

Padrão Reconhecimento

diário página inicial www.elsevier.com/locate/patco g



Previsão de séries de tempo financeiro com rede neural de gráfico multi-modalidade



Dawei Cheng^{a,*}, Fangzhou Yang^b, Sheng Xiang^c, Jin Liu^b

^a Departamento de Ciência e Tecnologia da Computação, Universidade de Tongji,

Xangai, China^b AI Lab, Emoney Inc, Xangai, China

^c Center for Artificial Intelligence, University of Technology Sydney, Sydney, Austrália

artile info

Histórico do artigo:

Recebido em 15 de dezembro de 2020

Revisado em 22 de julho de 2021

Aceito em 31 de julho de 2021

Disponível online em 2 de agosto de 2021

Keywords:

Rede neural de gráficos

Atenção gráfica

Aprendizado profundo

Investimento quantitativo

abstract

A análise da série de tempo financeiro desempenha um papel central na cobertura dos riscos do mercado e na otimização das desilusões de investimentos. Esta é uma tarefa desafiadora, pois os problemas são sempre acompanhados por fluxos multi-modalidade e efeitos de lead-lag. Por exemplo, os movimentos de preços das ações são reflexos de estados de mercado complicados em diferentes velocidades de difusão, incluindo séries históricas de preços, notícias da mídia, eventos associados, etc. Além disso, o setor financeiro exige que os modelos de previsão sejam interpretáveis e compatíveis. Há, antes, neste artigo, a WE propõe uma rede neural de gráfico multimodal (MAGNN) para aprender com esses insumos multimodais para previsão de séries temporais financeiras. A rede de gráficos heterogênea é construída pelas fontes como nós e relações em nosso gráfico de conhecimento financeiro como bordas. Para garantir a interpretação do modelo, utilizamos um mecanismo de atenção em duas fases para otimização conjunta, permitindo que os usuários finais investiguem a importância da modalidade interna e fontes interseções. Experimentos extensivos em conjuntos de dados do mundo real demonstram o desempenho superior da MAGNN na previsão do mercado financeiro. Nosso método fornece aos investidores uma opção rentável e interpretável e permite que eles tosem decisões de investimento informadas.

© 2021 Elsevier Ltd. Todos os direitos reservados.

1. Introdução

A capitalização do mercado financeiro das com-panies listadas nos EUA chega a 30 trilhões de dólares em 2019, representando mais de 1,5 vezes o Produto Interno Bruto (PIB) nos Estados Unidos [1]. Neste mercado massivo, mas volátil, a previsão do movimento dos preços das ações é muito importante tanto para as instituições financeiras quanto para os in-vestors. De acordo com a eficiente hipótese de mercado (EMH) [2], idealmente, os preços das ações refletem todas as in emissões disponíveis em um mercado ef-iciente, que inclui preços históricos, notícias, eventos, etc. No entanto, em uma situação do mundo real, diferentes ações que respondem a diferentes eventos são não intuitivas e não sincronizadas. Assim, é desafiador modelar esse fenômeno intrincado, chamado de efeito lead-lag [3], em uma estrutura de previsão de séries temporais.

A indústria financeira tem pesquisado modelos de previsão de preços desde o início do século XX [4] e aperfeiçoou essas tecnologias desde então, investindo milhões de dólares em esse processo. Os métodos quantitativos tradicionais dependem de dados históricos de preços de séries temporais para a previsão do movimento dos preços das ações [5,6]. Esses modelos visam reduzir a estocástica e capturar pat-

<https://doi.org/10.1016/j.patco.2021.108218> 0031-3203/©
2021 Elsevier Ltd. Todos os direitos reservados.

terns extraindo indicadores técnicos significativos [7] e/ou la-tenda características [8]. Ultimamente, com o desenvolvimento de mídias sociais e tecnologias de processamento de linguagem natural, notícias não estruturadas têm sido aproveitadas para melhorar a capacidade do modelo de previsão [9]. Mas essas tecnologias não captam relações internas entre os laços equi, o que limita seus potenciais para o modelo de previsão. Para ex-amplio, o recurso de nível de termo de um evento "Qualcomm arquiva processo contra a Apple" não pode diferenciar o appellor "Qualcomm" e appellee "Apple", por isso é difícil inferir o movimentações de preços correspondentes das equities relacionadas, Qualcomm e Apple Inc.

Recentemente, pesquisadores [10] tendem a melhorar a representação de informações de mercado extraindo tuplas de eventos estruturais e indi-cators (ou seja, indicador de sentimento) [11] de notícias da mídia. A ideia principal é aprender representações distribuídas de que eventos semelhantes ou notícias de sentimento semelhantes poderiam ter características semelhantes. Essas características são então vinculadas a empresas listadas e integradas com séries temporáveis históricas para previsão de preços [12]. Mas dois eventos semelhantes podem ser silenciosos sem relação, como "Steve Jobs deixa a Apple" e "David Peter deixa o Starbucks". Para superar isso, os estudos [13,14] empregam informações externas a partir de gráficos de conhecimento (KG) no processo de aprendizagem de recursos [15]. Em seguida, os dois eventos acima

* Autor Correspondente.

Endereço de e-mail : dcheng@tongji.edu.cn (D. Cheng).

podem ter representações diferentes- ent de acordo com as diferenças semânticas no KG, porque Steve Jobs é o fundador da Apple, enquanto David Peter é mais como ser um cliente no Starbucks.

No entanto, os movimentos de preços das ações no mercado financeiro não dependem apenas de eventos individuais de si mesmo, mas também relacionados às conexões de outras ações [16]. Essas multissídulas, incluindo séries temporais numéricas, textos não estruturados e gráficos relacionais, contribuem de forma diferente como efeito de sinergia no movimento dos preços. Por exemplo, um evento "Qualcomm suits against Apple" também influenciará outros jogadores (ou seja, concorrentes, empresas upstream e downstream) do mercado de smartphones em diferentes difusões-velocidades, como Samsung, Foxconn e Google, etc. Prever efetivamente os preços das ações relacionadas a partir dos efeitos de lead-lag é desafiador, devido à incompletude do domínio financeiro knowledge e intrínseca sequencial Padrões.

Portanto, neste artigo, propomos um modelo de rede neural de gráficos multi-modalidade para prever os movimentos de preços por fontes de relações de lead-lag, incluindo preços históricos, eventos de mídia, e conhecimento correspondente de KG. Em particular, primeiro extraímos relações de entidades vinculadas a partir de notícias cruas e, em seguida, armazenamos-nas em nossos gráficos de conhecimento financeiro (FinKG). Em seguida, propomos um trabalho de atenção numérica de grafo para aprender a representação unificada de séries temporais de destino, em que as fontes multi-modalidade são definidas como nós de origem e o patrimônio previsto como nó alvo. Utilizamos um mecanismo de atenção em duas fases (atenção interna e intersameria) para inferir os padrões sequenciais internos e as relações de lead-lag entre fontes. O mecanismo de atenção da modalidade interna foi projetado para automatizar- aprender diferentes contribuições de fontes estruturadas de gráficos para o nó alvo dentro de cada entrada de modalidade. Enquanto a atenção intermoda é proposta para aprender pesos entre diferentes modali-las dinamicamente para uma previsão decente de movimento de preços de alcatrão- obter nódulos, como modalidade diferente contribui de forma diferente em período de tempo diferente. Posteriormente, as características informativas aprendidas são alimentadas em camada de previsão para previsão de movimento de preços. Experimentos exten-sive em dados reais de mercado mostram a eficácia do nosso método e a interpretação do mecanismo de atenção bidáculo proposto.

Em poucas palavras, a principal contribuição deste artigo inclui:

- Formalizamos o problema dos efeitos de lead-lag na previsão de séries de tempo financeiro e identificamos seus desafios únicos decorrentes de aplicações reais do setor financeiro.
- Propomos uma nova rede neural de gráficos multi-modalidade (MAGNN) para aprender os efeitos de lead-lag para a previsão de tempo financeiro, que preserva informações informativas do mercado como insumos, incluindo preços históricos, texto de notícias brutas e relações em KG. Para nosso melhor conhecimento, este é o primeiro estudo a explorar os efeitos de lead-lag incorporando fontes informativas em uma estrutura neural de gráfico uni-fied para movimentos de preços predição.
- A fim de seguir processos altamente regulados no in-dustry financeiro, projetamos e implementamos um mecha-nismo de atenção em duas fases para inferir a interpretabilidade tanto da modalidade interna e fontes inter-modalidades. Também validamos a eficácia das tecnologias de atenção projetadas na aprendizagem dos padrões internos se-quential e das relações entre fontes de lead-lag por meio de estudos empíricos.
- Resultados experimentais extensos em 3714 ações demonstram o desempenho superior do nosso método proposto. Além disso, nosso modelo foi implantado em um grande provedor de serviços financeiros da China e validamos seu desempenho de profabilidade e interpretabilidade em cenários do mundo real. Os códigos-fonte serão lançados em breve.

