

Listas de conteúdo disponíveis em [ScienceDirect](#)

Diário de Computação Suave Aplicada

página inicial da revista: www.elsevier.com/locate/asoc

Previsão de séries temporais financeiras com aprendizado profundo : Uma revisão sistemática da literatura: 2005-2019

Omer Berat Sezer, Mehmet Ugur Gudelek, Ahmet Murat Ozbayoglu[∗]

Departamento de Engenharia da Computação, TOBB University of Economics and Technology, Ankara, Turquia

informações do artigo

Historia do artigo:

Recebido em 11 de maio de 2019

Recebido em formulário revisado em 9 de fevereiro de 2020

Aceito em 12 de fevereiro de 2020

Disponível on-line em 18 de fevereiro de 2020

Palavras-chave:

Aprendizado profundo

Finança

Inteligência computacional

Aprendizado de máquina

Previsão de séries temporais

CNN

LSTM

RN

resumo

A previsão de séries temporais financeiras é, sem dúvida, a principal escolha de inteligência computacional para pesquisadores de finanças, tanto na academia quanto no setor financeiro, devido às suas amplas áreas de implementação e impacto substancial. Os pesquisadores de Machine Learning (ML) criaram vários modelos e um grande número de estudos foi publicado em conformidade. Como tal, existe um número significativo de pesquisas cobrindo estudos de ML sobre previsão de séries temporais financeiras. Ultimamente, modelos de Deep Learning (DL) apareceram no campo, com resultados que superaram significativamente seus equivalentes tradicionais de ML. Embora haja um interesse crescente no desenvolvimento de modelos para previsão de séries temporais financeiras, há uma falta de artigos de revisão que se concentrem apenas em DL para finanças. Assim, a motivação deste artigo é fornecer uma revisão abrangente da literatura de estudos de EAD sobre a implementação de previsão de séries temporais financeiras. Nós não apenas categorizamos os estudos de acordo com suas áreas de implementação de previsão pretendidas, como índice, forex e previsão de commodities, mas também os agrupamos com base em suas escolhas de modelo DL, como Redes Neurais Convolucionais (CNNs), Redes de Crenças Profundas (DBNs) e Memória de Longo Prazo (LSTM). Também tentamos vislumbrar o futuro do campo destacando seus possíveis retrocessos e oportunidades em benefício dos pesquisadores interessados.

© 2020 Elsevier BV Todos os direitos reservados.

1. Introdução

O setor financeiro sempre se interessou pela previsão bem-sucedida de dados de séries temporais financeiras. Numerosos estudos foram publicados sobre modelos de ML com desempenhos relativamente melhores do que as técnicas clássicas de previsão de séries temporais.

Enquanto isso, a ampla aplicação de sistemas automatizados de negociação eletrônica, juntamente com a crescente demanda por rendimentos mais altos, continua forçando pesquisadores e profissionais a continuar trabalhando na implementação de melhores modelos. Assim, novas publicações e implementações continuam agregando à literatura de finanças e inteligência computacional.

Nos últimos anos, DL emergiu fortemente como a classe de previsão de melhor desempenho no campo de ML em várias áreas de implementação. A previsão de séries temporais financeiras não é exceção e, como tal, um número crescente de modelos de previsão baseados em várias técnicas de DL foi introduzido nas conferências e periódicos apropriados nos últimos anos. Apesar do grande número de artigos de pesquisa cobrindo sistemas de previsão e negociação de séries temporais financeiras usando técnicas tradicionais de computação soft, até onde sabemos, nenhuma revisão foi realizada na literatura para DL. Por isso, decidimos trabalhar em tal

um estudo abrangente, com foco em implementações DL de previsão de séries temporais financeiras. Nossa motivação é dupla; não apenas visamos fornecer um instantâneo do estado da arte das perspectivas acadêmicas e da indústria de modelos de EAD desenvolvidos, mas também identificar as características importantes e distintivas de cada modelo estudado para evitar que pesquisadores e praticantes façam escolhas insatisfatórias durante seus estudos. desenvolvimento de sistema.

Também queríamos prever para onde a indústria está indo, indicando possíveis direções futuras.

Nossa motivação fundamental foi responder às seguintes questões de pesquisa:

- Quais modelos de DL são usados para previsão de séries temporais financeiras no?
- Como o desempenho dos modelos DL se compara ao de seus equivalentes tradicionais de ML? • Qual é a direção futura da pesquisa DL para previsão de séries temporais financeiras?

Nosso foco foi exclusivamente em implementações de DL para previsão de séries temporais financeiras. Para outras aplicações financeiras baseadas em DL, como avaliação de risco e gerenciamento de portfólio, os leitores interessados podem consultar outro documento de pesquisa recente [1]. Como queríamos destacar estudos de previsão de séries temporais financeiras em nossa pesquisa, omitimos outros estudos de previsão de séries temporais que não estavam focados em dados financeiros. Enquanto isso, incluímos artigos de pesquisa de séries temporais que tinham casos de uso financeiro ou exemplos, mesmo que

[∗] Autor correspondente.Endereços de e-mail: oberatsezer@etu.edu.tr (OB Sezer), mgudelek@etu.edu.tr (MU Gudelek), mobayoglu@etu.edu.tr (AM Ozbayoglu).

os próprios jornais não estavam diretamente preocupados com questões financeiras. previsão de séries temporais. Além disso, decidimos incluir algoritmos papéis comerciais que foram baseados em previsões financeiras, mas ignoram aqueles que não possuíam um componente de previsão de séries temporais.

Revisamos principalmente periódicos e conferências para nossa pesquisa, mas também incluímos teses de mestrado e doutorado, capítulos de livros, artigos arXiv e publicações técnicas notáveis que surgiram em pesquisas na web. Decidimos incluir apenas artigos publicados em Língua Inglesa.

Durante nossa pesquisa, percebemos que a maioria dos artigos que utilizam o termo "deep learning" em sua descrição foram publicados em os últimos cinco anos. No entanto, também encontramos alguns mais antigos estudos que implementaram modelos profundos, como o Recurrent Neural Redes (RNNs) e redes Jordan–Elman. No entanto, ao seu momento da publicação, o termo "deep learning" não era comum uso. Por isso, decidimos incluir também esses papéis.

De acordo com nossas descobertas, este será um dos primeiros trabalhos abrangentes de pesquisa de "previsão de séries temporais financeiras" com foco em DL. Existem muitas revisões de ML para previsão de séries temporais financeiras na literatura, mas não encontramos nenhum estudo sobre LD. Assim, queríamos preencher essa lacuna analisando o desenvolvimento modelos e aplicações em conformidade. Esperamos que como resultado neste artigo, pesquisadores e desenvolvedores de modelos terão uma melhor ideia de como eles podem implementar modelos de EAD em seus estudos.

O restante deste artigo está estruturado da seguinte forma. Na Seção 2, as pesquisas existentes se concentraram em ML e estudos de computação flexível para previsão de séries temporais financeiras são mencionados. Na Seção 3, abordaremos os modelos DL existentes que são usados, como CNN, LSTM e Deep Reinforcement Learning (DRL). A Seção 4 irá foco nas várias áreas de implementação da previsão de séries temporais financeiras usando DL, ou seja, previsão de ações, previsão de índices, previsão de tendências, previsão de commodities, previsão de volatilidade, previsão de câmbio e previsão de criptomoeda. Dentro cada subseção, a definição do problema será dada, seguida por as implementações de DL específicas. Na Seção 5, estatísticas gerais resultados sobre nossos achados serão apresentados, incluindo histogramas relacionados com as distribuições anuais de diferentes subcampos, modelos, tipos de publicação, etc. Um instantâneo de última geração de finanças estudos de previsão de séries temporais serão dados por meio dessas estatísticas. Ao mesmo tempo, também mostrarão as áreas que são já maduros em comparação com áreas promissoras ou novas que ainda tem espaço para melhorias. A Seção 6 discute as conquistas acadêmicas e industriais que foram alcançadas e expectativas futuras. A seção incluirá destaques de áreas abertas que requerem mais pesquisas. Por fim, concluímos este artigo na Seção 7, resumindo nossas descobertas.

2. Previsão de séries temporais financeiras com ML

Previsão de séries temporais financeiras e aplicativos associados têm sido estudados extensivamente por muitos anos. Quando o ML começou ganhando popularidade, os aplicativos de previsão financeira baseados em modelos de computação soft também se tornaram disponíveis. Até embora nosso foco particular seja em implementações DL de estudos de previsão de séries temporais, será benéfico mencionar pesquisas existentes cobrindo séries temporais financeiras baseadas em ML estudos de previsão para fornecer alguma perspectiva histórica.

Em nosso estudo, não incluímos nenhum documento de pesquisa que fosse focada em áreas específicas de aplicação financeira que não sejam estudos de previsão. No entanto, fomos confrontados com algumas publicações de revisão que incluíam uma mistura de estudos financeiros de séries temporais e outras aplicações financeiras. Decidimos incluir esses papéis para manter a abrangência de nossa cobertura.

Exemplos dessas publicações acima mencionadas são fornecidos aqui. Foram publicados livros sobre previsão do mercado de ações [2], desenvolvimento de sistema de negociação [3], exemplos práticos de forex e

aplicações de previsão de mercado [4] usando modelos de ML, como Redes Neurais Artificiais (RNAs), Computações Evolutivas (ECs), e Programação Genética (GP), e modelos baseados em Agentes [5].

Havia também algumas pesquisas de periódicos e conferências existentes. Bahrammirzaee et al. [6] pesquisou a previsão financeira e estudos de planejamento junto com outras aplicações financeiras usando vários Técnicas de Inteligência Artificial (IA) como RNA, Sistemas Especialistas e modelos híbridos. Zhang et al. [7] também comparou métodos de ML em diferentes aplicações financeiras, incluindo mercado de ações estudos de previsão. Em Mochon et al. [8], modelos de computação suave para o mercado, previsão de forex e sistemas de negociação foram analisados. Mullainathan e Spies [9] estudaram o processo de previsão em geral do ponto de vista econômico.

Havia também uma série de documentos de pesquisa concentrados em um modelo de ML específico. Embora esses trabalhos se concentrassem em uma técnica, as áreas de implementação geralmente abrangem várias aplicações financeiras, incluindo previsão de séries temporais financeiras. Entre esses métodos de computação suave, EC e ANN têm teve o maior interesse geral.

Em termos de estudos de EC, Chen escreveu um livro sobre Algoritmos Genéticos (GAs) e GP em Finanças Computacionais [10]. Mais tarde, os Algoritmos Evolutivos Multiobjetivos (MOEAs) foram amplamente pesquisados para várias aplicações financeiras, incluindo previsão de séries temporais [11-13]. Enquanto isso, Rada revisou EC aplicativos junto com Sistemas Especialistas para investimentos financeiros modelos [14].

Em termos de estudos de ANN, Li e Ma revisaram as implementações de RNA para previsão de preços de ações e algumas outras aplicações financeiras [15]. Tkac et al. [16] pesquisou diferentes implementações de RNA em aplicações financeiras, incluindo previsão de preços de ações. Recentemente, Elmsili e Outtaj pesquisaram aplicações de RNAs em pesquisa de economia e gestão, incluindo tempo econômico previsão em série [17].

Também houve várias pesquisas de mineração de texto focadas em aplicações financeiras, incluindo previsão de séries temporais financeiras. Mittermayer e Knolmayer compararam vários métodos de mineração de texto implementações que extraem as respostas do mercado às notícias para previsão [18]. Mitra et al. [19] focado em estudos de análise de notícias para previsão de retornos anormais para estratégias de negociação em seus pesquisa. Nassirtoossi et al. estudos de mineração de texto revisados para ações ou previsão do mercado forex [20]. Kearney et al. [21] também pesquisados previsão de séries temporais baseadas em mineração de texto e estratégias de negociação usando sentimento textual. Da mesma forma, Kumar e Ravi [22] revisou estudos de mineração de texto para previsão de forex e mercado de ações. Ultimamente, Xing et al. [23] pesquisados com base em linguagem natural estudos de previsão financeira.

Finalmente, havia pesquisas específicas de aplicação que se concentravam em implementações de previsão de séries temporais financeiras específicas. Entre esses estudos, a previsão do mercado de ações teve o mais interesse. Uma série de pesquisas foram publicadas para estudos de previsão do mercado de ações com base em vários métodos de computação suave em momentos diferentes [24-31]. Chatterjee et al. [32] e Katarya e Mahajan [33] concentrou-se no mercado financeiro baseado em ANN estudos de previsão, enquanto Hu et al. [34] centraram-se nas implementações da CE para previsão de ações e modelos de negociação algorítmica. Em uma aplicação de previsão de séries temporais diferente, os pesquisadores pesquisaram estudos de previsão de forex usando ANN [35] e vários outros técnicas de computação soft [36].

Embora existam muitas pesquisas para implementações de ML de previsão de séries temporais financeiras, DL ainda não foi pesquisado de forma abrangente, apesar do surgimento de várias implementações de EAD nos últimos anos. Esta foi a principal motivação para o nosso pesquisa. Na próxima seção, cobrimos os vários modelos de DL usados em estudos de previsão de séries temporais financeiras.

3. Aprendizado profundo

DL é um tipo de RNA que consiste em várias camadas de processamento e permite abstração de alto nível para modelar dados. A principal vantagem dos modelos DL é extrair os bons recursos dos dados de entrada automaticamente usando um procedimento de aprendizado de propósito geral.

Portanto, os modelos DL têm sido propostos para muitas aplicações, tais como: reconstrução de imagem, fala, vídeo e áudio, compreensão de linguagem natural (particularmente classificação de tópicos), análise de sentimentos, resposta a perguntas e tradução de linguagem [37]. As melhorias históricas dos modelos DL são pesquisadas em Schmidhuber et al. [38].

A previsão de séries temporais financeiras é muito popular entre os pesquisadores de ML há mais de 40 anos. A comunidade financeira foi impulsionada pela recente introdução de modelos DL para previsão financeira e suas respectivas publicações. O sucesso dos modelos DL sobre ML é o grande atrativo para pesquisadores de finanças. Com mais dados de séries temporais financeiras e

diferentes arquiteturas profundas, novos métodos de DL serão propostos. Em nossa pesquisa, a grande maioria dos estudos descobriu que os modelos DL são melhores do que seus equivalentes ML.

Na literatura, existem diferentes tipos de modelos DL: Deep Multilayer Perceptron (DMLP), RNN, LSTM, CNN, Restricted Boltzmann Machines (RBMs), DBN, Autoencoder (AE) e DRL [37, 38]. Em toda a literatura, a previsão de séries temporais financeiras é considerada principalmente como um problema de regressão. No entanto, há também um número significativo de estudos, principalmente sobre previsão de tendências, que utilizam modelos de classificação para lidar com problemas de previsão financeira. Na Seção 4, diferentes implementações de DL são apresentadas juntamente com suas escolhas de modelo.

3.1. Deep Multi Layer Perceptron (DMLP)

A DMLP foi uma das primeiras RNAs desenvolvidas. Sua diferença das redes rasas é que o DMLP contém mais camadas. Mesmo que arquiteturas de modelos particulares possam ter variações dependendo dos diferentes requisitos do problema, os modelos DMLP consistem principalmente em três camadas: entrada, oculta e saída. O número de neurônios em cada camada e o número de camadas são os hiperparâmetros da rede. Em geral, cada neurônio nas camadas ocultas tem termos de entrada (x), peso (w) e polarização (b). Além disso, cada neurônio tem uma função de ativação não linear, que produz uma saída cumulativa dos neurônios anteriores. Eq. (1) [39] ilustra a saída de um único neurônio na Rede Neural (NN). Existem diferentes tipos de funções de ativação não lineares. As funções de ativação não linear mais comumente usadas são sigmoid (Eq. (2)) [40], tangente hiperbólica (Eq. (3)) [41], Unidade Linear Retificada (ReLU)

(Eq. (4)) [42], ReLU com vazamento (Eq. (5)) [43], swish (Eq. (6)) [44] e softmax (Eq. (7)) [39]. As ativações não lineares foram comparadas em [44].

$$y_i = \tilde{y} \left(\sum_j W_{ji} x_j + b_i \right) \quad (1)$$

$$\tilde{y}(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

$$\tanh(z) = \frac{e^{-z} - e^z}{e^{-z} + e^z} \quad (3)$$

$$R(z) = \max(0, z) \quad (4)$$

$$R(z) = 1(x < 0)(\tilde{y}x) + 1(x \geq 0)(x) \quad (5)$$

$$f(x) = x\tilde{y}(\tilde{y}x) \quad (6)$$

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \quad (7)$$

Os modelos DMLP apareceram em várias áreas de aplicação [45, 37]. Usar um modelo DMLP tem vantagens e desvantagens dependendo dos requisitos do problema. Por meio de modelos DMLP, problemas como regressão e classificação podem ser resolvidos pela modelagem dos dados de entrada [46]. No entanto, se o número de recursos de entrada for aumentado (por exemplo, imagem como entrada), o tamanho do parâmetro na rede aumentará de acordo devido à natureza totalmente conectada do modelo, o que prejudicará o desempenho computacional e criará problemas de armazenamento. Para superar esse problema, diferentes tipos de métodos de Deep Neural Network (DNN) foram propostos (como CNN) [37]. Com o DMLP, processos de classificação e regressão muito mais eficientes podem ser realizados.

