overreactions.

4.4 Parameter Analysis

We examine stock ranking ability regarding different lengths of lookback days. From Fig. 4a, the performance first increases with larger ΔT, then begins to decline due to the inclusion of outdated media and market signals. Fig. 4b shows the impact of applying varied numbers of graph layers for semantic embedding, where the optimal setting is L = 2. This is reasonable because as nodes incrementally absorb information from high-order neighbors, the graph may become oversmooth thus may deteriorate the accuracy of distilled text signals. We further explore the change of profitability w.r.t. the number of selected stocks κ. Fig. 4c shows that HISN consistently performs better than FAST (the SOTA stock ranking model exploiting media reviews). This demonstrates HISN’s adaptability to trading strategies with different risk appetites.

4.5 Case Study

We look closer to interpreting the attention mechanism of our model in semantic and temporal modeling. The upper part of Fig. 5 visualizes the node- and type-level weights HAP assigned to a randomly sampled review of stock $ABB. The text itself and entity relationship receive higher attentions. It means that intrinsic text content and associated knowledge of
included entities (e.g., smart home, Philips which reveal fine-grained clues of corporate business innovation and partnership) are most informative to characterize the review semantics. The lower left of Fig. 5 elaborates the attentions paid to diverse modalities (price and media memory states) given the next input. It can be seen that the duplex information act together during the modeling of stock evolution profile. In general, previous semantic state plays a major role when factoring in next media input, while the situation reverses when reconciling new price signals. Moreover, we summarize the top-5 types of reviews that have attracted the most attention. As shown in the lower right of Fig. 5, equity analysis and release of corporate operating conditions would likely induce public’s sentiment toward underlying stock variations more significantly, and thereby driving investment reactions.

5 Conclusion

We propose HISN, a heterogeneous interactive snapshot network for stock profiling and recommendation. We first devise a flexible HAP framework for modeling media investment reviews of each stock’s dynamic snapshot, which incorporates additional information to capture implicit semantic relatedness. We also propose a twins-GRU that couples market and media signals in a cross-interaction fashion so as to catch their fine-granular interdependencies along the temporal trajectory. Extensive experiments demonstrate that HISN can identify evolving stock status to benefit the profitability of investment recommendation. In the future, we shall explore the correlation of enterprises and cold start problem of new stocks.

Acknowledgements

This research is supported by the National Key Research and Development Program of China (No. 2021ZD0111202), the National Natural Science Foundation of China (No.62176005, 91846303), PKU-Haier “Smart Home” Qingdao Joint Innovation Lab, Project 2020BD002 supported by PKU-Baidu Fund, and Research Funds for NSD Construction, University of International Relations.

References