2. Preliminares

Nesta seção, introduzimos o pano de fundo do efeito lead-lag e o processo de construção de gráficos heterogêneos.

2.1. Efeito de lag de chumbo

Em um mercado eficiente e informacional, a movimentação de preços das ações pode ser considerada como a reação de eventos financeiros ou notícias [17]. No entanto, quando um novo evento chegou ao mercado de ações, os preços de algumas ações responder mais rápido do que outros. Este fenômeno de movimento de preços correlacionado, porém assíncrono, é referido como efeito lead-lag [3]. Por exemplo, em Fig. 1, quando um novo evento ("Qualcomm suites against Apple") chegar ao mercado, ele não só trará fluctua preço de "Qualcomm" e "Apple", mas também influenciará empresas a montante e downstream, como a Samsung (fornecedora e principal concorrente da Apple no mercado de smartphones) e a Foxconn (fabricante de homens da Apple). Mas seus movimentos de preços são assíncronos porque a velocidade de difusão do evento é diferente em relação a diferentes laços de nti. Portanto, é uma tarefa desafiadora aprender com essa relação de lead-lag no mercado financeiro.

2.2. Construção de gráfico heterogêneo

No MAGNN, o gráfico heterogêneo multi-modalidade estende o gráfico heterogêneo convencional [18] com fontes multi-modalidade. Os nós gráficos são divididos em seis tipos (fonte, notícias, eventos, mercado, ponte e nós de destino) com três tipos in-puts (séries temporais numerais, textos de mídia e relações). Damos a definição da seguinte forma:

Definição 1ºgrama heterogêneo. Um gráfico heterogêneo é de- notado como

$$g = (V_T, V_S, E),$$

onde V_T representa o conjunto de nós de destino, V_S denota o conjunto de nós de origem e E é o conjunto de links que se conectam entre nós.

Definição 2Ssaia de origem. V_S estão associados a diferentes modais por uma função de mapeamento: $V_S \rightarrow$, onde denota o conjunto de modalidades, incluindo dados de mercado numeral, textos de mídia e rela-tions.

Definição 3Ss de alvo. V_T são nossas ações previstas no gráfico, que foi projetado para receber e agregar mensagens de outros nós através de links direcionados.

Definição 4Ponto deBridge. denota os nós conectados entre fontes multi-modalidade e nós de destino. Eles são extraídos do gráfico de conhecimento de domínio FinKG.

Definição 5 Nódulos atribuídos. incluem notícias, eventos e nodes de mercado, que só se conectam às suas empresas sujeitas.

As entradas multi-modalidade são parecem como nós em um gráfico heterogêneo, no qual podem passar mensagens para outros nós através de links. Uma empresa pode ser um nó de origem, alvo ou ponte, enquanto os nós atribuídos (notícias, eventos, mercado) só se conectam à sua empresa atual. Por exemplo, o nó de mercado (M) da Apple só conecta o nó de destino Apple, como mostrado na Fig. 1.

Definição 6. Bordas (E) são um conjunto de links que se conectam entre nós, que incluem bordas direcionadas e não direcionadas. A relação entre as empresas (nós de origem, alvo ou ponte) são di-rected, que seta do sujeito ao objeto. A conexão entre a empresa e seus nós atribuídos não são direcionadas.

Fig. 1 mostra um exemplo em execução de entradas heterogêneas de gráficos e multi-modalidades. Quando um evento (ou notícia) "Processos da Qualcomm contra a Apple" chegam ao mercado, extraímos a relação entre seu assunto (Qualcomm) e objeto (Apple), e estabelecemos uma borda (terno contra) direcionado do sujeito (Qualcomm) para o ob-ject (Apple). Então, se

dados de mercado, como mostrou a Fig. 2 a. Em seguida, os insumos multi-modalidade são alimentados na camada de atenção de gráficos da modalidade interna (InnGAT) em paralelo, em que cada entrada de modalidade é aprendida pela InnGAT independentemente sobre o gráfico heterogêneo. A atenção de origem interseção (IntSAT) tira a saída do InnGAT e aprende representações

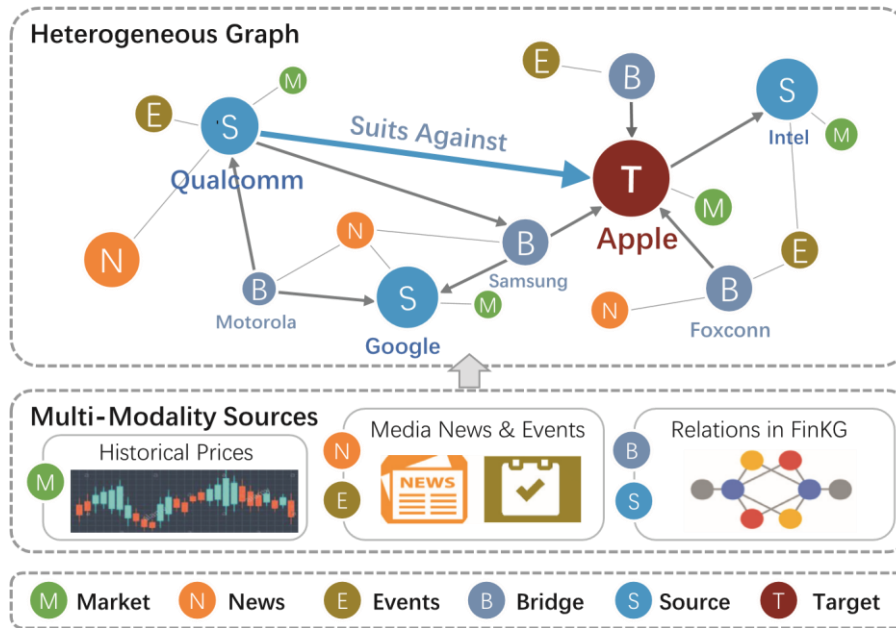


Fig. 1. Uma ilustração de insumos multi-modalidade e gráfico heterogêneo.

quisermos prever os movimentos de preços da Apple no dia seguinte, definimos-os como nó-alvo e extraímos insumos multi-modalidade em conformidade, o que inclui o evento (notícias) incorporações de semman, nó de fonte vinculado (Qualcomm) e seus preços históricos, dados de mercado do nó-alvo (Apple) e as relações (incluindo bordas e acenos de ponte), como Samsung, Motorola, Foxconn, etc.) de não-acenos de origem e alvo. Por exemplo, em telefone inteligente mar-ke, a Samsung é concorrente da Apple e um cliente a jusante do Google. Assim, é um nó de ponte do Google e da Apple neste cenário. Como mostrado na Fig. 1, normalmente, cada evento (notícia) é accomplished com um nó de origem e um nó alvo. Construímos gráficos heterogêneos por insumos multi-modalidade dos nós vinculados e relações correspondentes na FinKG. Os métodos detalhados de re-lação e construção de gráficos são apresentados na Seção 3.1. Essas entradas informativas são então alimentadas em MAGNN para aprendizado conjunto e inter-pretoble.

3. Metodologia

Nesta seção, primeiro introduzimos o quadro geral e os insumos multi-modalidades de nossa abordagem proposta e, em seguida, apresentamos atenção numérgria de modalidade interna e atenuação da fonte interde modalidade, respectivamente. Por fim, introduzimos a previsão de metas de trabalho líquido e otimização de modelos.

3.1. Estrutura do modelo e insumos

Fig. 2 mostra o quadro geral da rede neural de gráfico multide modalidade proposta para previsão de séries temporais financeiras. Construímos o gráfico heterogêneo primeiro pelos eventos, notícias, relações em KG e os

de alta ordem de todas as modalidades. Finalmente, o os recursos aprendidos são alimentados em um feed-forward e rede de classificação para previsão de alvos.

Na implementação, empregamos um BERT [19]¹ pré-treinado como nosso modelo de incorporação de notícias, e multamos o modelo BERT do nosso corpo de notícias financeiras em larga escala. Para a extração de tupla de eventos, aproveitamos o OpenIE [20] amplamente utilizado e utilizamos a incorporação de tuplas estruturadas aprendidas pela rede neural tensor [21] como característica do evento. Na construção da FinKG, empregamos o OpenNRE² para extrair relações de textos de notícias massivas e armazená-las em nosso gráfico de conhecimento FinKG. Se a entidade de um evento (ou notícia) é uma empresa listada, nós os marcamos como o nó de origem. As demais entidades são denotadas como nós de ponte no gráfico de conhecimento. Quando um conjunto de eventos atingiu o FinKG, extraímos os nós adjacentes e as re-lações correspondentes das entidades mencionadas como gráfico base. Então, marcamos as ações previstas como nós alvo. Posteriormente, as notícias, eventos e dados de mercado estão ligados a cada entidade e, finalmente, formam o gráfico heterogêneo, como mostra Fig. 2 a. Atualizamos o gráfico heterogêneo a cada dia de negociação.