Na Fig. 1, as camadas de um modelo DMLP, neurônios e pesos entre os neurônios são ilustrados.

A etapa de aprendizado do DMLP é implementada por meio de retropropagação. O erro nos neurônios da camada de saída é propagado de volta para as camadas anteriores. Os algoritmos de otimização são usados para encontrar os parâmetros/variáveis ótimos dos NNs. Eles são usados para atualizar os pesos das conexões entre as camadas. Diferentes algoritmos de otimização foram desenvolvidos: Stochastic Gradient Descent (SGD), SGD com Momentum, Adaptive Gradient Algorithm (AdaGrad), Root Mean Square Propagation (RMSProp) e Adaptive Moment Estimation (ADAM) [47–51]. O gradiente de perfunto é um método iterativo para encontrar parâmetros ótimos da função que minimiza a função de custo. SGD é um algoritmo que seleciona aleatoriamente algumas amostras para cada iteração em vez de todo o conjunto de dados [47]. SGD com Momentum lembra a atualização em cada iteração, o que acelera a descida do gradiente [48]. AdaGrad é um SGD modificado que melhora o desempenho de convergência do algoritmo SGD padrão [49].

RMSProp é um algoritmo de otimização que adapta a taxa de aprendizado para cada parâmetro. No RMSProp, a taxa de aprendizado é dividida por uma média de execução das magnitudes dos gradientes recentes para aquele peso [50]. O ADAM é uma versão atualizada do RMSProp que usa médias corridas dos gradientes e segundos momentos dos gradientes. O ADAM combina as vantagens do RMSProp (funciona bem em configurações online e não estacionárias) e AdaGrad (funciona bem com gradientes esparsos) [51].

Conforme mostrado na Fig. 1, o efeito da retropropagação é transferido para as camadas anteriores. Se o efeito do SGD é gradualmente perdido quando o efeito atinge as camadas iniciais durante a retropropagação, esse problema é chamado de problema do gradiente de fuga [52]. Nesse caso, as atualizações entre as camadas iniciais ficam indisponíveis e o processo de aprendizado é interrompido. O alto número de camadas na rede neural e a complexidade crescente causam o problema do gradiente de fuga.

A questão importante no DMLP são os hiperparâmetros das redes e o método de ajuste desses hiperparâmetros.

Os hiperparâmetros são as variáveis da rede que afetam a arquitetura de rede e o desempenho das redes. O número de camadas ocultas, número de unidades em cada camada, técnicas de regularização (dropout, L1, L2), inicialização do peso da rede (zero, random, He [53], Xavier [54]), funções de ativação (Sigmoid, ReLU, tangente hiperbólica, etc.), taxa de aprendizado, taxa de decaimento, valores de momento, número de épocas, tamanho do lote (tamanho do mini lote) e algoritmos de otimização (SGD, AdaGrad, RMSProp, ADAM, etc.) são os hiperparâmetros do DMLP. Escolher melhores valores/variáveis de hiperparâmetros para a rede resulta em melhor desempenho. Portanto, encontrar os melhores hiperparâmetros para a rede é uma questão importante. Na literatura, existem diferentes métodos para encontrar os melhores hiperparâmetros: Manual Search (MS), Grid Search (GS), RandomSearch (RS) e Bayesian Methods (Sequential Model-Based Global Optimization (SMBGO), The Gaussian Process Approach (GPA)), Abordagem do Estimador de Parzen Estruturado em Árvore (TSPEA)) [55,56].

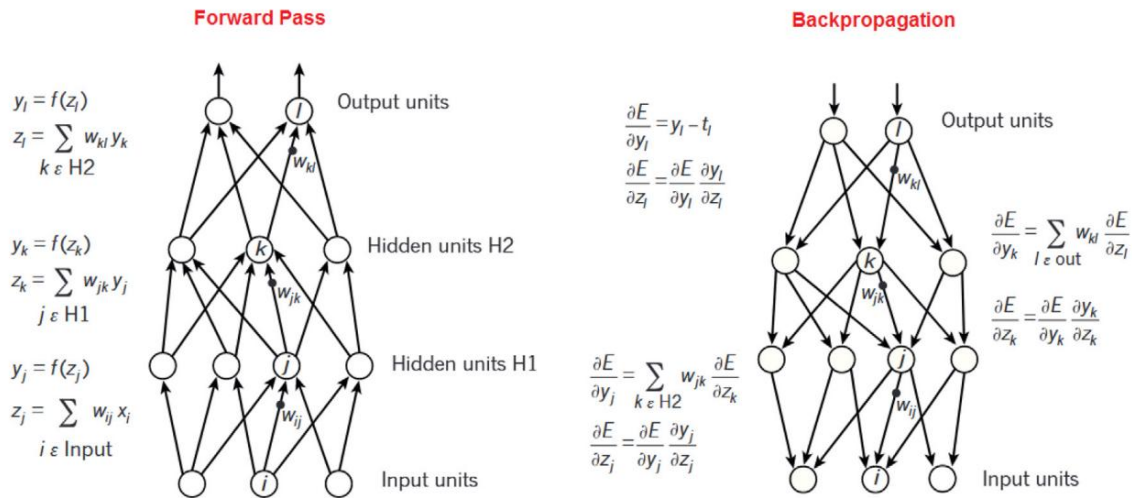


Fig. 1. Rede neural multicamada profunda com passagem direta e retropropagação [37].

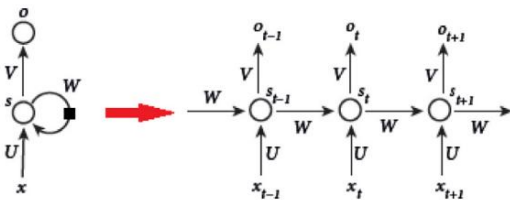


Fig. 2. Célula RNN ao longo do tempo [37].

3.2. Rede Neural Recorrente (RNN)

A RNN é outro tipo de rede DL usada para séries temporais ou dados sequenciais, como linguagem e fala. RNNs também são usados em modelos tradicionais de ML (Back Propagation Through Time (BPTT), redes Jordan–Elman, etc.); no entanto, os períodos de tempo em tais modelos são geralmente menores do que aqueles usados em modelos RNN profundos.

Deep RNNs são preferidos devido à sua capacidade de incluir períodos de tempo mais longos. Ao contrário das Redes Neurais Totalmente Conectadas (FNNs), as RNNs usam memória interna para processar as entradas de entrada. RNNs são usados para analisar dados de séries temporais em vários campos (reconhecimento de manuscrito, reconhecimento de fala, etc.). Conforme declarado na literatura, as RNNs são boas em prever o próximo caractere no texto, aplicativos de tradução de idiomas e processamento sequencial de dados [45,37].

A arquitetura do modelo RNN consiste em diferentes números de camadas e diferentes tipos de unidades em cada camada. A principal diferença entre RNN e FNN é que cada unidade RNN leva o dados de entrada atuais e anteriores ao mesmo tempo. A saída depende dos dados anteriores no modelo RNN. As RNNs processam as sequências de entrada uma a uma a qualquer momento durante sua operação.

As unidades na camada oculta contêm informações sobre o histórico da entrada no "vetor de estado". Quando a saída das unidades na camada oculta é dividida em diferentes passos de tempo discretos, um RNN é convertido em um DMLP [37]. Na Fig. 2, o fluxo de informações na camada oculta da RNN é dividido em tempos discretos. O estado do nó S em diferentes momentos de t é mostrado como s_t , o valor de entrada x em diferentes momentos é x_t , e o valor da saída o em diferentes momentos é mostrado como o_t .

Os valores dos parâmetros (U , W , V) são sempre usados na mesma etapa.

As RNNs podem ser treinadas usando o algoritmo BPTT. Algoritmos de otimização (SGD, RMSProp, ADAM) são usados para o processo de ajuste de peso. Com o método de aprendizado BPTT, a mudança de erro no tempo t é refletida na entrada e nos pesos dos tempos t anteriores. A dificuldade de treinar um RNN é que o

A estrutura RNN tem uma dependência inversa ao longo do tempo. Portanto, as RNNs se tornam cada vez mais complexas à medida que o período de aprendizado aumenta. Embora o objetivo principal de usar uma RNN seja aprender dependências de longo prazo, estudos na literatura mostram que quando o conhecimento é armazenado por longos períodos de tempo, não é fácil aprender com uma RNN [57]. Para resolver este problema em particular, foram desenvolvidos LSTMs com diferentes estruturas de RNA [37]. Eqs. (8) e (9) ilustram formulações de RNN mais simples. Eq. (10) mostra o erro total, que é a soma dos erros de cada iteração de tempo. $1 \text{ ht} = Wf(ht\bar{y}1) + W(hx) x[t]$

(8)

$$yt = W(S) f(ht)$$

(9)

$$\frac{\bar{y}E}{\bar{y}W} = \sum_{t=1}^T \frac{\bar{y}Et}{\bar{y}W}$$

(10)

Os hiperparâmetros do RNN também definem a arquitetura da rede, e o desempenho da rede é afetado pelas escolhas dos parâmetros, como no caso do DMLP. O número de camadas ocultas, número de unidades em cada camada, técnicas de regularização, inicialização do peso da rede, funções de ativação, taxa de aprendizado, valores de momento, número de épocas, tamanho do lote (tamanho do minilote), taxa de decaimento, algoritmos de otimização, modelo (Vanilla RNN, Gated Recurrent Unit (GRU), LSTM) e o comprimento da sequência são os hiperparâmetros de RNN. Encontrar os melhores hiperparâmetros para a rede é um problema significativo. Na literatura, existem diferentes métodos para encontrar os melhores hiperparâmetros: MS, GS, RS e métodos Bayesianos (SMBGO, GPA, TSPEA) [55,56].

3.3. Memória de longo prazo (LSTM)

LSTM [58] é um tipo de RNN onde a rede pode lembrar valores de curto e longo prazo. As redes LSTM são a escolha preferida de muitos desenvolvedores de modelos DL ao lidar com problemas complexos, como fala automática e reconhecimento de caracteres manuscritos. Os modelos LSTM são usados principalmente com dados de séries temporais. Suas aplicações incluem Processamento de Linguagem Natural (NLP), modelagem de linguagem, tradução de linguagem, reconhecimento de fala, análise de sentimento, análise preditiva e análise de séries temporais financeiras [59,60]. Com módulos de atenção e estruturas de EA, as redes LSTM podem ser mais bem-sucedidas na análise de dados de séries temporais, como tradução de idiomas [59].

¹ Richard Socher, CS224d: Aprendizado profundo para processamento de linguagem natural, Notas de Aula

As redes LSTM consistem em unidades LSTM. As unidades LSTM se fundem para formar uma camada LSTM. Uma unidade LSTM é composta de células, cada uma com uma porta de entrada, uma porta de saída e uma porta de esquecimento. Esses portões regulam o fluxo de informações. Com esses recursos, cada célula lembra os valores desejados em intervalos de tempo arbitrários. Eqs. (11)–(15) mostram a forma da passagem direta da unidade LSTM [58] (x_t : vetor de entrada para a unidade LSTM, f_t : vetor de ativação da porta de esquecimento, i_t : vetor de ativação da porta de entrada, o_t : ativação da porta de saída, h_t : vetor de saída da unidade LSTM, c_t : vetor de estado da célula, γ_g : função sigmóide, γ_c , γ_h : função tangente hiperbólica, $\dot{\gamma}$: produto elemento a elemento (Hadamard), W , U : matrizes de peso a serem aprendidas, b : parâmetros do vetor de polarização a serem aprendidos) [60].

$$f_t = \gamma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (11)$$

$$i_t = \gamma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (12)$$

$$o_t = \gamma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (13)$$

$$c_t = f_t \dot{\gamma} c_{t-1} + i_t \dot{\gamma} \gamma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (14)$$

$$h_t = o_t \dot{\gamma} \gamma_h(c_t) \quad (15)$$

LSTM é uma versão especializada do RNN. Portanto, as atualizações de peso e os métodos de otimização preferidos são os mesmos. Além disso, os hiperparâmetros do LSTM são iguais aos do RNN: número de camadas ocultas, número de unidades em cada camada, inicialização do peso da rede, funções de ativação, taxa de aprendizado, valores de momento, número de épocas, tamanho do lote (tamanho do minilote), taxa de decaimento, algoritmos de otimização, comprimento de sequência para LSTM, recorte de gradiente, normalização de gradiente e dropout [60,61]. Para encontrar os melhores hiperparâmetros do LSTM, os métodos de otimização de hiperparâmetros usados para RNN também são aplicáveis ao LSTM [55,56].

3.4. Redes neurais convolucionais (CNNs)

A CNN é um tipo de DNN que consiste em camadas convolucionais baseadas na operação convolucional. É o modelo mais comum usado para problemas de classificação baseados em visão e processamento de imagens (classificação de imagens, detecção de objetos, segmentação de imagens, etc.) [62–64]. A vantagem da CNN é o número de parâmetros em comparação com os modelos DL vanilla, como o DMLP.

A filtragem com a função de janela do kernel oferece a vantagem de processamento de imagem para arquiteturas CNN com menos parâmetros, o que é benéfico para computação e armazenamento. Nas arquiteturas CNN, existem diferentes camadas: convolucional, max-pooling, dropout e camada Multilayer Perceptron (MLP) totalmente conectada.

A camada convolucional consiste em uma operação de convolução (filtragem). Uma operação de convolução básica é mostrada na Eq. (16), onde t denota tempo, s denota mapa de características, w denota kernel, x denota entrada e a denota variável. Além disso, a operação de convolução é implementada em imagens bidimensionais. Eq. (17) mostra a operação de convolução para uma imagem bidimensional, onde I denota a imagem de entrada, K denota o kernel, (m, n) denota as dimensões da imagem e i e j denotam variáveis. Camadas convolucionais consecutivas e de pool máximo constroem a rede profunda.

Eq. (18) descreve a arquitetura NN, onde W denota pesos, x denota entrada, b denota polarização e z denota a saída de neurônios. No final da rede, a função softmax é usada para obter a saída. Eqs. (19) e (20) ilustram a função softmax, onde y denota saída [39].

$$s(t) = (x \dot{\gamma} w)(t) = \dot{\gamma} \dot{\gamma} x(a)w(t \dot{\gamma} a) \quad (16)$$

$$S(i, j) = (I \dot{\gamma} K)(i, j) = \dot{\gamma} \dot{\gamma} I(m)K(i \dot{\gamma} m, j \dot{\gamma} n). \quad (17)$$

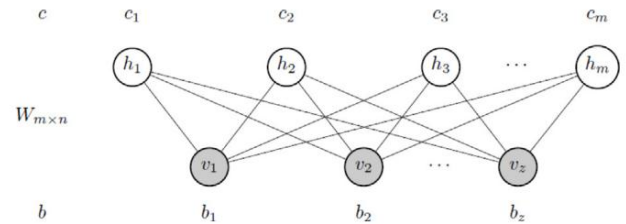


Fig. 3. Camadas visíveis e ocultas do RBM [65].

$$z_i = \dot{\gamma} W_{ij} x_j + b_i \quad (18)$$

$$y = \text{softmax}(z) \quad (19)$$

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \quad (20)$$

O processo de retropropagação é usado para aprendizado do modelo CNN. Os algoritmos de otimização mais comumente usados (SGD, RMSProp) são usados para encontrar os parâmetros ótimos da CNN. Os hiperparâmetros da CNN são semelhantes a outros hiperparâmetros do modelo DL: número de camadas ocultas, número de unidades em cada camada, inicialização do peso da rede, funções de ativação, taxa de aprendizado, valores de momento, número de épocas, tamanho do lote (tamanho do minilote), decaimento taxa, algoritmos de otimização, dropout, tamanho do kernel e tamanho do filtro. Para encontrar os melhores hiperparâmetros CNN, os seguintes algoritmos de busca são comumente usados: métodos MS, GS, RS e Bayesiano. [55,56].