3.2. Atenção do gráfico de modalidade interna

Dado cada característica de entrada da modalidade e o gráfico heterogêneo construído, a atenção num gráfico de modalidade interna é projetada para prop-ágata e agregar informações de nódulos de origem ao alvo nodo. Como mostrado na Fig. 2 b, as entradas do InnGAT incluem o pré-...

S φ τ π incorporações treinadas do nó de origem e_i e do nó alvo e onde $\varphi \in \{n, e, p\}$ denota o tipo

¹ <https://github.com/google-research/bert>

² <https://github.com/thunlp/OpenNRE>

de modalidade e $eu \in N_s$ indica os vizinhos do nó S . N_s é o conjunto de vizinhos.

Projetamos projeções em duas fases para mapear os insumos multi-modalidade em representações latentes, denominada projec-tion de origem e projeção de alvo. Eles são parametrizados por peso ma-trix $WS^\phi \in \mathbb{R}^{d_h \times d_\phi}$ e $W_T^\phi \in \mathbb{R}^{d_h \times d_t}$, respectivamente. d_ϕ , d_t e d_h de-note a dimensão da incorporação do nó de origem, a incorporação de nó de destino e as características ocultas projetadas. Em seguida, um mecanismo de atenção compartilhada é introduzido para calcular coefientes de atenção de nível de nó, que é parametrizado por um vetor de peso $\vec{\alpha}^* \in \mathbb{R}^{2d_h}$. Fi-nally, O coeficiente de atenção da modalidade interna para ϕ tipo de fonte

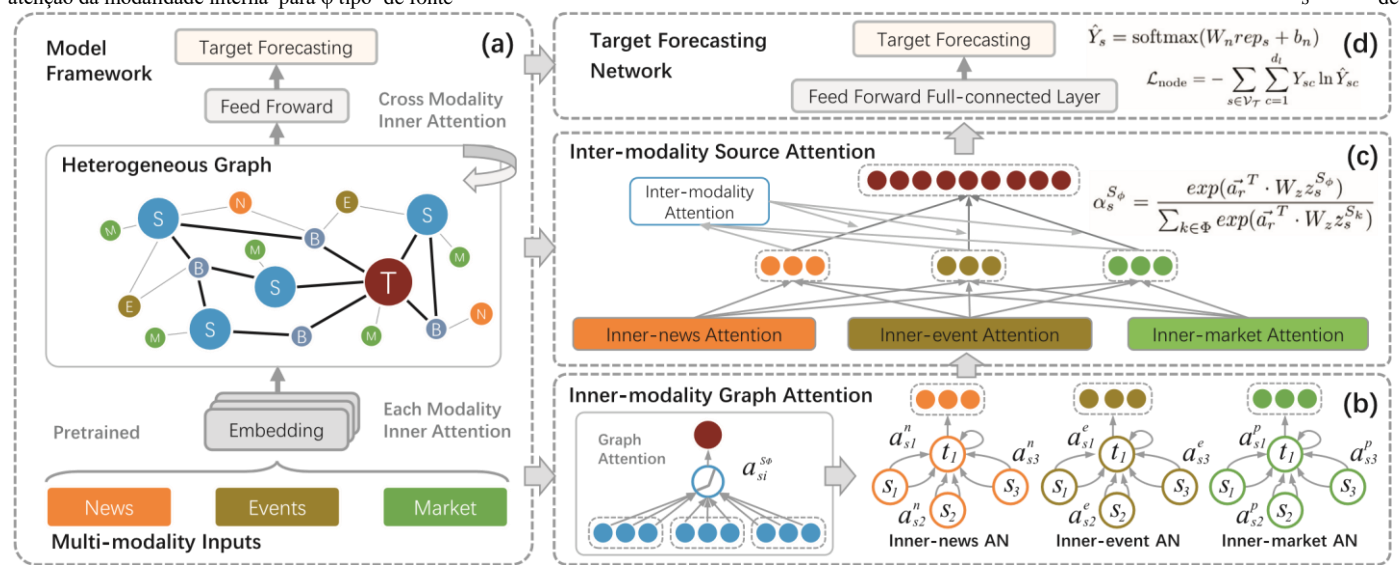


Fig. 2. O quadro geral da rede neural de gráfico multide modalidade proposta. Inclui entradas multi-modalidade, camada de atenção de gráficos de modalidade interna, camada de atenção de origem entre modalidades e a rede de previsão de destino. No gráfico heterogêneo, o símbolo de S, B, E, N, M, T denota a fonte, ponte, eventos, notícias, mercado e nós alvos, respectivamente.

entre o nó de origem i e o nó *de* destino s é formulado como:

$$\alpha_{si\varphi} = \frac{\exp_{\in N_s^{\varphi}}(\bar{a}^{\top}[W^{\phi} \cdot e^{\top}||W^{\phi} \cdot \bar{a}_{\phi}^{\top}[W_I^{\phi} \cdot e_s^{\top}||W_S^{\phi} \cdot \text{(LeakyReLU}(\text{(LeakyReLUe)}))_{kk}])}{\varphi^T \text{ }_{sSej} \text{ }_{i} \text{ }_{\varphi} \text{ }_{\text{S}}}, \quad (1)$$

onde \top uma ϕ representa a transposição \vec{a}_ϕ e \parallel é a operação de concatenação.

Posteriormente, calculamos as características de saída do nó de destino S para o tipo de modalidade φ como média ponderada das características ocultas de origem com função sigmoid, que é formulada como:

$$\phi = \sigma \left(\sum_{i \in \mathcal{N}_s^{\phi}} \alpha_{si}^{S_{\phi}} W_S^{\phi} e_i^{S_{\phi}} \right) \mathbf{z}_s \quad (2)$$

onde WS^ϕ denota os pesos aprendidos, e σ é o func sigmoid

s^φ . N_s indica o conjunto vizinho de nó S . z_s denota o recurso out-put do InnGAT para o nó S na modalidade φ . Na implemen-tation, estendemos o InnGAT com atenção multi-cabeça, a fim de estabilizar o processo de aprendizagem.

3.3. Atenção de fonte interseção

A atenção de origem interde modalidade (IntSAT) é proposta para agregar as informações de fontes multide modalidades para representação de nó alvo. Como ilustrado em Fig. 2 c, as entradas

$$S = \Phi \text{ de}$$

Target Forecasting $\hat{Y}_s = \text{softmax}(W_n \text{rep}_s + b_n)$ (d)

Feed Forward Full-connected Layer

$$\mathcal{L}_{\text{node}} = - \sum_{s \in V_T} \sum_{c=1}^{d_l} Y_{sc} \ln \hat{Y}_{sc}$$

Source Attention

$$\alpha_s^\phi = \frac{\exp(\vec{a}_r^T \cdot W_z z_s^{\phi})}{\sum_{k \in \Phi} \exp(\vec{a}_r^T \cdot W_z z_s^{S_k})}$$

(c)

IntSAT para nó de destino incluem os recursos de saída z_s de In-nGAT de todas

IntSAT para nó de destino incluem os recursos de saída z_s de In-nGAT de todas as modalidades, onde $\varphi \in \{n, e, P\}$. Na rede de atenção de origem intermodal, uma transformação linear compartilhada param-eterizada por uma matriz de peso $W_z \in \mathbb{R}^{d_r \times d_z}$ e um *atten* multi-fonte- O mecanismo de ponderação parametrizado por um vetor *de* peso $r \in \mathbb{R}^{d_r}$ são em-*ployed* para calcular coeficientes de atenção de origem, respectivamente. d_z em-

s° acede a dimensão de z_s e d_r é a dimensão das características ocultas trans. Matematicamente, o coeficiente de atenção do tipo de modalidade ϕ para o nó de destino pode ser formulado por:

$$\sum_{k \in \Phi} \exp(\tilde{a}_r^\top \cdot \alpha) = \sum_{k \in \Phi} \exp(\tilde{a}_r^\top \cdot W_z z_s) \quad (3)$$

s^φ onde a_r e W_z são os pesos aprendidos, e α_s denota o coeficiente de atenção do tipo modalidade φ .