3.5. Máquinas Boltzmann restritas (RBMs)

Uma RBM é uma RNA estocástica produtiva que pode aprender uma distribuição de probabilidade no conjunto de entrada [65]. RBMs são usados principalmente para aprendizado não supervisionado [66]. RBMs são usados em aplicações como redução de dimensão, classificação, aprendizado de recursos e filtragem colaborativa [67]. A vantagem dos RBMs é sua capacidade de encontrar padrões ocultos de maneira não supervisionada. A desvantagem dos RBMs é seu difícil processo de treinamento. "RBMs são complicados porque, embora existam bons estimadores do gradiente de probabilidade logarítmica, não existem maneiras baratas conhecidas de estimar a probabilidade logarítmica em si" [68].

Um RBM é um modelo gráfico de duas camadas, bipartido e não direcionado que consiste em duas camadas: visível e oculta (Fig. 3). As camadas não estão conectadas entre si. Cada célula é um ponto computacional que processa a entrada e toma decisões estocásticas sobre se esse nódulo nervoso transmitirá a entrada.

As entradas são multiplicadas por pesos específicos, determinados valores de limite (bias) são adicionados aos valores de entrada e, em seguida, os valores calculados são passados por uma função de ativação. No estágio de reconstrução, os resultados das saídas reentram na rede como entrada antes de finalmente sair da camada visível como saída. Os valores da entrada anterior e os valores após os processos são comparados. O objetivo desta comparação é reduzir a diferença.

Eq. (21) ilustra a semântica probabilística para um RBM usando sua função de energia, onde P denota a semântica probabilística para um RBM, Z denota a função de partição, E denota a função de energia, h denota unidades ocultas e v denota unidades visíveis. Eq. (22) ilustra a função de partição ou constante de normalização. Eq. (23) mostra a energia de uma configuração (em notação matricial) de um RBM padrão com unidades ocultas e visíveis de valor binário, onde a denota pesos de polarização (desvios) para as unidades visíveis, b denota pesos de polarização para as unidades ocultas, W denota o peso da matriz da conexão entre oculto e visível

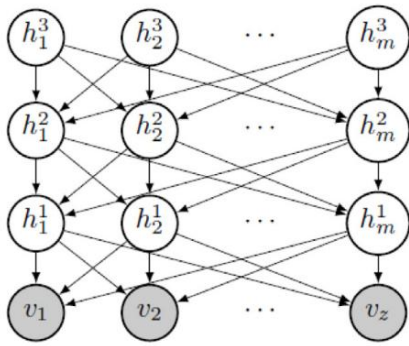


Fig. 4. Rede de crenças profundas [65].

unidades, T denota a transposição da matriz, v denota unidades visíveis, e h denota unidades ocultas [69,70].

$$P(v, h) = \frac{1}{Z} \exp(\tilde{y} E(v, h)) \quad (21)$$

$$Z = \sum_h \tilde{y} \exp(\tilde{y} E(v, h)) \quad (22)$$

$$E(v, h) = \tilde{y} a^T v \tilde{y} b h \tilde{y} v T W h \quad (23)$$

O aprendizado é realizado várias vezes na rede [65].

O treinamento de RBMs é implementado minimizando a probabilidade logarítmica negativa do modelo e dos dados. O algoritmo Contrastive Divergence (CD) é usado como algoritmo de aproximação estocástica, que substitui a expectativa do modelo usando uma estimativa usando Gibbs Sampling com um número limitado de iterações [66]. No algoritmo CD, o algoritmo Kullback Leibler Divergence (KL-Divergence) é usado para medir a distância entre sua distribuição de probabilidade reconstruída e a distribuição de probabilidade original da entrada [71].

Momentum, taxa de aprendizado, peso-custo (taxa de decaimento), tamanho do lote (tamanho do minilote), método de regularização, número de épocas, número de camadas, inicialização de pesos, tamanho das unidades visíveis, tamanho das unidades ocultas, tipo de unidades de ativação (sigmoid, softmax, ReLU, unidades gaussianas), função de perda e algoritmos de otimização são os hiperparâmetros de RBMs. Semelhante a outras redes profundas, os hiperparâmetros são pesquisados com os métodos MS, GS, RS e Bayesiano (processo Gaussiano). Além disso, a Amostragem de Importância Recozida (AIS) é usada para estimar a função de partição. O algoritmo CD também é usado para a otimização de RBMs [55, 56, 72, 73].

3.6. Redes de crenças profundas (DBNs)

Um DBN é um tipo de RNA profunda que consiste em uma pilha de redes RBM (Fig. 4). Um DBN é um modelo generativo probabilístico que consiste em variáveis latentes. Em um DBN, não há link entre as unidades em cada camada. DBNs são usados para encontrar características independentes discriminadas no conjunto de entrada usando aprendizado não supervisionado [69].

A capacidade de codificar estruturas de rede de ordem superior e inferência rápida são as vantagens dos DBNs [74]. Os DBNs têm as mesmas desvantagens dos RBMs porque os DBNs são compostos de RBMs.

Quando um DBN é treinado de maneira não supervisionada, ele pode aprender a reconstruir o conjunto de entrada de maneira probabilística. Então, as camadas da rede começam a detectar características discriminantes na entrada. Após esta etapa de aprendizado, o aprendizado supervisionado é conduzido para classificação [75]. Eq. (24) ilustra a probabilidade de gerar um vetor visível (W : peso da matriz de conexão entre a unidade oculta h e a unidade visível v , $p(h|W)$: distribuição a priori sobre vetores ocultos) [69].

$$p(v) = \prod_h p(h|W) p(v|h, W) \quad (24)$$

O processo de treinamento do DBN pode ser dividido em duas etapas:

aprendizado RBM empilhado e aprendizado por retropropagação. No aprendizado RBM empilhado, um algoritmo de CD iterativo é usado [66]. No aprendizado de backprop, algoritmos de otimização (SGD, RMSProp, ADAM) são usados para treinar a rede [74]. Os hiperparâmetros de um DBNs são semelhantes aos de um RBM. Momentum, taxa de aprendizado, peso-custo (taxa de decaimento), método de regularização, tamanho do lote (tamanho do mini lote), número de épocas, número de camadas, inicialização de pesos, número de pilhas RBM, tamanho das unidades visíveis nas camadas dos RBMs, tamanho das unidades ocultas nas camadas dos RBMs, tipo de unidades (sigmoid, softmax, retificadas, unidades gaussianas, etc.), inicialização do peso da rede e algoritmos de otimização são os hiperparâmetros dos DBNs. Semelhante a outras redes profundas, os hiperparâmetros são pesquisados com os métodos MS, GS, RS e Bayesiano. O algoritmo CD também é usado para a otimização de DBNs [55, 56, 72, 73].

3.7. Codificadores de carro (AEs)

As redes AE são RNAs usadas como modelos de aprendizado não supervisionado. Além disso, as redes de EA são comumente usadas em modelos DL, onde elas remapeam as entradas (características) de forma que as entradas sejam mais representativas para classificação. Em outras palavras, as redes de EA realizam um processo de aprendizado de recursos não supervisionado, que se encaixa muito bem com o framework DL. Uma representação de um conjunto de dados é aprendida reduzindo a dimensionalidade com EAs. Os AEs são semelhantes às Redes Neurais Feedforward (FFNNs) em sua arquitetura. Eles consistem em uma camada de entrada, uma camada de saída e uma ou mais camadas ocultas que os conectam. O número de nós na camada de entrada e o número de nós na camada de saída são iguais entre si nos EAs e possuem uma estrutura simétrica. As vantagens mais notáveis dos EAs são a redução de dimensionalidade e o aprendizado de recursos. Enquanto isso, a redução da dimensionalidade e a extração de recursos em EAs causam algumas desvantagens. Concentrar-se em minimizar a perda de relacionamento de dados na codificação de EAs causa a perda de alguns relacionamentos de dados significativos. Portanto, isso pode ser considerado uma desvantagem dos EAs [76].

Em geral, os AEs contêm dois componentes: codificador e decodificador. A entrada $x \in [0, 1]$ é convertida através da função $f(x)$ ($W1$ denota uma matriz de peso, $b1$ denota um vetor de polarização, $\tilde{y}1$ função de ativação sigmóide do codificador). A saída h é a parte codificada dos EAs (código), variáveis latentes ou representação latente. O inverso da função $f(x)$, chamada função $g(h)$, produz a reconstrução da saída r ($W2$ denota uma matriz de pesos, $b2$ denota um vetor de polarização e $\tilde{y}2$ é uma função de ativação sigmóide elemento a elemento do decodificador). Eqs. (25) e (26) ilustram o processo simples de EA [77]. Eq. (27) mostra a função de perda do AE, o Erro Quadrado Médio (MSE). Na literatura, EAs têm sido usados para extração de características e redução de dimensionalidade [39, 77].

$$h = f(x) = \tilde{y}1(W1x + b1) \quad (25)$$

$$r = g(h) = \tilde{y}2(W2h + b2) \quad (26)$$

$$L(x, r) = \frac{1}{2} \|x - r\|^2 \quad (27)$$

EAs são uma versão especializada de FFNNs. O aprendizado de retropropagação é usado para atualizar os pesos na rede [39]. Algoritmos de otimização (SGD, RMSProp, ADAM) são usados para o processo de aprendizagem de EAs. O MSE é usado como uma função de perda em EAs. Além disso, algoritmos de recirculação também podem ser usados para o treinamento de EAs [39]. Os hiperparâmetros dos EAs são semelhantes aos dos hiperparâmetros DL. Taxa de aprendizado, custo de peso (taxa de decaimento), fração de desistência, tamanho do lote (tamanho do minilote), número de épocas, número de camadas, número de nós em cada camada do codificador, tipo

de funções de ativação, número de nós em cada camada do decodificador, inicialização do peso da rede, algoritmos de otimização e número de nós na camada de código (tamanho da representação latente) são os hiperparâmetros dos EAs. Semelhante a outras redes profundas, os hiperparâmetros são pesquisados com métodos MS, GS, RS e Bayesianos [55,56].

3.8. Aprendizado por reforço profundo (DRL)

O aprendizado por reforço (RL) é um tipo de aprendizado que difere dos modelos de aprendizado supervisionado e não supervisionado. Não requer um conjunto de dados preliminar que tenha sido rotulado ou agrupado anteriormente. RL é uma abordagem de ML inspirada na ação/comportamento de aprendizagem, que trata de quais ações devem ser tomadas pelos sujeitos para alcançar a maior recompensa em um ambiente. Existem diferentes áreas em que é usado: teoria dos jogos, teoria do controle, sistemas multi-agentes, pesquisa operacional, robótica, teoria da informação, gerenciamento de portfólio de investimentos, otimização baseada em simulação, jogos Atari e estatística [78]. Algumas vantagens do uso de RL para problemas de controle são que um agente pode ser facilmente retreinado para se adaptar às mudanças no ambiente e que o sistema é continuamente aprimorado enquanto o treinamento é realizado constantemente.

Um agente RL aprende interagindo com seu entorno e observando os resultados dessas interações. Este método de aprendizagem imita o básico de como os humanos aprendem.

O RL é baseado principalmente em um Processo de Decisão de Markov (MDP). Um MDP é usado para formalizar o ambiente RL. Um MDP consiste em cinco tuplas: estado (conjunto finito de estados), ação (conjunto finito de ações), função de recompensa (sinal de feedback escalar), matriz de probabilidade de transição de estado, e γ denota o fator de desconto (valor presente das recompensas futuras). O objetivo do agente é maximizar a recompensa cumulativa. O retorno (G_t) é a recompensa total descontada. A Eq. (28) ilustra o retorno total, onde R denota recompensas, t denota tempo e k denota uma variável no tempo.

$$G_t = R_t + \gamma R_{t+1} + \gamma^2 R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (28)$$

A função valor é a previsão de valores futuros. Fornece informações sobre a qualidade do estado/ação. Eq. (29) ilustra a formulação da função de valor, onde $v(s)$ denota a função de valor, $E[\cdot]$ denota a função de expectativa, G_t denota a recompensa total descontada, s denota o estado dado, R denota as recompensas, S denota o conjunto de estados e t denota o tempo.

$$v(s) = E[G_t | S_t = s] = E[R_t + \gamma v(S_{t+1}) | S_t = s] \quad (29)$$

Política (π) é a estratégia de comportamento do agente. É como um mapa do estado à ação. Existem dois tipos de funções de valor para expressar as ações na política: função estado-valor ($v(s)$) e função ação-valor ($q(s, a)$). A função estado-valor (Eq. (30)) é o retorno esperado de começar de s para seguir a política π ($E\pi[\cdot]$ denota função de expectativa). A função valor-ação (Eq. (31)) é o retorno esperado de começar de s e realizar a ação a para seguir a política π (A denota o conjunto de ações e a denota a ação dada).

$$v(s) = E\pi[G_t | S_t = s] = E\pi[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s] \quad (30)$$

$$q\pi(s, a) = E\pi[G_t | S_t = s, A_t = a] \quad (31)$$

A função valor-estado ótimo (Eq. (32)) é a função valor máximo sobre todas as políticas. A função ótima de ação-valor

(Eq. (33)) é a função valor-ação máxima sobre todas as políticas.

$$v^*(s) = \max_{\pi}(v^{\pi}(s)) \quad (32)$$

$$q^*(s, a) = \max_{\pi}(q^{\pi}(s, a)) \quad (33)$$

As soluções e métodos de RL na literatura são muito amplos para serem revisados neste artigo. Portanto, resumimos as questões importantes de RL e soluções e métodos importantes de RL. Os métodos RL podem ser divididos principalmente em dois tipos: métodos baseados em modelos e métodos livres de modelos. Os métodos baseados em modelo usam um modelo que é conhecido pelo agente antes, valor/política e experiência. A experiência pode ser real (amostra do ambiente) ou simulada (amostra do modelo). Métodos baseados em modelos são usados principalmente em aplicações de robótica e algoritmos de controle [79]. Os métodos livres de modelo são divididos principalmente em dois grupos: baseados em valor e baseados em políticas. Nos métodos baseados em valor, uma política é produzida diretamente da função de valor (por exemplo, ϵ -greedy). Nos métodos baseados em política, a política é parametrizada diretamente. Nos métodos baseados em valor, existem três soluções principais para problemas de MDP: Programação Dinâmica (DP), Monte Carlo (MC) e Diferença Temporal (TD).

No método DP, os problemas são resolvidos com subestrutura ótima e subproblemas sobrepostos. O modelo completo é conhecido e é utilizado para planejamento no MDP. Existem duas iterações (algoritmos de aprendizado) no DP: iteração de política e iteração de valor. O método MC aprende a experiência diretamente executando um episódio de jogo/simulação. MC é um tipo de método sem modelo que não requer transições/recompensas do MDP. Ele coleta estados e obtém a média dos retornos para a função valor. O TD também é um método livre de modelo que aprende a experiência diretamente executando o episódio. Além disso, o TD aprende episódios incompletos como o método DP usando bootstrapping. O método TD combina os métodos MC e DP. SARSA (estado, ação, recompensa, estado, ação; $S_t, A_t, R_t, S_{t+1}, A_{t+1}$) é um tipo de algoritmo de controle TD. O valor Q (função valor-ação) é atualizado com as ações do agente. É um modelo de aprendizagem on-policy que aprende com as ações de acordo com a política atual π . Eq. (34) ilustra a atualização da função valor-ação no algoritmo SARSA, onde S_t denota estado atual, A_t denota ação atual, t denota tempo, R denota recompensa, γ denota taxa de aprendizado, γ denota fator de desconto. Q-learning é outro algoritmo de controle TD. É um modelo de aprendizagem fora da política que aprende com diferentes ações que não requerem a política π . Eq. (35) ilustra a atualização da função valor-ação no algoritmo Q-Learning (todo o algoritmo é descrito em [78], um

γ
denota ação).

$$Q(S_t, A_t) = Q(S_t, A_t) + \gamma[R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)] \quad (34)$$

$$Q(S_t, A_t) = Q(S_t, A_t) + \gamma[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \quad (35)$$

Nos métodos baseados em valor, uma política pode ser gerada diretamente da função de valor (por exemplo, usando ϵ -greedy). Os métodos baseados em política usam a política diretamente em vez de usar a função de valor. Tem vantagens e desvantagens sobre os métodos baseados em valor. Os métodos baseados em políticas são mais eficazes em espaços de ação de alta dimensão ou contínuos e têm melhores propriedades de convergência do que os métodos baseados em valor. Ele também pode aprender políticas estocásticas. Por outro lado, os métodos baseados em políticas avaliam uma política que normalmente é ineficiente e tem alta variação. Normalmente converge para um ótimo local em vez de global. Nos métodos baseados em políticas, existem também soluções diferentes: Gradiente de política, Reforçar (Gradiente de política de Monte Carlo) e Ator-Crítico [78] (detalhes dos métodos baseados em políticas podem ser encontrados em [78]).