Finalmente, construímos a representação do *representante* do nó-alvo pela concatenação das características projetadas ponderadas pela atenção das três modalidades, formuladas como:

$$rep_s = [\alpha_{sS_n} W_{z_{sS_n}} \parallel \alpha_{sS_e} W_{z_{sS_e}} \parallel \alpha_{sS_p} W_{z_{sS_p}}], \quad (4) \text{ onde o } \alpha_{sS^n},$$

α_{ss} e α_{ss}^p são o coeficiente de atenção do IntSAT. W_z denota os pesos aprendidos e *os representantes* significa a represtração de saída da rede de atenção de fonte interde modalidade.

3.4. Rede de previsão de alvos e otimização

Dada a representação aprendida do nó alvo da InnGAT e da IntSAT, então empregamos uma rede neural rasa para a previsão de preços de piche, como mostrado na Fig. 2 d. Em particular, nós nos formamos tarde da tarefa de previsão como um problema de classificação, o que significa que dividimos a tendência dos movimentos de preços em três categorias { *para cima*, *neural*, *para baixo* }. Detalharemos as configurações na seção experiment. A rede de previsão consiste em duas camadas de conexão completa e uma camada softmax. Eles são definidos como:

$$Y_{...s} = \text{softmax}(\text{NN}_f(W_n \text{rep}_s + b_n)) \quad (5)$$

onde NN_f denota uma rede neural rasa com duas camadas de conexão completa. $W_n \in \mathbb{R}^{d_s \times d_l}$ e $b_n \in \mathbb{R}^{d_l}$ são a matriz de peso e o viés, respectivamente. d_l é o número de categorias-alvo. Neste pa-per, definimos o $d_l = 3$.

Finalmente, definimos a função de perda do modelo proposto pela entropia cruzada da probabilidade na camada de saída como abaixo:

$$L_{\text{alvo}} = \sum_{com \in \mathbb{V}_T} \sum_{c=1}^{d_l} -Y_{sc} \ln Y_{...sc} \quad (6)$$

onde Y_{sc} é o rótulo de verdade moída da categoria cth movimento para ações, que é marcado como 1 para os movimentos de preço "up", 0 para o "neural" e -1 para o movimento "para baixo", respectivamente. \mathbb{V}_T

denota o conjunto de nódulos-alvo.

Nossa rede neural de gráfico multi-modalidade proposta pode ser treinada de forma ponta a ponta minimizando a perda de entropia transversal de classificação. Teoricamente, podemos otimizar o modelo pelo processo de descida de gradiente estocástico padrão. Na prática, nós em-ploy algoritmo Adam [22] como o otimizador do nosso modelo. Nós definimos a taxa inicial de aprendizagem para 0,001, e o tamanho do lote para 64 por default.

4. Experimentos

Nesta seção, realizamos extensos experimentos para validar a eficácia de nossas tecnologias propostas. Introduzimos primeiro a aquisição de dados e as configurações experimentais e, em seguida, reportamos o resultado de cada experimento.

4.1. Conjuntos de dados e configurações experimentais

Aquisição de dados Gerar conjuntos de dados informativos a partir de fontes multidefismas massivas é um desafio em nosso experimento. Para ter certeza, coletamos eventos financeiros, notícias, preços de mercado e o gráfico de conhecimento para todas as 3714 empresas públicas listadas no mercado de ações A da China,³ a partir de 01 de janeiro 2018 a 31 de dezembro de 2019. Em particu-lar, nós rastreamos os anúncios públicos e aproveitamos os métodos de extração de eventos para construir eventos para empresas de lista. São 143.884 eventos estruturados em 41 categorias em nosso conjunto de dados, como relatórios sazonais/anuais, reestruturação de ativos, aumento/diminuição das classificações de crédito, mudança do presidente ou membros do conselho, acidente de produção, etc. Para notícias financeiras, rastreamos informações de 87 grandes sites que cobrem os relatórios mais importantes do mar-ket. Há 5,13 milhões de notícias durante o intervalo de tempo. Nós alavancamos o reconhecimento de entidades (e vinculando) e as tecnologias de ex-tração de relação neural para extrair entidades e relações a partir de textos brutos. Essas empresas listadas vinculadas são armazenadas como nós no gráfico de

conhecimento, no qual cada relação é armazenada como um nó de ponta. Por fim, são 5,26 milhões de entidades e 6,93 milhões de relações na FinKG.

Reunimos os dados de preços das ações da China listadas compa-nies das fontes da Bolsa de Valores de Xangai e Shenzhen de 2018 a 2020, incluindo 500 dias de negociação. Os dados diários do mercado em-clude os preços das ações (aberto, fechado, alto e baixo) e a negociação em formação (volume de negociação e taxa de rotatividade) daquele dia para cada estoque. No experimento, removemos o estoque de suspensão de negociação e os preços intradáveis (como estoques limitados e limitados) do conjunto de dados. No mercado de ações da China, os investidores precisam seguir um mecanismo de limite de 10% para baixo estritamente.

Configurações experimentais Prevemos o movimento do preço em três categorias { *para cima*, *neural*, *para baixo* }. Para o estoque i no dia t , a taxa de retorno pode ser calculada por $R_{r,i} = p_{t,i} / p_{t-1,i} - 1$. Nós estabelecemos o rótulo de verdade dos movimentos de preços como:

$$\begin{aligned} \text{até} \quad R_r &\geq r_{up}, \\ f(R_r) &= \text{neural } r_{para\ baixo} < R_r < r_{para\ cima}, \quad (7) \text{ para baixo } R_r \leq \\ r_{para\ baixo} \end{aligned}$$

onde configuramos $r_{up} = 0.01$ e $r_{para\ baixo} = -0.01$. Em nosso conjunto de dados, há 226.585 amostras na categoria "up", 327.851 "neurais" e 238.630 na categoria "down".

No experimento, utilizamos os dados do ano de 2018 como conjunto de treinamentos e avaliamos o desempenho no ano de 2019. Partic-ularly, construímos características a partir de dados multi-modalidade nos 60 dias de negociação re-cent e aplicamos o movimento de preços do dia seguinte como o rótulo. Em seguida, aplicamos uma janela deslizante de cada dia de negociação e reportamos o resultado médio de 2019 no experimento. Nas configurações de estratégia de trad-ing, simplesmente compramos o previsto "up", vendemos as ações "para baixo" e não mantemos nenhuma ação sobre ações "neurais". O percentual de trad-ing é alocado pelos pesos lineares da capacidade de prob prevista. Por favor, note que existem muitas técnicas para o desenvolvimento de uma estratégia de negociação, que está além do escopo deste artigo. Nós ig-nore os custos de transação de todos os métodos comparados para simplicidade e equidade no experimento.

Métodos comparados e métricas de avaliação Utilizamos as abordagens fol-lowing amplamente utilizadas como linhas de base para validar a effectiveness do nosso método proposto: Stock-LSTM [13], News-ATT [9],

Tabela 1

A comparação do desempenho de previsão.

	Micro-F1	Macro-F1	F1 ponderado
Stock-LSTM	0.4540	0.4233	0.4489
News-ATT	0.4551	0.4551	0.4502
Stock-GAT	0.4656	0.4396	0.4654
Event-NTN	0.4720	0.4478	0.4718
MAGNN-G	0.4815	0.4607	0.4798
MAGNN-S	0.4813	0.4604	0.4793
MAGNN-tudo	0,4838 **	0,4627 **	0,4825 **

** indica que as melhorias são estatisticamente significastas para $p < 0.01$ julgado por t-test emparelhado.

Stock-GAT [23], e Event-NTN [21]. Todos os parâmetros são definidos com base em suas sugestões padrão no papel. Por exemplo, Stock-LSTM é definido como duas camadas com um tamanho oculto de 100 e 50. Nosso método possui duas variações: MAGNN-G e MAGNN-S, que empregam apenas a atenção do

³ <https://www.investopedia.com/terms/a/a-shares.asp>

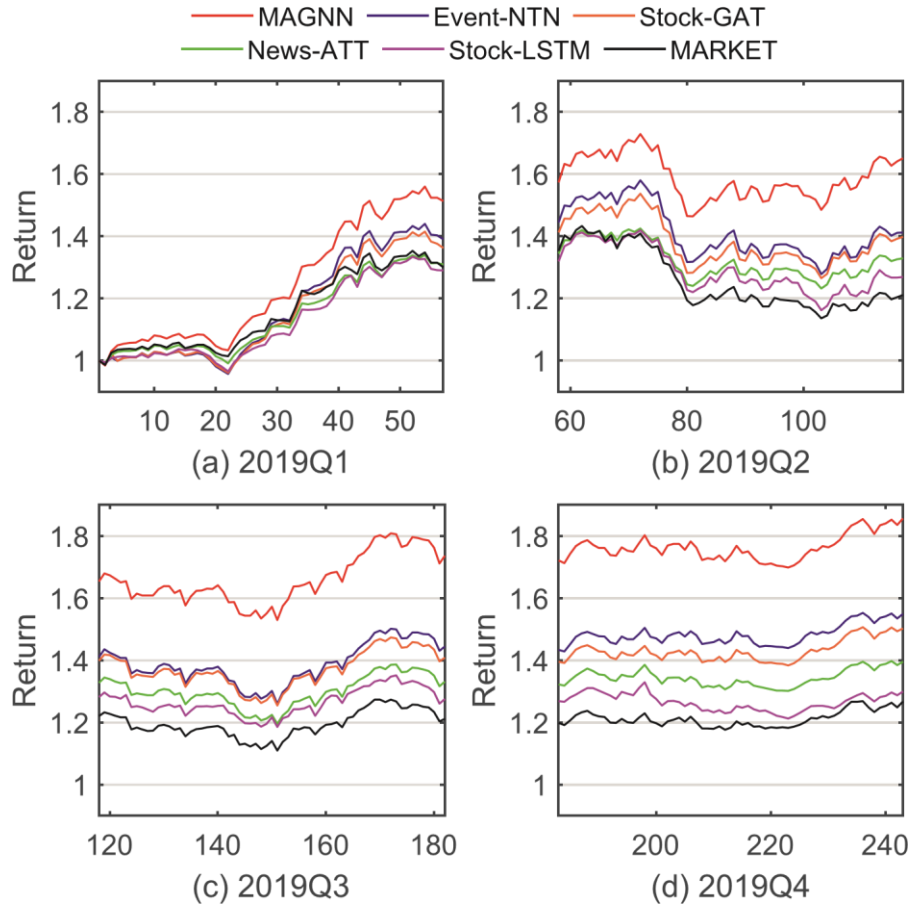


Fig. 3. Os retornos acumulados obtidos no conjunto de testes (2019) pelo método proposto e compararam as linhas de base. Para melhor ilustração, dividimos em quatro quartos de vista. gráfico da modalidade interna ou a fonte inter-modalidade apenas em tenção. MAGNN-all denota a versão completa de nossas técnicas propostas.

Para a avaliação, aplicamos a pontuação Micro-F1, Macro-F1 e Weighted-F1 para medir o desempenho da previsão de cura. Para a carteira construída, empregamos retorno acumulado de ativos (Retorno), retorno diário médio (Retorno D) e Índice Sharpe amplamente utilizado [24] como métricas de avaliação. Um Retorno é formulado como

$$art = \frac{1}{|S|} \sum_{i \in S} \frac{p_{t-1} - p_{t-2}}{p_{t-1} - p_{t-2}}, \quad (8)$$

onde S_{t-1} denota o conjunto de ações em carteira no momento $t-1$. p_{t-1} é o preço para ações i no momento t e $|\cdot|$ denota o número de itens definidos. A taxa sharpe (RS) é o retorno médio obtido acima da taxa livre de risco por unidade de volatilidade, que é expressa como:

$SR = (R_p - R_f) / \sigma_p$ onde R_p é o retorno da carteira, R_f é a taxa livre de risco, σ_p é o desvio padrão do retorno excessivo da carteira. Usamos o rendimento de títulos do governo chinês de 1 ano⁴ como taxa livre de risco.

4.2. Previsão financeira

Nesta seção, avaliamos a precisão de previsão das séries de tempo financeiro, que é a principal tarefa deste artigo. A Tabela 1 relata a pontuação micro-F1, Macro-F1 e Weighted-F1 de cada abordagem. ** denota que as melhorias são estatisticamente significativas para $p < 0.01$ julgado por t-test emparelhado.

As quatro primeiras linhas da Tabela 1 mostram o resultado da classificação das linhas de base comparadas. É claro que, Stock-LSTM e News-ATT não são satisfatórios, demonstrando que nem ações nem notícias por si só poderiam alcançar o desempenho ideal. Stock-GAT é ligeiramente melhor que o Stock-LSTM, provando a eficácia da preservação da estrutura gráfica em um modelo de previsão de séries temporais. Em toda a linha de base, Event-NTN é mais competitivo, o que supera consideravelmente o Novo-ATT. O processo de extração de eventos estruturados a partir de notícias brutas mostra-se útil na aprendizagem de incorporações representativas. As linhas 5 e 6 exibem o desempenho das variações do nosso método proposto. Como podemos ver, MAGNN-S é semelhante ao MAGNN-G. Ambos são apostas melhores do que as linhas de base mais competitivas. A validade da integração de insumos multi-modalidade em nossa tarefa é fortemente comprovada. É um momento para projetar um modelo inovador para aprender com as fontes acima, que é a principal motivação deste artigo. O MAGNN-all outperforms todas as linhas de base, demonstrando sua superioridade em aprender com insumos multi-modalidade para previsão financeira.

Tabela 2

O retorno de carteiras com método diferente.

Métodos	Um Retorno	D Retorno	Razão sharpe
---------	------------	-----------	--------------

⁴ <http://yield.chinabond.com.cn/>

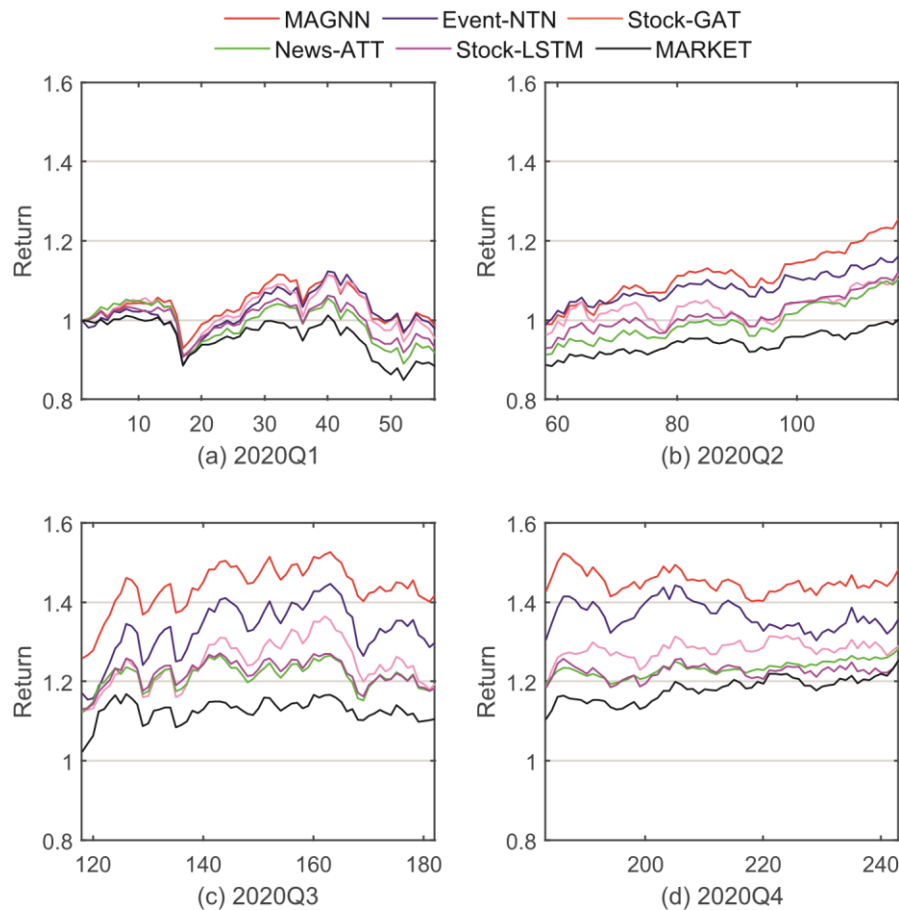


Fig. 4. Os retornos acumulados obtidos na janela de validação (2020) pelo método proposto e comparados às linhas de base.

Stock-LSTM	0.3002	0.0012	2.7919
News-ATT	0.3960	0.0015	3.5097
Stock-GAT	0.5035	0.0018	3.4506
Event-NTN	0.5507	0.0019	3.4467
MAGNN-G	0.6521	0.0021	3.5082
MAGNN-S	0.6775	0.0021	3.5561
MAGNN-tudo	0.8571 **	0.0027 **	3.7619 **

Neste experimento, gostaríamos de salientar que a previsão de tendência de mercado é muito desafiadora e uma pequena fração de improvement já pode trazer uma grande quantidade de receita na indústria financeira. De acordo com a prática de Marcos et al., [25] mesmo 0,005 melhorias na precisão da previsão é muito difícil para novos pesquisadores, o que normalmente poderia levar a mais de 12% de lucros excedentes. Nosso método melhora as melhores linhas de base acima de 1% na Tabela 1 e, consequentemente, leva a melhorias de lucro de quase 30% nos retornos acumulados, conforme relatado na Tabela 2. Portanto, podemos afirmar com segurança que nossos métodos propostos superam significativamente as linhas de base de última geração na tarefa de previsão.