Os métodos DRL contêm NNs. Portanto, os hiperparâmetros DRL são semelhantes aos hiperparâmetros DL. Taxa de aprendizagem, peso-custo (taxa de decaimento), fração de abandono, método de regularização, tamanho do lote (tamanho do minilote), número de épocas, número de camadas, número de nós em cada camada, tipo de funções de ativação, inicialização do peso da rede, algoritmos de otimização, fator de desconto e número de episódios são os hiperparâmetros do DRL. Semelhante a outras redes profundas, os hiperparâmetros são pesquisados com os métodos MS, GS, RS e Bayesiano [55,56].

4. Previsão de séries temporais financeiras

A área de aplicação financeira mais estudada é a previsão de uma determinada série temporal financeira, particularmente a previsão de preços de ativos. Embora existam algumas variações, o foco principal é prever o próximo movimento do ativo subjacente. Mais da metade das implementações existentes de EAD estão focadas nesta área. Embora existam vários subtópicos desse problema geral, incluindo previsão de preço de ações, previsão de índice, previsão de preço de forex, previsão de preço de commodities (petróleo, ouro, etc.), previsão de preço de títulos, previsão de volatilidade, previsão de preço de criptomoeda, a dinâmica subjacente são os mesmos em todas essas aplicações.

Os estudos também podem ser agrupados em dois grupos principais com base em seus resultados esperados: previsão de preço e previsão de movimento de preço (tendência). Embora a previsão de preços seja essencialmente um problema de regressão, na maioria das aplicações de previsão de séries temporais financeiras, a previsão correta do preço não é percebida como tão importante quanto identificar corretamente o movimento direcional. Como resultado, os pesquisadores consideram a previsão de tendências, ou seja, a previsão de como o preço mudará, uma área de estudo mais crucial em comparação com a previsão exata de preços. Nesse sentido, a previsão de tendências torna-se um problema de classificação. Em alguns estudos, apenas os movimentos para cima ou para baixo são levados em consideração (problema de 2 classes), embora também existam problemas de 3 classes (movimentos para cima, para baixo ou neutros).

O LSTM e suas variações, juntamente com alguns modelos híbridos, dominam o domínio de previsão de séries temporais financeiras. O LSTM, por sua natureza, utiliza as características temporais de qualquer sinal de série temporal; portanto, a previsão de séries temporais financeiras é uma implementação bem estudada e bem-sucedida do LSTM. No entanto, alguns pesquisadores preferem extrair características apropriadas da série temporal ou transformar a série temporal de modo que os dados financeiros resultantes se tornem estacionários de uma perspectiva temporal, o que significa que mesmo se embaralharmos a ordem dos dados, ainda poderemos treinar adequadamente o modelo e alcançar um desempenho de teste bem-sucedido fora da amostra. Para essas implementações, CNN e Deep Feedforward Neural Network (DFNN) são os modelos DL mais comumente escolhidos.

Várias implementações de previsão de séries temporais financeiras usando modelos DL existem na literatura. Abordaremos cada um deles nas subseções a seguir. Neste trabalho de pesquisa, examinamos os artigos usando os seguintes critérios: Primeiro, agrupamos os artigos de acordo com seus assuntos. Em seguida, agrupamos os artigos relacionados de acordo com seu conjunto de recursos. Por fim, agrupamos cada subgrupo de acordo com modelos/métodos de EAD.

Para cada área de implementação, os artigos relacionados são subagrupados e tabulados. Cada tabela contém os seguintes campos para fornecer informações sobre os detalhes de implementação para os artigos dentro do grupo: Artigo (Art.) e Conjunto de dados são triviais, Período refere-se ao período de tempo para treinamento e teste. Conjunto de recursos lista os recursos de entrada usados no estudo. Lag é a duração de tempo do vetor de entrada (por exemplo, 30 d significa que o vetor de entrada tem uma janela de 30 dias), e horizonte mostra até onde o modelo prevê no futuro. Algumas abreviações são usadas para os dois campos acima mencionados: min é minutos, h é horas, d é dias, w é semanas,

m é meses, y é anos, s é passos e * é misto. O método mostra os modelos DL que são utilizados no estudo. Os critérios de desempenho fornecem as métricas de avaliação e o Ambiente (Ambiente) lista a estrutura/software/ferramentas de desenvolvimento. Alguns valores de coluna podem estar vazios, indicando que não havia informações relevantes no papel para o campo correspondente.

4.1. Previsão do preço das ações

A previsão de preços de qualquer ação é a aplicação financeira mais estudada de todas. Observamos a mesma tendência nas implementações de EAD. Dependendo do horizonte de tempo de previsão, diferentes parâmetros de entrada são escolhidos, variando de negociação de alta frequência (HFT) e movimentos de preços intradiários a preços de fechamento de ações diários, semanais ou até mensais. Além disso, análise técnica, fundamental, feeds de mídia social e sentimento estão entre os diferentes parâmetros usados para os modelos de previsão.

Nesta pesquisa, agrupamos os primeiros artigos de previsão de preços de ações de acordo com seus conjuntos de recursos, como estudos usando apenas os dados brutos de séries temporais (dados de preços, Aberto, Fechado, Alto, Baixo, Volume (OCHLV)) para previsão de preço; estudos usando vários outros dados e estudos usando técnicas de mineração de texto. Em relação ao primeiro grupo, os modelos DL correspondentes foram implementados diretamente usando séries temporais brutas para previsão de preços. A Tabela 1 tabula os estudos de previsão de preços de ações que utilizaram apenas dados brutos de séries temporais na literatura. Na Tabela 1, diferentes métodos/modelos também são listados com base em quatro subgrupos: DNN (redes profundas, mas sem detalhes de topologia) e modelos LSTM, multimodelos, modelos híbridos, métodos novos.

Os modelos DNN e LSTM foram usados apenas em 3 artigos. Em Chong et al. [80], DNN e retornos de ações defasados foram usados para prever os preços das ações no Índice de Preços de Ações Composto da Coreia (KOSPI). Chen et al. [81], e Dezsi e Nistor [82] aplicaram dados brutos de preços como entrada para modelos LSTM.

Enquanto isso, alguns estudos implementam vários modelos DL para comparação de desempenho usando apenas dados brutos de preço (OCHLV) para previsão. Entre os estudos de destaque, Samarawickrama et al. [83] compararam RNN, Stacked Recurrent Neural Network (SRNN), LSTM e GRU. Hiransha et al. [84] compararam LSTM, RNN, CNN e MLP, enquanto em Selvin et al. [85], RNN, LSTM, CNN e Autoregressive Integrated Moving Average (ARIMA) foram os preferidos. Lee e Yoo [86] compararam 3 modelos RNN (SRNN, LSTM, GRU) para previsão de preços de ações e então construíram um portfólio baseado em limiares selecionando ações de acordo com as previsões. Li et al. [87] implementou DBN. Finalmente, os autores de [88] compararam 4 modelos de ML diferentes para a previsão de preço seguinte em dados de preço de 1 minuto: um modelo de 1 DL (AE e RBM), MLP, Rede Neural de Função de Base Radial (RBF) e Extreme Learning Machine (ELM). Eles também compararam os resultados para conjuntos de dados de tamanhos diferentes.

Os autores de [89] usaram dados de preços e métodos DNN, Gradient Boosted Trees (GBT) e Random Forest (RF) para a previsão de ações no Standard's & Poor's 500 Index (S&P500).

Chandra e Chan [90] usaram neuroevolução cooperativa, RNN (rede Elman) e DFNN para a previsão de preços de ações na National Association of Securities Dealers Automated Quotations (NASDAQ) (ACI Worldwide, Staples e Seagate).

Enquanto isso, modelos híbridos foram usados em alguns trabalhos. Liu et al. [91] aplicaram CNN+LSTM. Heaton et al. [92] implementaram a indexação inteligente com AE. Batres et al. [93] combinaram DBN e MLP para construir uma carteira de ações prevendo o log-retorno mensal de cada ação e escolhendo apenas ações que deveriam ter um desempenho melhor do que a ação mediana.

Além disso, novas abordagens foram adaptadas em alguns estudos. Yuan et al. [94] propuseram o romance Deep and Wide Neural Network (DWNN), que é uma combinação de RNN e CNN. Zhang et al. [95] implementou uma rede recorrente State Frequency Memory (SFM).

tabela 1

Previsão de preços de ações usando apenas dados brutos de séries temporais.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Critérios de desempenho	Env.	
[80]	38 ações na KOSPI 2010–2014		Estoque defasado retorna	50 minutos	5 minutos	DN	NMSE, RMSE, MAE, MIM	–
[81]	ações da China mercado, 3049 Ações	1990–2015	OCHLV	30 dias	3 dias	LSTM	Precisão	Theano, Keras
[82]	Devoluções diárias de Estoque 'BRD' em Mercado romeno	2001–2016	OCHLV	–	1 dia	LSTM	RMSE, MAE	Python, Theano
[83]	297 listados empresas do CSE	2012–2013	OCHLV	2 dias	1 dia	LSTM, SRNN, GRU MAD, MAPE		Difícil
[84]	5 ações na NSE	1997–2016	OCHLV, Dados de preços, volume de negócios e número de negócios.	200 dias	1..10 dias	LSTM, RNN, CNN, MLP	MAPA	–
[85]	Ações da Infosys, TCS e CIPLA do NSE	2014	Dados de preço	–	–	RNN, LSTM e CNN	Precisão	–
[86]	10 ações em S&P500	1997–2016	OCHLV, Dados de preço 36 m 1 m			RNN, LSTM, GRU	Precisão, Mensal Retorna	Difícil, Tensorflow
[87]	Dados de ações de S&P500	2011–2016	OCHLV	1 dia	1 dia	DBN	MSE, norma-RMSE, MUITO DE	–
[88]	Alta frequência dados de transação de os futuros CSI300	2017	Dados de preço	–	1 minuto	DNN, ELM, RBF	RMSE, MAPE, Precisão	Matlab
[89]	Ações no S&P500	1990–2015	Dados de preço	240 dias	1 dia	DNN, GBT, RF	Retorno médio, MDD, proporção calma	H2O
[90]	ACI Worldwide, grampos, e Seagate em NASDAQ	2006–2010	Fechamento diário preços	17 dias	1 dia	RN, ANN	RMSE	–
[91]	Ações chinesas	2007–2017	OCHLV	30 dias	1..5d	CNN + LSTM	Retorno Anualizado, Retração Mxm	Pitão
[92]	20 ações em S&P500	2010–2015	Dados de preço	–	–	AE + LSTM	Devoluções semanais	–
[93]	S&P500	1985–2006	Mensalmente e diariamente log-retornos	*	1 dia	DBN + MLP	Validação, Erro de Teste	Theano, Python, Matlab
[94]	12 ações da SSE Índice composto	2000–2017	OCHLV	60 dias	1..7d	DWNN	MSE	Tensorflow
[95]	50 ações de NYSE	2007–2016	Dados de preço	–	1d, 3d, 5 dias	SFM	MSE	–

Em outro grupo de estudos, alguns pesquisadores voltaram a focar em modelos baseados em LSTM. No entanto, seus parâmetros de entrada vieram de várias fontes, incluindo dados brutos de preços, dados técnicos e/ou análise fundamental, dados macroeconômicos, demonstrações financeiras, notícias e sentimento dos investidores. A Tabela 2 resume esses estoques documentos de previsão de preços. Na Tabela 2, diferentes métodos/modelos também são listados com base em cinco subgrupos: modelo DNN; LSTM e modelos RNN; modelos múltiplos e híbridos; modelo CNN; e romance métodos.

Modelos DNN foram usados em alguns artigos de previsão de preços de ações desse grupo. Em Abe et al. [96], um modelo DNN e 25 características fundamentais foram usadas para a previsão do Índice do Japão constituintes. Feng et al. [97] também usaram características fundamentais e um modelo DNN para predição. Um modelo DNN e macroeconomia dados, como PIB, taxa de desemprego e estoques, foram usado pelos autores de [98] para a previsão de US de baixo nível séries temporais macroeconômicas desagregadas.

Os modelos LSTM e RNN foram escolhidos em alguns estudos. Kraus e Feuerriegel [99] implementou LSTM com transferência de aprendizado usando mineração de texto através de notícias financeiras e dados do mercado de ações. Sim ilarmente, Minami et al. [100] usou LSTM para prever o dia seguinte do estoque preço usando eventos de ação corporativa e índice macroeconômico. Zhang e Tan [101] implementaram o DeepStockRanker, um modelo baseado em LSTM para classificação de ações usando 11 indicadores técnicos. Dentro Zhuge et al. [102], os autores utilizaram a série temporal de preços e dados emocionais de mensagens de texto para prever o preço de abertura das ações do dia seguinte com uma rede LSTM. Akita et al. [103] usado informações textuais e preços de ações por meio do Parágrafo Vector + LSTM para a previsão de preços e as comparações foram fornecidas

com classificadores diferentes. Ozbayoglu [104] usou indicadores técnicos junto com dados de ações em uma rede Jordan-Elman para preço predição.

Havia também modelos múltiplos e híbridos que utilizavam principalmente recursos de análise técnica como suas entradas para o modelo DL. Diversos indicadores técnicos foram alimentados em redes LSTM e MLP em Khare et al. [105] para previsão de preço intradiário. Recentemente, Zhou et al. [106] usou um GAN para minimizar a perda de erro de previsão e o modelo de perda de previsão de direção (GAN-FD) para previsão de preço de ações e compararam o desempenho de seus modelos com ARIMA, ANN e Máquina de vetores de suporte (SVM). Singh et al. [107] usou vários recursos de indicadores técnicos e dados de séries temporais com o Principal Análise de Componentes (PCA) para redução de dimensionalidade em cascata com um DNN (FFNN de 2 camadas) para previsão de preços de ações. Karaoglu et al. [108] usou indicadores comerciais baseados em microestrutura de mercado como entradas em um RNN com Graves LSTM detectando o buy-pressão de venda dos movimentos no Índice da Bolsa de Valores de Istambul (BIST) para realizar a previsão de preços para negociação de ações inteligente. Dentro Zhou et al. [109], previu-se o retorno do próximo mês e construíram-se carteiras a serem executadas. Bons retornos mensais foram alcançados com os modelos LSTM e LSTM-MLP.

Enquanto isso, em alguns jornais, os modelos da CNN eram os preferidos. Abroyan et al. [110] usou 250 recursos, incluindo detalhes do pedido, para a previsão de dados reais de uma corretora privada de transações arriscadas. Eles usaram CNN e LSTM para previsão de preço das ações. Os autores de [111] usaram um modelo CNN e fundamental, dados técnicos e de mercado para previsão.