4.3. Desempenho da carteira

Na avaliação de desempenho da carteira construída, relatamos o retorno do ativo (Retorno), retorno diário médio (Retorno D) e Índice Sharpe primeiro. Em seguida, apresentamos a curva de retorno acumulada através do intervalo de tempo do período de teste. Como descrito acima, compramos as ações previstas "para cima" e vendemos as "para baixo". A posição é simplesmente definida linear à probabilidade de prever saídas de modelos.

A Tabela 2 relata o desempenho das carteiras de investimentos

construídas pelo nosso método proposto e outras linhas de base. Podemos servir que, em todas as métricas de avaliação, nossa técnica proposta for- executa a linha de base significativamente. Particularmente, a Stock-LSTM e a News-ATT alcançam menor desempenho em "A Return" e "Sharpe Ratio", o que indica os baixos retornos rentáveis. Ao classificar o gráfico de conhecimento e eventos estruturados, o retorno do Stock-GAT e event-NTN é maior do que as linhas de base clássicas. O mesmo fenômeno é observado nas métricas de proporção sharpe. As últimas três linhas exibem o resultado do nosso método proposto e seus sub modelos. O MAGNN tem um desempenho melhor do que todos os métodos comparados em três métricas de avaliação amplamente utilizadas. A eficácia do nosso método proposto na construção de carteiras rentáveis é fortemente comprovada.

Para avaliar ainda mais o retorno do nosso método proposto através do intervalo de tempo de teste, examinamos o retorno acumulado em cada dia de negociação e reportamos os resultados comparados em Fig. 3. Observamos que a Stock-LSTM está muito próxima do índice CSI 300 do mercado⁵ ao longo do ano de 2019. News-ATT são melhor do que Stock-LSTM. Desde o final de 2019Q1, nosso método lidera os retornos e amplia a lacuna em 2019T2 com linhas de base comparadas, o que é notável. Em seguida, realizamos estudos empíricos com os principais especialistas financeiros sobre a vitrine das curvas de retorno. A razão parece ser que quando o mercado cai (em 2019T2), nosso método poderia prever o sinal "para baixo" com antecedência, que é aprendido com o implicitamente efeitos de lead-lag em fontes multi-modalidade. Como um re-

⁵ <http://www.csindex.com.cn/en/indices/index-detail/0 0 030 0>

sult, magnn realiza o melhor desempenho desde o final de 2019Q1 e lidera a superioridade até o final de 2019Q4, com mais de 80% dos lucros do investimento.

4.4. Generalização do desempenho

Para observar a generalização de desempenho do nosso método pró-posed, optamos por uma duração maior para avaliar o retorno do porto-folio em comparação com as linhas de base. Em particular, treinamos nosso modelo pelos dados históricos durante 2019 e depois prevemos os movimentos dos preços em 2020. A estratégia de negociação é definida como a experiência anterior e a porcentagem de negociação também é alocada pelos pesos lineares da probabilidade prevista. [Figo. 4](#) relatar o desempenho dos retornos acumulados pelo método proposto e comparar as linhas de base. Como podemos ver, o Stock-LSTM é o mais baixo, que está muito próximo do retorno do mercado no final de 2020. Enquanto o News-ATT e Stock-GAT têm melhor desempenho que a Stock-LSTM e o Market, provando a eficácia de incluir notícias e re-lações de gráficos no preço das ações tarefa de previsão. Em todas as linhas de base, Event-NTN é mais competitivo. Nosso método alcança o melhor resultado na janela ob-servation e liderando constantemente o desempenho desde o início do segundo trimestre de 2020. O resultado do experimento de erradicação de geração de modelos por meio de dados de treinamento em 2019 é consistente com o resultado do treinamento do modelo com 2018, o que demonstra a eficácia e generalização do nosso método proposto.

4.5. Interpretação do modelo de atenção

Como descrito na seção preliminar, os movimentos de preços das ações podem ser parecidos como a reação de eventos financeiros ou notícias, que também está relacionada ao seu próprio desempenho histórico. Assim, precisamos incluir fontes multi-modalidade (notícias, eventos, mercado) como a entrada do nosso modelo. No entanto, diferentes fontes contribuem com differently. O mecanismo de atenção intermoda pode automaticamente aprender seus pesos na previsão de preços e assim alcançar o desempenho de última geração. Além disso, dentro de cada modalidade, a relação interna de diferentes empresas também é muito importante. Por exemplo, um evento "Microsoft compra LinkedIn" foi relatado em 13 de junho de 2016. Imediatamente, o preço do LinkedIn subiu 46,81% e o Mi-crosoft caiu 3,2% naquele dia. Curiosamente, os preços das vendas- força, que é o principal concorrente do LinkedIn, caíram mais de 6% nas duas semanas seguintes. Salesforce não é o tema direto deste evento, mas também é profundamente influenciado. O modelo de atten-tion de modalidade interna poderia aprender essa relação estruturada em gráficos entre as fontes de entrada e a previsão de destino. Portanto, nosso método poderia ajudar a prever o movimento dos preços das ações com mais precisão e os resultados de ex-periment demonstram fortemente seu desempenho superior.

Então, para explorar a interpretação do nosso método proposto, visualizamos os pesos de atenção tanto da atenção do gráfico da modalidade interna quanto da atenção de origem interde modalidade na [Fig. 5](#). Localizamos cada equidade de acordo com o retorno previsto (x-eixo) e sua situação no gráfico heterogêneo ou fonte de modalidade (y - eixo) no calor mapa. Em seguida, coloremos pelos pesos médios de atenção no modelo de previsão.

[Fig. 5](#) a exhibe os pesos aprendidos do InnGAT. Como podemos ver, ações com aproximadamente três a seis nós de bairro gen- erally contribuem mais importante no modelo. Além disso, não há diferença notável para estruturas de nó em diferentes mercados (por exemplo, através do eixo x). O resultado prova que o retorno de um nó com cerca de quatro vizinhos no gráfico heterogêneo é

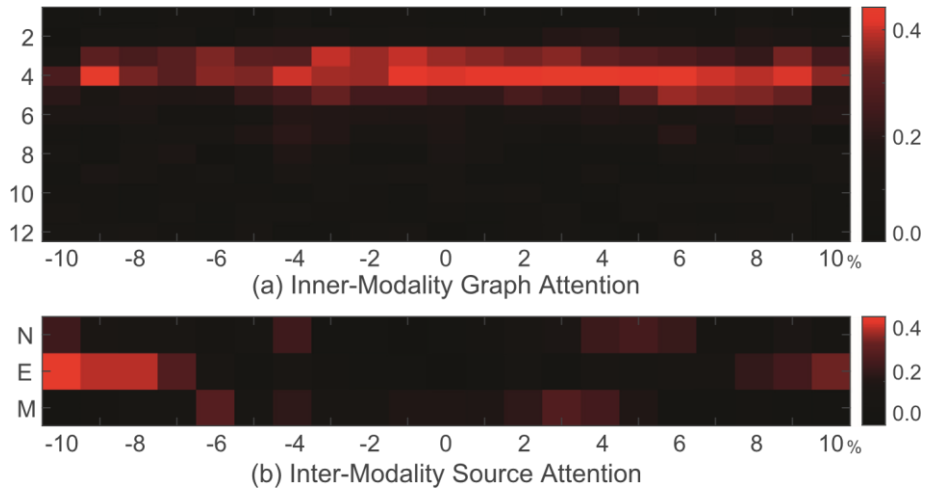


Fig. 5. Visualização dos coeficientes atencionais. X-eixo denota o retorno diário das ações. O eixo Y denota o número de vizinhos no subgrafo (a) e o tipo de modalidade em (b), no qual N,E,M denota as notícias, eventos e preços de mercado.



Fig. 6. A interface do nosso MAGNN proposto no sistema de gerenciamento de portfólio baseado na Web, que é implantado em um grande provedor de serviços financeiros da China. Nós traduzimos o informações-chave por laranja escrito à mão e palavras sublinhadas.

mais propensos a serem influenciados por nós adjacentes, o que significa que os efeitos de lead-lag são mais proeminentes nesta situação.