Novos métodos também foram desenvolvidos em alguns estudos. Em Trans et al. [112], com o conjunto de dados FI-2010, bid/ask e volume foram

mesa 2

Previsão de preços de ações usando vários dados.								
Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte		Crítérios de desempenho	Env.
[96]	Índice do Japão constituintes de WorldScope	1990–2016	25 Fundamentos Recursos	10 dias	1 dia	DN	Correlação, precisão, MSE	Tensorflow
[97]	Devolução do S&P500 1926–2016		Fundamental Recursos:	–	1 segundo	DN	MSPE	Tensorflow
[98]	baixo nível dos EUA desagregado macroeconômico série temporal	1959–2008	PIB, Desemprego taxa, Estoques, etc.	–	–	DN	R2	–
[99]	Estoque CDAX dados de mercado	2010–2013	notícias financeiras, dados do mercado de ações	20 dias	1 dia	LSTM	MSE, RMSE, MAE, Precisão, AUC	TensorFlow, Theano, Python, Scikit-Learn
[100]	Estoque de Tsugami Corporação	2013	Dados de preço	–	–	LSTM	RMSE	Difícil, Tensorflow
[101]	Ações na China Uma partilha	2006–2007	11 técnico indicadores	–	1 dia	LSTM	AR, IR, IC	–
[102]	Preços SCI	2008–2015	OCHL de mudança taxa, preço	7 dias	–	Análise Emocional + LSTM	MSE	–
[103]	10 ações em Nikkei 225 e notícia	2001–2008	Textual informação e Preços das ações	10 dias	–	Vetor de parágrafo + LSTM	Lucro	–
[104]	Ações da TKC na NYSE e QQQQ ETF	1999–2006	Técnico indicadores, preço	50 dias	1 dia	RN (Jordan–Elman)	Lucro, MSE	Java
[105]	10 Ações na NYSE -		Dados de preços, Técnico indicadores	20 minutos	1 minuto	LSTM, MLP	RMSE	–
[106]	42 ações em SSE da China	2016	OCHLV, Técnico Indicadores	242 minutos	1 minuto	GAN (LSTM, CNN)	RMSRE, DPA, GAN-F, GAN-D	–
[107]	diário do Google dados de estoque	2004–2015	OCHLV, Técnico indicadores	20 dias	1 dia	(2D) ² PCA + DNN	SMAPE, PCD, MAPE, RMSE, HR, TR, R2	R, Matlab
[108]	GarantiBank em BIST, Turquia	2016	OCHLV, Volatilidade, etc.	–	–	PLR, Graves LSTM	MSE, RMSE, MAE, CSR, R2	Fagulha
[109]	Ações na NYSE, AMEX, NASDAQ, TAQ comércio intradiário	1993–2017	Preço, 15 empresa características	80 dias	1 dia	LSTM + MLP	Retorno mensal, SR	Python, Difícil, Tensorflow em AWS
[110]	Corretagem privada real da empresa dados de risco transações	–	250 recursos: detalhes do pedido, etc.	–	–	CNN, LSTM	F1-Pontuação	Difícil, Tensorflow
[111]	Fundamentos e Dados técnicos, Dados econômicos	–	Fundamental técnico e mercado em formação	–	–	CNN	–	–
[112]	O LOB de 5 ações da finlandesa Mercado de ações	2010	Conjunto de dados FI-2010: oferta/pedido e volume	–	*	WMTR, MDA	Acurácia - Precisão, Recall, F1-Score	–
[113]	Retorna na NYSE, AMEX, NASDAQ	1975–2017	57 empresa características	*	–	Fama–Francês modelo de fator n DL	R ² , RMSE	Tensorflow

usado como o conjunto de recursos para previsão. No estudo, eles propuseram a Regressão de Séries Temporais Multicanal Ponderada (WMTR), e Análise Discriminante Multilinear (MDA). Feng et al. [113] usou 57 características, incluindo Market equity, Market Beta, Industry momentum e Asset growth, como entradas para a Fama-French n-factor DL para prever o patrimônio mensal dos EUA retornos na Bolsa de Valores de Nova York (NYSE), American Stock Exchange (AMEX) ou NASDAQ.

Vários trabalhos de pesquisa também usaram técnicas de mineração de texto para extração de recursos, mas usaram modelos não LSTM para estoque previsão de preços. A Tabela 3 resume a previsão do preço das ações artigos que usaram técnicas de mineração de texto. Na Tabela 3, diferentes métodos/modelos são agrupados em três subgrupos: CNN e modelos LSTM; modelos GRU, LSTM e RNN; e novos métodos.

Os modelos CNN e LSTM foram adaptados em alguns dos artigos. Em Ding et al. [114], eventos foram detectados pela Reuters e notícias da Bloomberg por meio de mineração de texto, e essa informação foi usado para previsão de preços e negociação de ações através da CNN modelo. Vargas e col. [115] usou mineração de texto no índice S&P500 notícias da Reuters através de um modelo híbrido LSTM+CNN para previsão de preço e estimativa de movimento direcional intradiário juntos. Lee et al. [116] usaram dados de notícias financeiras e implementaram a incorporação de palavras com Word2vec junto com MA e

oscilador estocástico para criar entradas para uma CNN recorrente (RCNN) para previsão do preço das ações. Iwasaki et al. [117] também usou sentimento análises por meio de mineração de texto e incorporação de palavras do analista relatórios e recursos de sentimento usados como entradas para um modelo DFNN para previsão do preço das ações. Em seguida, diferentes seleções de portfólio foram implementados com base nos retornos de ações projetados.

Os modelos GRU, LSTM e RNN foram os preferidos no próximo grupo de papéis. Das et al. [118] implementou a análise de sentimento em Publicações no Twitter junto com dados de ações para previsão de preços usando um RN. Da mesma forma, os autores de [119] usaram classificação de sentimento (neutro, positivo e negativo) para abertura ou fechamento de estoque previsão de preços com vários modelos LSTM. Eles compararam seus resultados com o SVM e alcançou um desempenho geral mais alto. Dentro Zhongshengz et al. [120], dados de texto e preço foram usados para o previsão dos preços do SSE Composite Index (SCI).

Novas abordagens foram relatadas em alguns artigos. Nascimento et al. [121] usaram incorporações de palavras para extrair informações de páginas da web e, em seguida, combinou com dados de preços de ações para previsão do preço das ações. Eles compararam a Autoregressiva (AR) modelo e RF com e sem novidades. Os resultados mostraram a incorporação de informações de notícias melhorou o desempenho. Han et al. [122] usou notícias financeiras e o corpus chinês ACE2005. Diferentes tipos de eventos de empresas chinesas foram classificados

baseado em um novo algoritmo de classificação de padrão de tipo de evento em Han et al. [122], e também a mudança no preço das ações no dia seguinte também foi prevista usando entradas adicionais.

4.2. Previsão de índice

Em vez de tentar prever o preço de uma única ação, vários pesquisadores preferiram prever o índice do mercado de ações. Os índices geralmente são menos voláteis do que as ações individuais porque são compostos de várias ações de diferentes setores e são mais indicativos do momento geral e do estado geral da economia.

Na literatura, diferentes dados de índices do mercado de ações têm sido utilizados para experimentos. Os dados de índice mais usados são os seguintes: S&P500, China Securities Index (CSI)300, National Stock Exchange of India (NIFTY), Tokyo Nikkei Index (NIKKEI)225, Dow Jones Industrial Average (DJIA), Shanghai Stock Exchange (SSE) 180, Hong Kong Hang Seng Index (HSI), Shenzhen Stock Exchange Composite Index (SZSE), London Financial Times Stock Exchange Index (FTSE)100, Taiwan Capitalization Weighted Stock Index (TAIEX), BIST, NASDAQ, Dow Jones Industrial Average 30 (DOW30), KOSPI, S&P500 Volatility Index (VIX), NASDAQ100 Volatility Index (VIXN), Bolsa de Valores Brasileira (Bovespa), Bolsa de Valores de Estocolmo (OMX) e NYSE. Os artigos [123–134,114] usaram o S&P500 como seu conjunto de dados.

Os autores dos estudos [123,124,135–137] usaram o NIKKEI como seu conjunto de dados. KOSPI foi usado em Li et al. [135], Jeong et al. [131], Baek et al. [132]. DJIA foi usado como o conjunto de dados nos artigos [123,136–139]. Os autores dos artigos [123,135,137,131] usaram HSI como o conjunto de dados em seus estudos, e SZSE foi usado nos estudos [140, 135,141,142].

Além disso, na literatura, existem diferentes métodos para a previsão de dados de índice. Enquanto alguns estudos usaram apenas dados brutos de séries temporais, outros usaram vários outros dados, como indicadores técnicos, dados de índice, feeds de mídia social, notícias da Reuters e Bloomberg e características estatísticas dos dados (desvio padrão, assimetria, curtose, proporção ômega, fundo alfa). Nesta pesquisa, primeiro agrupamos os artigos de previsão de índice de acordo com seus conjuntos de recursos, como estudos usando apenas dados brutos de séries temporais (dados de preço/índice, OCHLV); em seguida, agrupamos os estudos usando vários outros dados. A [Tabela 4](#) resume os papéis de previsão do índice usando apenas dados brutos de séries temporais. Além disso, diferentes métodos (modelos) foram usados para a previsão do índice. Os métodos MLP, RNN, LSTM e DNN (DFNN ou DMLP) foram os mais utilizados para a previsão do índice. Na [Tabela 4](#), esses vários métodos/modelos também são listados como quatro subgrupos: modelos ANN, DNN, MLP e Fuzzy Deep Direct Reinforcement Learning (FDDR); modelos RL e DL; modelos LSTM e RNN; e novos métodos.

Os modelos ANN, DNN, MLP e FDDR foram usados em alguns estudos. Em Lachiheb et al. [143], os retornos de log dos dados do índice foram usados com um DNN com entrada hierárquica para a previsão dos dados TUNIN DEX. Yong et al. [144] usaram um FFNN profundo e Open, Close, High, Low (OCHL) dos últimos 10 dias de dados de índice para previsão. Além disso, MLP e ANN foram usados para a previsão de dados de índice. Em Yumlu et al. [145], dados brutos de índice foram usados com MLP, RNN, Mixture of Experts (MoE) e Exponential GARCH (EGARCH) para previsão. Em Yang et al. [142], ensembles de ANN com OCHLV de dados foram usados para predição do índice composto de Shanghai.

Além disso, os métodos RL e DL foram usados em conjunto para predição de dados de índice em alguns estudos. Em Deng et al. [141], os métodos FDDR, DNN e RL foram usados para prever 300 ações a partir de dados do índice SZSE e preços de commodities. Em Jeong et al. [131], os métodos Deep Q Learning e DNN e um conjunto de dados de preços de ações de 200 dias foram usados juntos para previsão do índice S&P500.

A maioria dos métodos preferidos para previsão de dados de índice usando dados brutos de séries temporais foram baseados em LSTM e RNN.

Em Bekiros et al. [139], um RNN foi usado para previsão de retornos de log do índice DJIA. Em Fischer et al. [125], o LSTM foi usado para prever os dados do índice S&P500. Althelaya et al. [128] usaram métodos LSTM empilhados e LSTM bidirecional (Bi-LSTM) para S&P500 na previsão de preço. Yan et al. [146] usou uma rede LSTM para prever o preço de fechamento do dia seguinte do índice de ações de Xangai. Em seu estudo, eles usaram a decomposição wavelet para reconstruir a série temporal financeira para eliminação de ruído e melhor aprendizado. Em Pang et al. [140], o LSTM foi usado para a previsão do índice composto de ações A de Xangai. Namini et al. [136] usaram LSTM para prever dados de índice NIKKEI225, IXIC, HIS, GSPC e DJIA. Em Takahashi et al. [147] e Baek et al. [132], o LSTM também foi utilizado para a previsão dos índices S&P500 e KOSPI200. Baek et al. [132] desenvolveram um modelo de previsão de índice de ações baseado em LSTM chamado ModAugNet.

O método proposto foi capaz de vencer Buy and Hold (B&H) no longo prazo com um mecanismo de prevenção de overfitting. Elliot et al. [134] compararam diferentes modelos de ML (modelos lineares), Modelos Lineares Generalizados (GMLs) e vários modelos LSTM e RNN para previsão de preços de índices de ações. Em Hansson et al. [133], o LSTM e a parte autorregressiva dos dados do índice de séries temporais foram usados para predição dos índices S&P500, Bovespa50, OMX30.

Além disso, alguns estudos adaptaram novas abordagens. Em Zhang et al. [138], um DNN genético foi usado para a previsão do índice DJIA. Borovykh et al. [127] propuseram um novo modelo DNN chamado rede convolucional Wavenet para previsão de séries temporais. Bildirici et al. [148] propuseram um modelo de correção de erro vetorial (VEC)-modelo de Elman híbrido recorrente (RHE) Threshold Autoregressive (TAR) para forex e índice de ações de previsão de retorno e compararam vários modelos. Parida et al. [124] propuseram um método chamado Locally Recurrent Neuro-fuzzy Information System (LRNFIS) com Firefly Harmony Search Optimization (FHSO) Evolutionary Algorithm (EA) para prever os dados de preços S&P500 e NIKKEI225 e USD Exchange. Psaradellis et al. [149] propuseram Processo Autoregressivo Heterogêneo (HAR) com GA com um modelo SVR (GASVR) chamado HAR-GASVR para predição dos índices VIX, VIXN, Dow Jones Industrial Average Volatility Index (VIXD).

Na literatura, alguns estudos usaram vários dados de entrada, como indicadores técnicos, dados de índices, notícias de mídia social, notícias da Reuters e Bloomberg, e características estatísticas dos dados (desvio padrão, assimetria, curtose, índice ômega, alfa do fundo). A [Tabela 5](#) resume os documentos de previsão do índice usando esses vários dados mencionados. Os métodos DNN, RNN, LSTM e CNN foram os modelos mais usados na previsão de índices. Na [Tabela 5](#), diferentes métodos/modelos também são listados em quatro subgrupos: modelo DNN; modelos RNN e LSTM; modelo CNN; e novos métodos.

Um DNN foi usado como modelo de classificação em alguns artigos. Em Chen et al. [150], um DNN e algumas características dos dados (Retorno, Razão Sharpe (SR), Desvio Padrão (STD), Distorção, Curtose, Razão Omega, Alfa de Fundo) foram usados para predição. Em Widegren et al. [126], DNN, RNN e indicadores técnicos foram usados para previsão dos índices FTSE100, OMX30, S&P500.

Além disso, modelos RNN e LSTM com diversos outros dados também foram utilizados para previsão dos índices. Hsieh et al. [137] usaram RNN e OCHLV de índices e indicadores técnicos para prever os índices DJIA, FTSE, Nikkei e TAIEX. Mourelatos et al. [151] usaram GASVR e LSTM para previsão. Chen et al. [152] utilizaram quatro modelos LSTM (análise técnica, mecanismo de atenção e incorporação de vetores de mercado) para previsão do índice de retorno diário do índice HSI300. Em Li et al. [135], LSTM com redução de ruído wavelet e dados de índice, volume e indicadores técnicos foram usados para a previsão dos índices HSI, SSE, SZSE, TAIEX, NIKKEI e KOSPI. Si et al. [153] usaram um método MODRL+LSTM para prever índices de contratos de ações chinesas-IF-IH-IC. Bao et al. [123] usaram EAs empilhados para gerar recursos profundos usando OCHL de preços de ações, indicadores técnicos e condições macroeconômicas para alimentar o LSTM para prever preços futuros de ações.

Tabela 3
Previsão de preços de ações usando técnicas de mineração de texto para extração de recursos.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método	Horizonte	Critérios de desempenho	Env.
[114]	Índice S&P500, 15 ações no S&P500	2006–2013	Notícias de Reuters e Bloomberg	–	–	CNN	Precisão, MCC	–
[115]	Índice S&P500 notícias da Reuters	2006–2013	Notícias financeiras títulos, técnico indicadores	1 dia	1 dia	RCNN	Precisão	–
[116]	Índice TWSE, 4 ações em TWSE	2001–2017	Técnico indicadores, preço dados, notícias	15 dias	–	CNN + LSTM	RMSE, Lucro	Keras, Python, TALIB
[117]	Relatórios de analistas sobre o TSE e Osaka	2016–2018	Texto	–	–	LSTM, CNN, Bi-LSTM	Precisão, R-quadrado	R, Python, MeCab
[118]	Intercâmbio Ações do Google, Microsoft e	2016–2017	Sentimento do Twitter e preços das ações	–	–	RN	–	Faisca, Flume, API do Twitter,
[119]	Maça Ações de CSI300 índice, OCHLV de Índice CSI300	2009–2014	Postagens de sentimento, Dados de preço	1 dia	1 dia	Naive Bayes + LSTM	Precisão, Recordação, Pontuação F1, Precisão	Python, alto
[120]	Preços SCI	2013–2016	Dados de texto e Dados de preço	7 dias	1 dia	LSTM	Precisão, F1-Measure Python, Keras	
[121]	Ações de S&P500	2006–2013	Texto (notícias) e Dados de preço	7 dias	1 dia	LAR+News, RF+Notícias	MAPE, RMSE	–
[122]	Notícias de Sina.com, ACE2005 corpus chinês	2012–2016	Um conjunto de texto de notícias –	–	–	Sua única algoritmo	Precisão, Recordação, Pontuação F1	–

Além disso, diferentes implementações de CNN com vários dados (indicadores técnicos, notícias e dados de índice) têm sido usados em a literatura. Em Dingli et ai. [129], CNN e dados de índice, e indicadores técnicos foram usados para os índices S&P500, DOW30, NAS DAQ100 e preços de commodities, Forex e Bitcoin. Dentro Ding et ai. [114], um modelo da CNN com notícias da Reuters e Bloomberg foram usados para previsão do índice S&P500 e Preços de 15 ações no S&P500. Em Lee et ai. [116], CNN + LSTM e indicadores técnicos, dados de índice e notícias foram usados para previsão do índice da Bolsa de Valores de Taiwan (TWSE) e 4 preços das ações em TWSE.