Uma saída mais interpretável é observada em Fig. 5 b. Ao visualizar os pesos de atenção de cada modalidade, constatamos que todas as fontes (N, E e M) contribuem de forma importante para o modelo de previsão, o que comprova fortemente a declaração deste papel. Além disso, as notícias são semelhantes em diferentes situações de mercado, com um pequeno retorno positivo ligeiramente proeminente. O mesmo fenômeno é observado na modalidade de mercado. Pelo contrário, a contribuição dos eventos é significativamente maior em grandes retornos positivos e negativos do que os pequenos. A razão pode ser que a mudança acentuada da equitidade é principalmente impulsionada por eventos, em vez de notícias rotineiras ou dados de preços. O resultado demonstra os efeitos da entrada multidefismam e do modelo de atenção proposto.

4.6. Implementação e implantação do sistema

Em seguida, implantamos nosso método em cenários reais e avaliamos seu desempenho no experimento de rastreamento de mercado. Fig. 6 mostra a

interface do sistema de gestão de portfólio do nosso método proposto. Na visualização de navegação na página inicial (Fig. 6 a), podemos investigar as ações previstas "para cima" e "para baixo", o portfólio construído e os principais eventos. O navegador de esquerda pro-vides os efeitos de cada fonte de modalidade, incluindo notícias, eventos e preços de mercado, e o status de lead-lag das ações em eventos selecionados. Figo. 6 b relata a visão de previsão sobre um patrimônio típico da Spring Airlines (Código: 601021) Ltd., que é a primeira e maior companhia aérea de baixa tarifa da China. A parte superior mostra o preço das ações de verdade, e a parte inferior exhibe a relação de movimentos de preços previstos desde 01 de janeiro de 2020.

Como podemos ver, nosso método previu com sucesso quatro flutuações significantes antes da Spring Airlines. Por fim, informamos os detalhes de desempenho da carteira na Fig. 6 c. O resultado mostra que nosso método melhora significativamente os retornos com mais de 60% do lucro excess. Além disso, ao aprender os efeitos de lead-lag de fontes multimodais, nosso método poderia evitar grandes perdas no mercado, o que diminui o saque máximo de $-16,08\%$ a $-12,48\%$.

Na implementação, empregamos o Scrapy distribuído como o rastreador web, Redis como o banco de dados na memória. O modelo proposto é escrito em Tensorflow em Python e requer duas horas para treinamento em duas peças

da GPU Telsa P100. O sistema integrado de gerenciamento de portas-fólio é implementado pelos micro-serviços Spring Cloud e escrito em Java.

5. Obras relacionadas

Nesta seção, introduzimos alguns trabalhos relacionados à nossa pesquisa, incluindo análise de séries de tempo financeiro e rede neural de gráficos multi-modalidade.

Análise de séries de tempo financeiro Nas últimas décadas, inúmeras obras têm proposto prever a série de tempo financeiro [26]. As abordagens tradicionais incluem modelo autoregressivo, método médio móvel, análise de fatores [16], otimização de indicadores [27], etc. Af-terwards, as técnicas de aprendizado de máquina têm sido empregadas para a previsão do preço das ações, como máquina vetorial de suporte [28], impulsionamento de árvores [29] e rede neural, especialmente para o RNN e LSTM [30]. Recentemente, alguns pesquisadores demonstraram a efetividade de aproveitar notícias e eventos de texto não estruturados para aprender incorporações representativas para a previsão do preço das ações [31,32]. Advanced transfer learning [33] e o aprendizado não supervisionado [34] niques tecnológicos também são introduzidos para aprender as incorporações significativas para análise de séries temporais. No entanto, as abordagens existentes não aprendem as relações internas dos movimentos das ações sobre os efeitos de lead-lag de eventos (ou notícias). A maioria deles só emprega uma única fonte de modalidade para previsão, o que pode descartar muitas informações úteis.

Multi-modalidade e rede de atenção gráfica Graph neural net-work (GNN) mostrou seu desempenho superior para representar dados estruturados em gráficos [35,36]. O modelo de atenção gráfica melhora a representação do nó ajustando pesos significativos no processo de agregação com nós adjacentes [37], indicando a importância de os nós correspondentes [38] e as relações atribuídas [39]. A GNN com mecanismo de atenção tem mostrado sua eficácia em uma ampla gama de campos, incluindo finanças [40], saúde [41], computer vision [42], e-commerce [43,44] etc. Recentemente, alguns trabalhos exploram a aplicação da GNN ao aprendizado de insumos multi-modalidade, como o diagnóstico de doenças por imagens multiameus [45]. No entanto, há poucos estudos sobre previsão financeira por rede neural de gráficos multi-modalidade.

6. Conclusão

Neste artigo, propomos uma nova rede de grafia de várias modalidades para previsão de séries de tempo financeiro. Nosso método ad-veste o principal problema da previsão de preços no financeiro em dustry, interpretando os efeitos de lead-lag com fonte inform-tive, pela modalidade inner graf attention e mecanismo de atenção de origem intersecorra. Avaliamos minuciosamente a eficácia do método proposto, comparando-o com as linhas de base de última geração nos conjuntos de dados históricos maciços. Além disso, desesploramos o modelo em aplicações do mundo real, e o resultado prova que nosso trabalho poderia evitar perdas significativas de investimento financeiro.

Em conclusão, este é o primeiro trabalho a estudar o problema de previsão de séries de tempo financeiro por técnicas avançadas de GNN com fontes em formação, que podem inovar mais estudos sobre tanto a ciência da computação quanto as comunidades financeiras. Por um lado, estendemos o modelo de atenção gráfica a cenários multi-modalidade; por outro lado, melhoramos a previsão financeira com o aprendizado sobre dados alternativos.

Declaração de Interesse Concorrente

Os autores declaram que não têm interesses concorrentes ou relações pessoais que possam ter aparecido para influenciar o trabalho relatado neste artigo.

Confirmações

O trabalho é apoiado pelo Programa Nacional de P&D da China (2018YFB2100801), a Fundação Nacional de Ciência da China (6210070838), o Shanghai "Inovacientífica e Tecnológica- Plano de Ação" Projetos de Alta e Nova Tecnologia (19511101300). **Referências**

- [1] Worldbank, Capitalização de mercado de empresas nacionais listadas, 2020, Ac-cessed 16-março-2021, <https://data.worldbank.org/indicator/CM.MKT.LCAP.CD>.
- [2] E.F. Fama, O comportamento dos preços da bolsa de valores, J. Autocarro. 38.1 (1965) 34-105 APA.
- [3] M.L. O'Connor, A relação transversal entre os custos de negociação e efeitos de lead/lag nos mercados de ações & opções, Financ. Rev. 34 (4) (1999) 95-117.
- [4] A. Cowles 3^o, Os analistas do mercado de ações podem prever? Econometrica 1 (3) (1933) 309-324.
- [5] S.J. Taylor, Modelagem Financeira Time Series[M], World Scientific, 2008.
- [6] C.Q. Cao, R. S. Tsay, Análise não linear da série temporal das volatilidades das ações[J], J. Appl. O Econom. 7 (S1) (1992) S165-S185.
- [7] R.D. Edwards, J. Magee Bassetti, Análise Técnica das Tendências de Ações, 2012.
- [8] D. Cheng, Y. Tu, Z. Niu, L. Zhang, Aprendendo relações temporais entre sinais financeiros, em: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2018, pp. 2641-2645.
- [9] Z. Hu, W. Liu, J. Bian, X. Liu, T.-Y. Liu, Ouvindo sussurros caóticos: um quadro de aprendizado profundo para a previsão de tendência de ações orientada a notícias, em: Proceedings of the Eleventh ACM International Conference on Web Search & Mineração de Dados, WSDM '18, Association for Computing Machinery, Nova York, NY, EUA, 2018, pp. 261-269.
- [10] D. Tashiro, H. Matsushima, K. Izumi, H. Sakaji, Codificação de informações de alta frequência ou der e previsão de preços de ações de curto prazo por deep learning, Quant. Finanças 19 (9) (2019) 1499-1506.
- [11] S. Bharathi, A. Geetha, Análise de sentimento para previsão efetiva do mercado de ações, Int. J. Intell. Eng. Syst. 10 (3) (2017) 146-154.
- [12] Y. Xu, S.B. Cohen, Previsão do movimento de ações a partir de tweets e preços históricos, em: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 1970-1979.
- [13] T. Fischer, C. Krauss, Aprendizado profundo com redes de memória de longo prazo para previsões do mercado financeiro, Eur. J. Res. 270 (2) (2018) 654-669.
- [14] S. Deng, N. Zhang, W. Zhang, J. Chen, J.Z. Pan, H. Chen, Previsão e explicação de tendência de estoque orientada ao conhecimento via rede convolucional temporal, em: Processo Complementar da 2019 World Wide Web Conference, 2019, pp. 678-685.
- [15] D. Cheng, F. Yang, X. Wang, Y. Zhang, L. Zhang, Estrutura de incorporação de eventos baseados em gráficos de conhecimento para aplicações quantitativas financeiras, em: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Recuperação de Informações, 2020, pp. 2221-2230.
- [16] G. Ganesapillai, J. Gutttag, A. Lo, Conexões de aprendizagem em séries de tempo financeiro, em: Conferência Internacional sobre Machine Learning, 2013, pp. 109-117.
- [17] W.S. Chan, Reação do preço das ações às notícias e sem notícias: deriva e reversão após manchetes, J. O Financ. Economia. 70 (2) (2003) 223-260.
- [18] R. Hussein, D. Yang, P. Cudré-Mauroux, Os meta-caminhos são necessários? Revisitar-ing incorporações de gráficos heterogêneos, em: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 437-446.
- [19] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: pré-treinamento de transformadores bidirec profundos para compreensão da linguagem, 2018 arXiv:1810.04805
- [20] S. Saha, et al., Extração de informações abertas de sentenças conjuntivas, em: Proceedings of the 27th International Conference on Computational Linguistics, 2018, pp. 2288-2299.
- [21] X. Ding, K. Liao, T. Liu, Z. Li, J. Duan, Avaliação de eventos de aprendizagem aprimorada com conhecimento externo do senso comum, em: Proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9^o Conferência Conjunta In-ternacional sobre Processamento de Linguagem Natural (EMNLP-IJCNLP), 2019, pp. 4 896-4 905.
- [22] D.P. Kingma, J. Ba, Adam: um método para otimização estocástica, arXiv pré-impressão arXiv:1412.6980 (2014).
- [23] P. Velickovic, A. Casanova, P. Lio, G. Cucurull, A. Romero, Y. Bengio, Graph attention networks, in: 6^o Conferência Internacional sobre Representações de Aprendizagem, ICLR 2018 - Conference Track Proceedings, 2018.
- [24] O. Ledoit, M. Wolf, Teste de hipótese de desempenho robusto com o sharpe ratio, J. A Empir. Finanças 15 (5) (2008) 850-859.
- [25] M.L. De Prado, Avanços em Machine Learning Financeiro, John Wiley & Sons, 2018.
- [26] T.G. Andersen, R.A. Davis, J.-P. Kreiß, T.V. Mikosch, Manual da Série Tempo Financeiro, Springer Science & Business Media, 2009.
- [27] Z. Li, D. Yang, L. Zhao, J. Bian, T. Qin, T.-Y. Liu, Indicador individualizado para todos: otimização de indicadores técnicos em termos de estoque com incorporação de

- ações, em: *Proceedings da 25ª Conferência Internacional ACM SIGKDD sobre Conhecimento Descoberto & Data Mining*, 2019, pp. 894-902.
- [28] L.-J. Cao, F.E.H. Tay, Suporte a máquina vetorial com parâmetros adaptativos na previsão da série temporal financeira, *IEEE Trans. Netw. neural.* 14 (6) (2003) 1506-1518.
- [29] A. Agapitos, A. Brabazon, M. O'Neill, Regularização do gradiente para modelagem de séries de tempo financeiro, *Computação. O Manag. Sci.* 14 (3) (2017) 367-391.
- [30] A. Sagheer, M. Kotb, Pré-treinamento não supervisionado de um autoencoder empilhado profundo baseado em LSTM para problemas de previsão de séries temporais multivariadas, *Sci. Rep.* 9 (1) (2019) 1-16.
- [31] D. Zhou, L. Zheng, Y. Zhu, J. Li, J. Ele, Rede de atenuação neural multiportativa adaptativa de domínio para previsão financeira, em: *Proceedings of The Web Conference 2020*, 2020, pp. 2230-2240.
- [32] O.B. Sezer, M.U. Gudelek, A.M. Ozbayoglu, Série de tempo financeiro com aprendizado profundo: uma revisão sistemática da literatura: 2005-2019, *Appl. Soft Comput.* 90 (2020) 106181.
- [33] R. Ye, Q. Dai, Implementando o aprendizado de transferência em diferentes conjuntos de dados para previsão de séries temporais, *Pattern Recognit.* 109 (2021) 107617.
- [34] H. Wang, Q. Zhang, J. Wu, S. Pan, Y. Chen, Série time apresenta aprendizado com dados rotulados e não rotulados, *Pattern Recognit.* 89 (2019) 55-66.
- [35] N. Dehmamy, A.-L. Barabási Yu, Entendendo o poder de representação das redes neurais graph na topologia de gráficos de aprendizagem, em: *Avanços em Sistemas de Processamento de Informações Neurais*, 2019, pp. 15413-15423.
- [36] F. Manessi, A. Rozza, M. Manzo, Redes de gráficos dinâmicos, *Pattern Recognit.* 97 (2020) 10700.
- [37] B. Knyazev, G.W. Taylor, M. Amer, Compreensão da atenção e generalização em redes neurais gráficas, em: *Avanços no Processamento de Informações Neurais Systems*, 2019, pp. 4202-4212.
- [38] N. Park, A. Kan, X.L. Dong, T. Zhao, C. Faloutsos, Estimando a importância do nó em gráficos de conhecimento usando redes neurais de gráficos, em: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 596-606.
- [39] Feng, C. Xu, Y. Zuo, G. Chen, F. Lin, J. Xiahou, rede de atenção de gráficos tributosa para recomendação de ações[J], *Pattern Recognit.* (2021) 108-119, doi: 10.1016/j.patcog.2021.108119.
- [40] D. Cheng, Z. Niu, Y. Zhang, Classificação de risco de cadeia contagiosa para empréstimos em rede-guarantee, em: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, pp. 2715-2723.
- [41] E. Choi, M.T. Bahadori, L. Song, W.F. Stewart, J. Sun, Gram: modelo de atenção baseado em gráficos para aprendizagem de representação em saúde, em: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 787-795.
- [42] J. Wu, S.-h. Zhong, Y. Liu, Rede convolucional de gráfico dinâmico para resumo de vídeo multi-, *Padrão Recognit.* 107 (2020) 107382.
- [43] D. Cheng, S. Xiang, C. Shang, Y. Zhang, F. Yang, L. Zhang, Spatio-temporal attention-based rede neural para detecção de fraude de cartão de crédito, em: *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 2020, pp. 362-369.
- [44] D. Cheng, X. Wang, Y. Zhang, et al., rede neural gráfica para detecção de fraudes via atenção espaço-temporal[J], *IEEE Trans. O Knowl. Data Eng.* (2020) <https://ieeexplore.ieee.org/abstract/document/9204584/>.
- [45] X. Fang, Z. Liu, M. Xu, Conjunto de redes neurais convolucionais profundas baseadas em imagens multi-modalidade para o diagnóstico da doença de Alzheimer, *IET Image Proc.* 14 (2) (2019) 318-326.

Dawei Cheng é professor assistente do departamento de ciência da computação e tecnologia da Universidade de Tongji, Xangai, China. Antes disso, Dawei foi um associado de pós-doutorado no laboratório de inteligência artificial do MoE, departamento de ciência da computação e engenharia da Universidade De Xangai Jiao Tong. Ele recebeu o Ph.D. Degree em ciência da computação da Universidade de Xangai Jiao Tong, Xangai, China, em 2018. Seus interesses de pesquisa incluem aprendizado gráfico, mineração de dados e aprendizado de máquinas.

Fangzhou Yang é um cientista sênior de pesquisa do laboratório de inteligência artificial (equipe Seek-Data), Emoney Inc., Xangai, China. Fangzhou recebeu seu diploma de mestrado pela Universidade Técnica de Berlim e pela Universidade de Xangai Jiao Tong. Antes disso, ele recebeu sua B.Sc. graduado em ciência da computação pela Universidade de Xangai Jiao Tong. Seus interesses de pesquisa incluem reconhecimento de padrões e mineração de dados em finanças, gestão de portfólio e investimento quantitativo.

Sheng Xiang é ph.D. candidato no Centro de Inteligência Artificial, graduado em Ciência da Computação, Universidade de Tecnologia de Sydney, Austrália. Ele recebeu sua B.Sc. Graduado em Engenharia de Bioinformática pela Universidade de Xangai Jiao Tong. Seus

interesses de re-pesquisa incluem processamento e análise de gráficos grandes, processamento de consultas em fluxo de dados, dados incertos e gráficos.

Jin Liu é um cientista sênior de pesquisa do laboratório de inteligência artificial (Equipe Seek-Data), Emoney Inc., Xangai, China. Recebeu seu bacharelado e mestrado em ciência da computação e engenharia pela Universidade de Wuhan, Hubei, China. Seus interesses de pesquisa incluem mineração de dados financeiros, análise de fatores e machine learning.