Além disso, alguns novos métodos têm sido propostos para previsão. Rout et ai. [130] modelos RNN usados, Rede Neural de Link Funcional Computacionalmente Eficiente Recorrente (RCEFLANN), e Rede Neural de Link Funcional (FLANN), com seus pesos otimizado usando vários EAs como Particle Swarm Optimization (PSO) e Versão modificada do PSO (HMRPSO), para séries temporais previsão. Chen et ai. [154] usou notícias de mídia social para prever preço do índice e direção do índice com RNN-Boost com Latent Recursos de alocação de Dirichlet (LDA).

4.3. Previsão de preços de commodities

Vários estudos se concentraram particularmente na previsão de preços de qualquer commodity, como ouro, prata, petróleo e cobre. Com o aumento do número de commodities disponíveis para o público negociação através de bolsas de valores online, o interesse neste tópico será provavelmente crescerá nos anos seguintes.

Na literatura, existem diferentes métodos utilizados para previsão de preços de commodities. DNN, RNN, FDDR e CNN são os modelos mais usados para prever preços de commodities. A Tabela 6 lista os detalhes dos estudos de previsão de preços de commodities com DL.

Em Dingli et ai. [129], os autores usaram uma CNN para prever o movimento direcional de preços da próxima semana e do próximo mês. Enquanto isso, os modelos RNN e LSTM foram usados em alguns estudos de previsão de commodities. Em Dixon et ai. [155], um DNN foi usado para previsão de commodities. Em Widegren et al. [126], forex e índice conjuntos de dados foram usados. DNN e RNN foram usados para prever os preços de dados de séries temporais. Indicadores técnicos foram usados como recurso conjunto composto por Índice de Força Relativa (RSI), Porcentagem de Williams Faixa (William%R), Índice de Canal de Commodities (CCI), Porcentagem

Price Oscillator (PPOSC), momentum e movimento exponencial Média (EMA). Em Lasheras et ai. [156], os autores usaram um Elman RNN para prever o preço spot do cobre COMEX (através de Nova York Mercantile Exchange (NYMEX)) dos preços de fechamento diários.

Modelos híbridos e novos foram adaptados em alguns estudos. Em Zhao et ai. [157], FNN e Autoencoders Stacked Denoising (SDAE) modelos profundos foram comparados com o Support Vector Regressor (SVR), Random Walk (RW) e Markov Regime Switching (MRS) para a previsão do preço do petróleo WTI. Como critérios de desempenho, precisão, erro percentual médio absoluto (MAPE) e Foi utilizado o Erro Quadrado Médio (RMSE). Em Chen e col. [158], o autores tiveram como objetivo prever os preços do petróleo bruto WTI usando vários modelos, incluindo combinações de DBN, LSTM, Autoregressive Moving Média (ARMA) e RW. MSE foi usado como o desempenho critério. Em Deng et ai. [141], os autores usaram FDDR para estoque previsão de preços e geração de sinais de negociação. Eles combinaram DN e RL. As curvas de lucro, retorno, SR e lucro-perda foram usadas como critério de desempenho.

4.4. Previsão de volatilidade

A volatilidade está diretamente relacionada às variações de preços em um determinado momento período e é usado principalmente para avaliação de risco e precificação de ativos. Alguns pesquisadores implementaram modelos para prever com precisão a volatilidade subjacente de qualquer ativo.

Na literatura, existem diferentes métodos utilizados para previsão de volatilidade, incluindo LSTM, RNN, CNN, MM e Heterocedasticidade condicional auto-regressiva generalizada (GARCH) modelos. A Tabela 7 resume os estudos que focaram na volatilidade previsão. Na Tabela 7, diferentes métodos/modelos também estão representados como três subgrupos: CNN; modelos RNN e LSTM; e modelos híbridos e novos.

Um modelo CNN foi usado em um estudo de previsão de volatilidade com base em dados HFT [159]. Enquanto isso, os modelos RNN e LSTM foram usado em alguns estudos. Em Tino et ai. [160], os autores usaram dados de séries temporais financeiras para prever mudanças de volatilidade com Markov Models e Elman RNN para negociação lucrativa de opções de straddle. Xiong et ai. [161] usou dados de preços e diferentes tipos de Google Tendências domésticas com LSTM. Zhou et ai. [162] usaram CSI300 e 28 palavras do volume diário de pesquisa com base no Baidu como conjunto de dados

Tabela 4
Previsão de índice usando apenas dados brutos de séries temporais.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte		Critérios de desempenho	Env.
[124]	S&P500, Nikkei225, USD Trocas	2011–2015	Dados do índice	–	1d, 5d, 7d, 10d	LRNFIS com Firefly-Harmony Procurar	RMSE, MAPE, MAE	–
[125]	Índice S&P500	1989–2005	Dados de índice, Volume	240 dias	1 dia	LSTM	Retorno, STD, SR, Precisão	Pitão TensorFlow, Duro, R, H2O
[127]	S&P500, VIX	2005–2016	Dados do índice	*	1 d uWN, cWN 1 d, 30 d LSTM		MASE, HIT, RMSE	–
[128]	Índice S&P500	2010–2017	Dados do índice	10 dias	empilhado, Bi-LSTM		MAIO, RMSE, R ao quadrado	Python, Dificil, Tensorflow
[131]	S&P500, KOSPI, HSI, e EuroStoxx50	1987–2017	estoque de 200 dias preço	200 dias	1 dia	Q-Learning Profundo e DN	Lucro total, Correlação	–
[132]	S&P500, KOSPI200, 10 ações	2000–2017	Dados do índice	20 dias	1 dia	ModAugNet: LSTM MSE, MAPE, MAE		Dificil
[133]	S&P500, Bovespa50, OMX30	2009–2017	Autoregressivo parte do tempo Series	–	1 dia	LSTM	MSE, Precisão	Tensorflow, Dificil, R
[134]	S&P500	2000–2017	Dados do índice	–	1..4 d, 1 em, 1,3 m	GLM, LSTM+RNN MAE, RMSE		Pitão
[136]	Nikkei225, IXIC, HSI, GSPC, DJIA	1985–2018	OCHLV	5 dias	1 dia	LSTM	RMSE	Python, Dificil, Theano
[138]	DJIA	–	Dados do índice	–	–	Genética Profunda Rede neural	MSE	Java
[139]	Registre os retornos do DJIA	1971–2002	Dados do índice	20 dias	1 dia	RN	TR, taxa de sinal, PT/HM teste, MSFE, SR, lucro	–
[140]	Ações de Xangai índice composto, SZSE	2006–2016	OCHLV	10 dias	–	Camada incorporada + LSTM	Precisão, MSE	Python, Matlab, Theano
[141]	300 ações de SZSE, Commodity	2014–2015	Dados do índice	–	–	FDDR, DNN + RL Lucro, retorno, SR,	curvas de lucro-perda	Dificil
[142]	Xangai índice composto e SZSE	1990–2016	OCHLV	20 dias	1 dia	Conjuntos de Precisão de ANN		–
[143]	TUNINDEX	2013–2017	Registrar retornos de dados de índice	–	5 minutos	DN com entrada hierárquica	Precisão, MSE	Java
[144]	Ações de Cingapura Índice de mercado	2010–2017	OCHL dos últimos 10 dias de índice	10 dias	3 dias	Feed-forward DNN RMSE, MAPE, Lucro, SR		–
[145]	ÉS TU	1990–2002	Dados do índice	7 dias	1 dia	MLP, RNN, MoE	HIT, positivo/negativo HIT, MSE, MAE	–
[146]	SCI	2012–2017	OCHLV, dados do índice -		1..10 dias	Wavelet + LSTM MAPE, eles são desiguais coeficiente		–
[147]	S&P500 [148] ISE100	1950–2016 1987–2008	Dados do índice Dados do índice	15 dias –	1 dia 2d, 4d, 8d, 12d, 18 dias	LSTM TAR-VEC-MLP TAR-VEC-RBF, TAR-VEC-RHE	RMSE RMSE	Dificil –
[149]	VIX, VXN, VXD	2002–2014	Primeiros cinco atrasos autorregressivos	5 dias	1d, 22d	HAR-GASVR	SIM, RMSE	–

com LSTM para prever a volatilidade do índice. Kim et ai. [163] desenvolveu vários modelos LSTM integrados com GARCH para volatilidade predição.

Abordagens híbridas e novas também foram adaptadas em alguns estudos. Em Nikolaev et ai. [164], foi proposto um modelo RMDN com GARCH (RMDN GARCH). Além disso, vários modelos, incluindo modelos tradicionais de previsão e modelos DL, foram comparados para estimativa de volatilidade. Psaradellis et ai. [149] propôs um romance método chamado HAR com GASVR (HAR-GASVR) para índice de volatilidade previsão.

4.5. Previsão de preço de títulos

Alguns especialistas financeiros seguem as mudanças nos preços dos títulos para analisar o estado da economia, alegando que os preços dos títulos representam a saúde da economia melhor do que o mercado de ações [165]. Historicamente, as taxas de longo prazo são mais altas do que as taxas de curto prazo em expansão econômica normal, enquanto imediatamente antes recessões, as taxas de curto prazo passam as taxas de longo prazo, ou seja, uma curva de rendimento. Portanto, a previsão precisa do preço dos títulos é muito útil. No entanto, as implementações de DL para previsão de preços de títulos são muito escasso. Em Bianchi et ai. [166], o excesso de retorno de títulos foi previsto

usando vários modelos de ML, incluindo redes RF, AE e PCA e um DFNN de 2-3-4 camadas. O NN de 4 camadas superou os outros modelos.

4.6. Previsão de preços forex

Os mercados de câmbio têm os maiores volumes entre todos os mercados financeiros existentes no mundo. Eles estão abertos 24/7, e trilhões de dólares em transações de câmbio acontecer em um único dia. De acordo com o Banco Internacional Liquidações, negociação de câmbio tem um volume de mais de 5 trilhões de dólares por dia [167]. Além disso, há um grande número de plataformas de negociação forex online que oferecem oportunidades de transações alavancadas para seus assinantes. Como resultado, há grande interesse em estratégias de negociação lucrativas por comerciantes. Conseqüentemente, há uma série de estudos de previsão e negociação de forex baseados nos modelos DL. Porque a maioria das transações financeiras globais são baseados no dólar americano, quase todas as pesquisas de previsão forex artigos incluem USD em suas análises. No entanto, dependendo diferenças regionais e foco de pesquisa pretendido, vários modelos foram desenvolvidos em conformidade.

Na literatura, diferentes métodos têm sido usados para forex previsão de preços, incluindo RNN, LSTM, CNN, DBN, DNN, AE e

Tabela 5
Previsão de índice usando vários dados.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Crítérios de desempenho	Env.
[114]	Índice S&P500, 15 ações no S&P500	2006–2013	Notícias de Reuters e Bloomberg	–	– CNN	Precisão, MCC	–
[116]	Índice TWSE, 4 ações em TWSE	2001–2017	Técnico indicadores, índice dados, notícias OCHLV, Técnico Indicadores	15 dias	– CNN + LSTM	RMSE, Lucro	Keras, Python, TALIB
[123]	CSI300, NIFTY50, HSI, NIKKEI225, S&P500, DJIA	2010–2016	OCHLV, Técnico Indicadores	–	1 dia WT, Empilhado codificadores automáticos, LSTM	MAPE, Correlação coeficiente, THEIL-U	–
[126]	FTSE100, OMXS 30, SP500, Mercadoria, Forex	1993–2017	Técnico indicadores	60 dias	1 dia DN, RN	Precisão, valor-p	–
[129]	S&P500, DOW30, NASDAQ100, Mercadoria, Forex, Bitcoin	2003–2016	Dados de índice, Técnico indicadores	–	1s, 1m CNN	Precisão	Tensorflow
[130]	BSE, S&P500	2004–2012	Dados de índice, técnico indicadores	5 dias	1 d..1 m PSO, HMRPSO, DE, RCEFLANN	RMSE, MAPE	–
[135]	HSI, SSE, SZSE, TAIEX, NIKKEI, KOSPI	2010–2016	Dados de índice, volume, técnico indicadores	2d..512d 1d	LSTM com redução de ruído wavelet	Precisão, MAPE	–
[137]	DJIA, FTSE, NIKKEI, TAIEX	1997–2008	OCHLV, Técnico indicadores	26 dias	1 dia RN	RMSE, MAE, MAPE, THEIL-U	C
[150]	Fundo de hedge retorno mensal dados	1996–2015	Retorno, SR, STD, Distorção, Curtose, Ômega proporção, fundo alfa	12m 3m,	6m, 12 m DN	Índice de Sharpe, anual volte, Cum. Retorna	–
[151]	Estoque Nacional Banco da Grécia (REFEÇÃO)	2009–2014	FTSE100, DJIA, GDAX, NIKKEI225, EUR/USD, Ouro	1d, 2d, 5d, 10d	1 dia GASVR, LSTM	Retorno, volatilidade, SR, Precisão	Tensorflow
[152]	Taxa de retorno diária do índice HS300	2004–2018	OCHLV, Técnico indicadores	–	– Vetor de mercado + Tecnologia ind. + LSTM + Atenção	MSE, MAE	Pitão Tensorflow
[153]	chinês estoque-IF-IH-IC contrato	2016–2017	Decisões para mudança de índice	240 minutos 1 minuto	MODRL+LSTM	Lucros e perdas, SR	–
[154]	HS300	2015–2017	Mídia social notícias, dados de índice	1 dia	1 dia RNN-Reforço com LDA	Precisão, MAE, MAPE, RMSE	Pitão Scikit-learn

Tabela 6
Previsão de preços de commodities.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Crítérios de desempenho	Env.
[129]	S&P500, DOW30, NASDAQ100, Mercadoria, Forex, Bitcoin	2003–2016	Dados de preços, Técnico indicadores	–	1s, 1m CNN	Precisão	Tensorflow
[155]	Mercadoria, FX futuro, ETF	1991–2014	Dados de preço	100 * 5 minutos 5 minutos	DN	SR, relação de capacidade, Retorna	C++, Python
[126]	FTSE100, OMX30, S&P500, Mercadoria, Forex	1993–2017	Técnico indicadores	60 dias	1 dia DN, RN	Precisão, valor-p	–
[156]	Preços do cobre da NYMEX	2002–2014	Dados de preço	–	– Elman RN	RMSE	R
[157]	Petróleo bruto WTI preço	1986–2016	Dados de preço	1m	1m SDAE, Bootstrap agregação	Precisão, MAPE, RMSE	Matlab
[158]	Petróleo Bruto WTI Preços	2007–2017	Dados de preço	–	– ARMA + DBN, RW + LSTM	MSE	Python, Difícil, Tensorflow
[141]	300 ações de SZSE, Commodity	2014–2015	Dados de preço	–	– FDDR, DNN + RL Lucro, retorno, SR, curvas de lucro-perda		Difícil

Métodos MLP. A Tabela 8 fornece detalhes sobre essas implementações. Na Tabela 8, diferentes métodos/modelos são listados como quatro subgrupos: Redes de Crenças Profundas de Valor Contínuo (CDBN), modelos DBN, DBN+RBM e AE; Rede DNN, RNN, Psi-Sigma (PSN) e modelos LSTM; modelos CNN; e modelos híbridos. Os modelos CDBN, DBN, DBN+RBM e AE têm sido usados em alguns estudos. Em [168], a granulação de informações fuzzy integrada com CDBN foi aplicado para prever EUR/USD e GBU/USD taxas de câmbio. Eles estenderam um DBN com uma máquina Boltzman restrita contínua (CRBM) para melhorar o desempenho. Dentro Chao et ai. [169], os preços semanais GBP/USD e INR/USD foram

previsto, enquanto em Zheng et al. [170], CNY/USD e INR/USD foram o foco principal. Em ambos os casos, o DBN foi comparado com FFnn. Da mesma forma, Shen et al. [171] implementou vários Redes DBN para prever semanalmente GBP/USD, BRL/USD e INR/USD retornos da taxa de câmbio. Shen et ai. [172] combinado AE empilhado e SVR para prever 28 pares de moedas normalizados usando o tempo dados de série de USD, GBP, EUR, JPY, AUD, CAD e CHF. Os modelos DNN, RNN, PSN e LSTM foram preferidos em alguns estudos. Em Dixon et ai. [155], vários modelos DMLP foram desenvolvidos para prever futuros de AD e BP usando dados de 5 minutos em um período de 130 dias. Serpinis et ai. [173] usaram MLP, RNN, GP e

Tabela 7

Previsão de volatilidade.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Críticos de desempenho	Env.	
[159]	Bolsa de Londres Intercâmbio	2007–2008	Limite de livro de pedidos estado, comércio, ordens de compra/venda, exclusões de pedidos	–	–	CNN	Precisão, kapa	Café
[160]	DAX, FTSE100, opções de compra/venda	1991–1998	Dados de preço	*	*	MM, RN	Ewa-medida, iv, média dos lucros diários e padrão	–
[161]	S&P500	2004–2015	Dados de preço, 25 Google doméstico dimensões de tendência	–	1 dia	LSTM	MAPE, RMSE	–
[162]	CSI 300, 28 palavras da busca diária volume baseado em Baidu	2006–2017	Dados de preço e texto 5 d		5 dias	LSTM	MSE, MAPA	Python, alto
[163]	KOSPI200, Coreia Título do Tesouro taxa de juro, grau AA vínculo corporativo taxa de juros, ouro, óleo cru	2001–2011	Dados de preço	22 dias	1 dia	LSTM + GARCH	MAE, MSE, HMAE, HMSE	–
[164]	DEM/GBP taxa de câmbio	–	Devoluções	–	–	RMDN-GARCH	NMSE, NMAE, RH, WHR	–
[149]	VIX, VXN, VXD	2002–2014	Primeiros cinco atrasos autorregressivos	5 dias	1d, 22d	HAR-GASVR	SIM, RMSE	–

outras técnicas de ML junto com métodos tradicionais de regressão para também prever séries temporais EUR/USD. Eles também integraram Filtro de Kalman, operador LASSO e outros modelos para melhorar ainda mais os resultados [174]. Eles estenderam ainda mais suas análises incluindo PSN e fornecendo comparações junto com os tradicionais meteorologistas ARIMA, RW e STAR [175]. Para melhorar o desempenho, eles também integraram alavancagem híbrida de volatilidade variável no tempo. Sol et al. [176] implementou a previsão da taxa de câmbio RMB contra JPY, HKB, EUR e USD comparando os desempenhos de RW, RNN e FFNN. Maknickiene et al. [177] previu vários horários forex séries e criaram carteiras compostas por esses investimentos. Cada rede usou LSTM (RNN EVOLINO), e diferentes apetites de risco para usuários foram testados. Maknickiene et al. [178] também usado EVOLINO RNN + dados de entrada ortogonais para prever USD/JPY e preços XAU/USD em diferentes períodos.

Diferentes modelos CNN foram usados em alguns estudos. Na Pérsio et al. [179], EUR/USD foi novamente previsto usando múltiplos Modelos DL, incluindo MLP, CNN, RNN e Wavelet+CNN. Korczak et al. [180] implementou negociação forex (GBP/PLN) usando vários diferentes parâmetros de entrada em um ambiente de negociação baseado em vários agentes. Um dos agentes usou AE+CNN como predição modelo e superou todos os outros modelos.

Modelos híbridos também foram adaptados em algumas das pesquisas. Bildirici et al. [148] desenvolveu vários (TAR-VEC-RHE) modelos para prever retornos mensais para TRY/USD e desempenhos de modelos comparados. Nikolaev et al. [164] compararam vários modelos, incluindo modelos de previsão tradicionais e DL modelos, para previsão de DEM/GBP. Parida et al. [124] previsto AUD, CHF, MAX e BRL em relação aos dados de séries temporais da moeda USD usando o LRFIS e comparou-o com diferentes modelos. Enquanto isso, em vez de usar a minimização de erros baseada em LMS durante o aprendizado, eles usaram FHSO.

4.7. Previsão de preços de criptomoedas

Desde que as criptomoedas se tornaram um tema quente nas finanças indústria nos últimos anos, muitos estudos e implementações têm sido conduzido. A maioria dos estudos de criptomoeda se concentrou em previsão de preços.

O aumento do Bitcoin de 1000 USD em janeiro de 2017 para 20.000 USD em janeiro de 2018 atraiu muita atenção, não só de

do setor financeiro, mas também do público em geral. Recentemente, foram publicados artigos sobre previsão de preços e desenvolvimento de estratégias de negociação para Bitcoin e outras criptomoedas. Dado a atenção que a tecnologia subjacente atraiu, há uma forte chance de que novos estudos apareçam em um futuro próximo.

Na literatura, DNN, LSTM, GRU, RNN e métodos clássicos (ARMA, ARIMA, Autoregressive Conditional Heteroscedasticity (ARCH), GARCH, etc.)

previsão. A Tabela 9 resume os estudos que utilizaram esses métodos. Lopes [181] combinou a opinião de mercado e preço previsão para negociação de criptomoedas. Mineração de texto combinada com 2 modelos, CNN e LSTM, foram usados para extrair opinião. Bitcoin, Litecoin e StockTwits foram usados como o conjunto de dados. OCHLV de preços, indicadores técnicos e análise de sentimento foram usados como o conjunto de recursos. McNally et al. [182] compararam RNN, LSTM e ARIMA otimizados Bayesian para prever a direção do preço do Bitcoin. Sensibilidade, especificidade, precisão, exatidão e RMSE foram usados como as métricas de desempenho.

4.8. Previsão de tendências

Embora a previsão de tendências e a previsão de preços compartilhem mesmas características de entrada, alguns pesquisadores preferem prever o direção do preço de um ativo em vez de seu preço real. Isso altera a natureza do problema da regressão à classificação, e as métricas de desempenho correspondentes também mudam. no entanto vale a pena notar que essas duas abordagens ainda são fundamentalmente o mesmo; a diferença está na interpretação da saída.

Na literatura, existem diferentes métodos para previsão de tendências. Nesta pesquisa, agrupamos os artigos de acordo com sua conjuntos de recursos, como estudos usando apenas dados brutos de séries temporais (somente dados de preços, OCHLV); estudos usando indicadores técnicos, preço dados e dados fundamentais ao mesmo tempo; estudos usando texto técnicas de mineração; e estudos usando vários outros dados. Tabela 10 resume os estudos de previsão de tendências usando apenas o tempo bruto dados da série.

Diferentes métodos e modelos têm sido usados para previsão de tendências.

Na Tabela 10, eles estão divididos em três subgrupos: ANN, modelos DNN e FFNN; Modelos LSTM, RNN e NN Probabilístico; e novos métodos. Métodos ANN, DNN, DFNN e FFNN foram usados em alguns estudos. Em Das et al. [183], NN com preço

Tabela 8

Previsão de preços forex.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte		Crítérios de desempenho	Env.
[168]	EUR/USD, GBP/USD 2009–2012	1976–2003	Dados de preço	–	1 d	CDBN-FG	Lucro	–
[169]	GBP/USD, INR/USD 1976–2003		Dados de preço	10 em	1 s	DBN	RMSE, MAE, MAPE, DA, PCC	–
[170]	CNY/USD,INR/USD 1997–2016		Dados de preço	–	1s	DBN	MAPE, R-quadrado	–
[171]	GBP/USD, BRL/USD, INR/USD		Dados de preço	10 em	1s	DBN + RBM	RMSE, MAE, MAPE, precisão, PCC	–
[172]	Combinação de USD, GBP, EUR, JPY, AUD, CAD, CHF	2009–2016	Dados de preço	–	–	AE empilhado + SVR	MAE, MSE, RMSE	Matlab
[155]	Mercadoria, FX futuro, ETF	1991–2014	Dados de preço	100 * 5 minutos	5 minutos	DN	SR, relação de capacidade, Retorna	C++, Python
[126]	FTSE100, OMX30, S&P500, Mercadoria, Forex	1993–2017	Técnico indicadores	60 dias	1 dia	DN, RN	Precisão, valor-p	–
[173]	EUR/USD	2001–2010	Fechar dados	11 dias	1 dia	RN e mais	MAE, MAPE, RMSE, THEIL-U	–
[174]	EUR/USD	2002–2010	Dados de preço	13 dias	1 dia	RNN, MLP, PSN	MAE, MAPE, RMSE, THEIL-U	–
[175]	EUR/USD, EUR/GBP, EUR/JPY, EUR/CHF	1999–2012	Dados de preço	12 dias	1 dia	RNN, MLP, PSN	MAE, MAPE, RMSE, THEIL-U	–
[176]	RMB contra USD, EUR, JPY, HKD	2006–2008	Dados de preço	10 dias	1 dia	RN, ANN	RMSE, MAE, MSE	–
[177]	EUR/USD, EUR/JPY, USD/JPY, EUR/CHF, XAU/USD, XAG / USD, QM, QG	2011–2012	Dados de preço	–	–	Evolino RNN	Correlação entre valores reais previstos	–
[178]	USD/JPY	2009–2010	Dados de preço, Ouro	–	5 dias	EVOLINO RN + entrada ortogonal dados	RMSE	–
[179]	S&P500, EUR/USD 1950–2016	2016	Dados de preço	30 dias, 30 d * min	1 dia, 1 minuto	Wavelet+CNN	Precisão, perda de log	Difícil
[180]	USD/GBP, S&P500, FTSE100, óleo, ouro		Dados de preço	–	5 minutos	AE + CNN	SR, % de volatilidade, média retorno/transferência, taxa de Retorna	H2O
[148]	ISE100, TRY/USD		Dados de preço	–	2d, 4d, 8 dias, 12 dias, 18 dias	TAR-VEC-MLP, TAR-VEC-RBF, TAR-VEC-RHE	RMSE	–
[164]	DEM/GBP taxa de câmbio	–	Devoluções	–	–	RMDN-GARCH	NMSE, NMAE, RH, WHR	–
[124]	S&P500, NIKKEI225, USD Trocas	2011–2015	Dados de preço	–	1d, 5d, 7d, 10d	LRNFS com FHSO	RMSE, MAPE, MAE	–

Tabela 9

Previsão de preço de criptomoeda.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte		Crítérios de desempenho	Env.
[181]	Bitcoin, Litecoin, StockTwits	2015–2018	OCHLV, técnico indicadores, análise de sentimentos	–	30 minutos, 4h, 1d	CNN, LSTM, Estado Modelo de frequência	MSE	Difícil, Tensorflow
[182]	Bitcoin	2013–2016	Dados de preço	100 dias	30 dias	Bayesiano RNN otimizado, LSTM	Sensibilidade, especificidade, precisão, exatidão, RMSE	Keras, Python, Hiperas

os dados foram usados para previsão de tendência dos índices de ações S&P500. Navon et ai. [184] combinou FNN profundo com uma negociação seletiva unidade de estratégia para prever o próximo preço. Yang et ai. [142] criou um rede ensemble de vários modelos Backpropagation e ADAM para previsão de tendências. Na literatura, LSTM, RNN e Rede Neural Probabilística (PNN) com dados brutos de séries temporais também foram usados para previsão de tendências. Saad et ai. [185] comparou o Neural de Atraso de Tempo Rede (TDNN), RNN e PNN para detecção de tendências usando 10 ações do S&P500. Pérsio et al. [186] comparou 3 RNN diferentes modelos básicos (RNN, LSTM e GRU) para prever o movimento de Preços das ações do Google. Hansson et ai. [133] usou LSTM (e outros técnicas clássicas de previsão) para prever a tendência das ações preços. Em Shen et ai. [187], foram usados os modelos GRU e GRU-SVM

para as tendências do HSI, The Deutscher Aktienindex (DAX), e Índices S&P500. Existem também novos métodos que usam apenas séries temporais brutas dados de preço/índice na literatura. Chen et ai. [188] proposto um método que utilizou uma CNN com Gramian Angular Field (GAF), Mapeamento de média móvel (MAM) e castiçal com conversão dados de imagem. Em Sezer et ai. [189], um novo método de CNN com recurso de imagem foi proposto para a previsão da compra/venda/manter posições dos preços dos Exchange-Traded Funds (ETFs) e Dow30 preços das ações. Zhou et ai. [190] propuseram um método que utiliza Decomposição em modo empírico e máquina de fatoração baseada Modelos de Rede Neural (EMD2FNN) para prever as direções de preços de fechamento de ações com precisão. Em Ausmees et ai. [191], DBN com dados de preço foram usados para previsão de tendência de 23 ações de grande capitalização do índice OMX30.

Tabela 10

Previsão de tendências usando apenas dados brutos de séries temporais.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Critérios de desempenho	Env.
[183]	Estoque S&P500 índices	1963–2016	Dados de preço	30 dias	1 dia NN	Acurácia - Precisão, recordação, pontuação F1, AUROC	R, H2o, Python, Tensorflow
[184]	Espião ETF, 10 ações de S&P500	2014–2016	Dados de preço	60 min 30 min FNN		Ganho cumulativo	MatConvNet, Matlab
[142]	Xangai índice composto e SZSE	1990–2016	OCHLV	20 dias	1 dia Conjuntos de Precisão de ANN		—
[185]	10 ações de S&P500	—	Dados de preço		TDNN, RNN, PNN	Oportunidades perdidas, proporção de falsos alarmes	—
[186]	Estoque GOOGL diariamente dados de preço	2012–2016	Janela de tempo de 30 dias de OCHLV	22 dias, 50 dias, 70 dias	*	LSTM, GRU, Precisão RNN, Logloss	Python, alto
[133]	S&P500, Bovespa50, OMX30	2009–2017	Autoregressivo parte do preço dados	30 dias	1..15 dias LSTM	MSE, Precisão	Tensorflow, Dificil, R
[187]	HSI, DAX, S&P500 1991–2017		Dados de preço	—	1 dia GRU, GRU-SVM	% de retorno diário	Pitão Tensorflow
[188]	Estoque de Taiwan Futuros de Índice	2001–2015	OCHLV	240 dias	1..2d CNN com GAF, MAM, castiçal	Precisão	Matlab
[189]	ETF e Dow30	1997–2007	Dados de preço		CNN com recurso imagem	Retorno anualizado	Dificil, Tensorflow
[190]	SSEC, NASDAQ, S&P500	2007–2016	Dados de preço	20 minutos 7 minutos	EMD2FNN	MAE, RMSE, MAPE	—
[191]	23 ações de limite do OMX30 índice na Nasdaq Estocolmo	2000–2017	Dados de preços e retorna	30 dias	* DBN	MUITO DE	Python, Theano

Tabela 11

Previsão de tendências usando indicadores técnicos e dados de preços e dados fundamentais.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Critérios de desempenho	Env.
[192]	Índice KSE100	—	Dados de preços, vários dados fundamentais	—	— ANN, SLP, MLP, RBF, DBN, SVM	Precisão	—
[193]	Ações no Dow30	1997–2017	RSI (Técnico Indicadores)	200 dias	1 dia DMLP com genética algoritmo	Retorno anualizado	Spark MLlib, Java
[194]	Composto SSE Índice, FTSE100, PingAnBank	1999–2016	Técnico indicadores, OCHLV	24 dias	1 dia RBM	Precisão	—
[195]	Ações Dow30	2012–2016	Dados de preços, vários técnico indicadores	40 dias	— LSTM	Precisão	Python, Dificil, Tensorflow, TALIB
[196]	Preço das ações de IBOVESPA index	2008–2015	Técnico indicadores, OCHLV de preço	—	15 minutos LSTM	Acurácia - Precisão, Recall, pontuação F1, % retorno, máximo rebaixamento	Dificil
[197]	20 ações de NASDAQ e NYSE	2010–2017	Dados de preços, técnico indicadores	5 dias	1 dia LSTM, GRU, SVM, XGBoostName	Precisão	Dificil, Tensorflow, Pitão
[198]	17 ETF	2000–2016	Dados de preços, técnico indicadores	28 dias	1 dia CNN	Precisão, MSE, Lucro, AUROC	Dificil, Tensorflow
[199]	Ações no Dow30 e 9 Volume Superior ETF	1997–2017	Dados de preços, técnico indicadores	20 dias	1 dia CNN com recurso imagem	Recordação, precisão, F1-score, anualizado	Python, Dificil, Tensorflow, Java
[200]	Borsa Istambul 100 Ações	2011–2015	75 técnico indicadores, OCHLV de preço	—	1 hora CNN	Retorna Precisão	Dificil

Alguns estudos usaram indicadores técnicos, dados de preços e dados fundamentais ao mesmo tempo. A Tabela 11 resume o

documentos de previsão de tendências que usaram indicadores técnicos, dados de preços, e dados fundamentais. Além disso, esses estudos são agrupados em três subgrupos: modelos ANN, MLP, DBN e RBM; LSTM e modelos GRU; e novos métodos. ANN, MLP, DBN e RBM métodos foram usados com indicadores técnicos, dados de preços e dados fundamentais em alguns estudos. Em Raza et al. [192], vários modelos clássicos e ML e DBN foram comparados para previsão de tendência. Em Sezer et al. [193], indicadores de análise técnica (RSI) os limites de compra e venda foram otimizados com GA, que foi usado para sinais de compra e venda. Após a otimização, o DMLP também foi usado para

aproximação da função. Liang et al. [194] usou análise técnica parâmetros, OCHLV de preços e RBM para previsão de tendência de ações.

Métodos LSTM e GRU com indicadores técnicos, dados de preços, e dados fundamentais também foram utilizados em alguns artigos. Em Troiano et al. [195], o cruzamento e a convergência média móvel e Sinais de divergência (MACD) foram usados para prever a tendência de Dow 30 preços das ações. Nelson et al. [196] usou LSTM para estoque estimativa de movimento de preços. Canção et al. [197] usaram preços de ações, recursos de análise técnica e quatro modelos diferentes de ML (LSTM, GRU, SVM e eXtreme Gradient Boosting (XGBoost)) para prever a tendência dos preços das ações.

Além disso, novos métodos usando CNN com os dados de preço e indicadores técnicos foram propostos. Gudelek et al. [198]

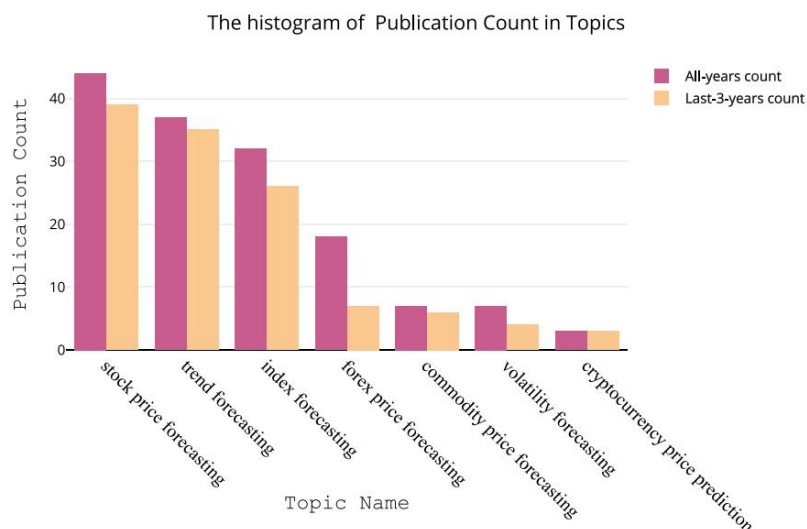


Fig. 5. O histograma da contagem de publicações em tópicos.

converteu as séries temporais de dados de preços em imagens bidimensionais usando análise técnica e as classificou com uma CNN profunda.

Da mesma forma, Sezer et al. [199] também propuseram uma nova técnica que converteu dados de séries temporais financeiras consistindo em saídas de indicadores de análise técnica para imagens bidimensionais e classificou essas imagens usando uma CNN para determinar os sinais de negociação. Gunduz et al. [200] propuseram um método usando uma CNN com características correlacionadas combinadas para prever a tendência dos preços das ações.

Além disso, também existem estudos utilizando técnicas de mineração de texto. A Tabela 12 resume os artigos de previsão de tendências usando técnicas de mineração de texto. Diferentes métodos/modelos são representados por quatro subgrupos: DNN, DMLP e CNN com modelos de mineração de texto; modelo GRU; modelos LSTM, CNN e LSTM+CNN; e novos métodos. No primeiro grupo de estudos, DNN, DMLP e CNN com mineração de texto foram usados para previsão de tendências. Em Huang et al. [201], os autores usaram diferentes modelos, incluindo Hidden Markov Model (HMM), DMLP e CNN usando humores do Twitter, para prever o movimento do dia seguinte. Peng et al. [202] usaram a combinação de mineração de texto e incorporação de palavras para extrair informações de notícias financeiras e um modelo DNN para previsão de tendências de ações.

Além disso, métodos GRU com técnicas de mineração de texto também têm sido usados para previsão de tendências. Huynh et al. [203] usaram notícias financeiras da Reuters e Bloomberg, dados de preços de ações e um modelo Bidirecional Gated Recurrent Unit (Bi-GRU) para prever futuros movimentos de ações. Dang et al. [204] usaram os modelos Stock2Vec e Two-stream GRU (TGRU) para gerar dados de entrada de notícias financeiras e preços de ações. Em seguida, eles usaram a diferença de sinal entre o fechamento anterior e a próxima abertura para a classificação dos preços das ações. Os resultados foram melhores do que os dos modelos de última geração.

Os modelos LSTM, CNN e LSTM+CNN também foram usados para previsão de tendências. Verma et al. [205] combinaram dados de notícias com dados financeiros para classificar o movimento dos preços das ações e os avaliaram com certos fatores. Eles usaram um modelo LSTM como arquitetura NN. Pinheiro et al. [206] propuseram um novo método que usava um modelo de linguagem neural baseado em caracteres usando notícias financeiras e LSTM para previsão de tendências. Em Prosky et al. [207], previsão de sentimento/humor e previsão de preço com base no sentimento, previsão de preço com mineração de texto e modelos DL (LSTM, NN, CNN) foram usados para previsão de tendências. Liu et al. [208] propuseram um método que usava duas redes LSTM separadas para construir uma rede ensemble. Um dos modelos LSTM foi usado para incorporação de palavras com word2Vec para criar uma matriz de entrada de informações para a CNN.

O outro foi usado para previsão de preços usando recursos de análise técnica e preços de ações.

Na literatura, também existem novos métodos para prever a tendência de dados de séries temporais. Yoshihara et al. [209] propuseram um novo método que usa uma combinação de RBM, DBN e incorporação de palavras para criar vetores de palavras para uma rede RNN-RBM-DBN para prever a tendência dos preços das ações. Shi et al. [210] propuseram um novo método chamado DeepClue que interpretava visualmente modelos DL baseados em texto na previsão de movimentos de preços de ações. No método proposto, notícias financeiras, gráficos e tweets de mídia social foram usados juntos para prever o movimento do preço das ações. Zhang et al. [211] propuseram um método que realizava a fusão de informações de várias fontes de notícias e mídias sociais para prever a tendência das ações.

Hu et al. [212] propuseram um novo método que utilizou técnicas de mineração de texto e Redes de Atenção Híbridas com base em notícias financeiras para previsão de tendências de ações. Wang et al. [213] combinou análise técnica e análise de sentimento de mídia social (tópicos financeiros relacionados) e criou um método Deep Random Subspace Ensembles (DRSE) para classificação. Matsubara et al. [214] propuseram um método que usou um Deep Neural Generative Model (DGM) com artigos de notícias usando um algoritmo de vetor de parágrafo para criar o vetor de entrada para previsão de tendências de ações. Li et al. [215] implementaram a classificação intradiária da direção do preço das ações usando notícias financeiras e preços de ações.

Além disso, os estudos também usaram diferentes variações de dados. A Tabela 13 resume os documentos de previsão de tendências usando esses vários dados agrupados em dois subgrupos: modelos LSTM, RNN e GRU e modelos CNN.

Os métodos LSTM, RNN e GRU com várias representações de dados têm sido usados em alguns artigos de previsão de tendências. Tsan tekidis et al. [216] usaram dados de séries temporais de livros de ordens limitadas e um método LSTM para previsão de tendências. Sirignano et al. [217] propuseram um novo método que usava o fluxo do livro de ordens limitado e informações de histórico para determinar os movimentos de estoque usando LSTM. Os resultados do método proposto foram notavelmente estacionários. Chen et al. [154] usaram notícias de mídia social, recursos LDA e um modelo RNN para prever a tendência dos preços do índice. Buczkowski et al. [218] propuseram um novo método que utilizava recomendações de especialistas (comprar, manter ou vender), conjunto de GRU e LSTM para prever a tendência dos preços das ações.

Modelos CNN com diferentes representações de dados também foram usados para previsão de tendências. Tsantekidis et al. [219] usaram as últimas 100 entradas do livro de ordens de limite para criar imagens para previsão de preços de ações usando uma CNN. Usando os dados do livro de ordens com limite

Tabela 12

Previsão de tendências usando técnicas de mineração de texto.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Horizonte	Método	Crítérios de desempenho	Env.
[201]	S&P500, NYSE Composto, DJIA, Composto NASDAQ	2009–2011	humores do Twitter, dados de índice	7 dias	1 dia	DN, CNN	Taxa de erro	Keras, Theano
[202]	Notícias da Reuters e Bloomberg, Estoque histórico dados de segurança	2006–2013	notícias, preço dados	5 dias	1 dia	DN	Precisão	–
[203]	Notícias da Reuters, Bloomberg	2006–2013	notícias financeiras, dados de preço	–	1d, 2d, 5d, 7d	Bi-GRU	Precisão	Python, alto
[204]	Notícias sobre a Apple, Airbus, Amazon de Reuters, Bloomberg, Preços das ações S&P500	2006–2013	Dados de preços, notícias, técnico indicadores	–	–	GRU de dois fluxos, estoque2vec	Acurácia - Precisão, AUROC	Alto, Python
[205]	Índice NIFTY50, NIFTY Banco/Automático/TI/Energia Índice, Notícias	2013–2017	Dados de índice, notícia	1d, 2d, 5 dias	1 dia	LSTM	MCC, Precisão	–
[206]	Notícias da Reuters, Bloomberg, ações dados de preço/índice de S&P500	2006–2013	Notícias e frases	–	1 hora, 1 dia	LSTM	Precisão	–
[207]	30 ações DJIA, S&P500, DJI, notícias da Reuters	2002–2016	Dados de preços e características de novos artigos	1m	1 dia	LSTM, NN, CNN e palavra2vec	Precisão	VADER
[208]	APPL de S&P500 e notícias de Reuters	2011–2017	Notícias, OCHLV, Técnico indicadores	–	1 dia	CNN + LSTM, CNN+SVM	Precisão, pontuação F1	Tensorflow
[209]	Notícias, ações Nikkei Média e 10-Empresas Nikkeis	1999–2008	Notícias, MACD	–	1 dia	Precisão RNN, RBM+DBN , valor P	–	–
[210]	Notícias da Reuters e Bloomberg para Ações S&P500	2006–2015	notícias financeiras, dados de preço	1 dia	1 dia	DeepClue	Precisão	Software Dynet
[211]	Dados de preço, índice dados, notícias, redes sociais dados de mídia	2015	Dados de preços, notícias de artigos e mídia social	1 dia	1 dia	Matriz acoplada e tensor	Precisão, MCC	Jieba
[212]	Notícias e chinês dados de estoque	2014–2017	Palavras selecionadas em uma notícia	10 dias	1 dia	ELE	Precisão, Anual Retorna	–
[213]	Sina Weibo, Estoque recordes de mercado	2012–2015	Técnico indicadores, frases	–	–	DRSE	Pontuação F1, precisão, recordação, precisão, AUROC	Pitão
[214]	Nikkei225, S&P500, notícias da Reuters e Bloomberg	2001–2013	Dados de preços e notícia	1 dia	1 dia	DGM	Precisão, MCC, %lucro	–
[215]	Notícias, preços das ações De Hong Kong Bolsa de Valores	2001	Dados de preços e TF-IDF de notícia	60 minutos	(1..6)*5 min	ELM, DLR, PCA, BELM, KELM, NN	Precisão	Matlab

para criar um formato de matriz 2D com uma CNN para prever movimento direcional foi inovador. Em Doering et al. [159], HFT a previsão de microestrutura foi implementada com uma CNN.

5. Instantâneo atual do campo

Depois de revisar todos os trabalhos de pesquisa especificamente direcionados às implementações de previsão de séries temporais financeiras usando modelos DL, agora estamos prontos para fornecer algumas estatísticas gerais sobre o estado atual do campo. O número de artigos incluídos em nosso pesquisa foi de 140. Classificamos os artigos de acordo com sua tipo de ativo previsto. Também analisamos os estudos com base em suas Escolhas de modelo DL, frameworks para o ambiente de desenvolvimento, conjuntos de dados, benchmarks comparáveis e alguns outros diferenciais critérios como conjuntos de recursos e números de citações, que podem não ser incluído neste trabalho devido a restrições de espaço. Vamos agora resumir nossas observações notáveis para fornecer informações pesquisas com destaques importantes dentro da área.

A Fig. 5 apresenta os vários tipos de ativos para os quais os pesquisadores desenvolveram seus modelos de previsão correspondentes. Como esperado, estudos de previsão relacionados ao mercado de ações dominam o campo. Estoque previsão de preço, previsão de tendência e previsão de índice foram as três principais escolhas para pesquisas de previsão de séries temporais financeiras.

Até agora, 46 artigos foram publicados para previsão de preços de ações, 38 para previsão de tendências e 33 para previsão de índices. Esses estudos constituem mais de 70% de todos os estudos, indicando alta interesse. Além do acima, havia 19 artigos sobre previsão cambial e 7 sobre previsão de volatilidade. Enquanto isso criptomoeda a previsão começou a atrair pesquisadores; porém, apenas 3 artigos sobre este tema já foram publicados, mas este número é deverá aumentar nos próximos anos [220]. Fig. 6 destaques a taxa de publicação conta para várias áreas de implementação através dos anos. Enquanto isso, a Fig. 7 fornece mais detalhes sobre a escolha de modelos DL em várias áreas de implementação.

A Fig. 8 ilustra o crescente apetite dos pesquisadores para desenvolver modelos DL para implementações de séries temporais financeiras. Enquanto isso, como a Fig. 9 indica, a maioria dos estudos foi publicada em periódicos (57 deles) e conferências (49 artigos), mas um número de artigos do arXiv (11) e teses de pós-graduação (6) também existem.

Uma das questões mais importantes para um pesquisador é onde eles podem publicar suas descobertas. Durante nossa revisão, também investigamos cuidadosamente onde cada artigo foi publicado. Nós tabulamos nossos resultados para os principais periódicos para previsão de séries temporais financeiras na Fig. 10. De acordo com esses resultados, os periódicos com os artigos mais publicados incluem Expert Systems with Applications, Neurocomputing, Applied Soft Computing, The Journal

Tabela 13
Previsão de tendências usando vários dados.

Arte.	Conjunto de dados	Período	Conjunto de características	Atraso	Método Horizonte	Crítérios de desempenho	Env.	
[216]	Nasdaq Nordic (Kesko Corporation, Outokumpu Oyj, Sampo, Rautaruukki, Wartsila Oyj)	2010	Dados de preço e volume em LOB	100 segundos	10 segundos, 20 segundos, desliza de 50	LSTM	Precisão, Recordação, Pontuação F1, k de Cohen	—
[217]	Alta frequência registro de todos ordens	2014–2017	Dados de preços, registro de todos pedidos, transações	2 horas	—	LSTM	Precisão	—
[154]	Chinês, O Xangai Shenzhen 300 Índice de ações (HS300)	2015–2017	Social media news (Sina Weibo), dados de preços	1 dia	1 dia	RNN-Reforço com LDA	Precisão, MAE, MAPE, RMSE	Python, Scikit aprender
[218]	Dados ISMIS 2017 Mineração Concorrência conjunto de dados	—	Identificador especialista, classe previsto por especialista	—	—	LSTM + GRU + FCNN	Precisão	—
[219]	Nasdaq Nordic (Kesko Corporation, Outokumpu Oyj, Sampo, Rautaruukki, Wartsila Oyj)	2010	Preço, dados de volume, 10 ordens do LOB	—	—	CNN	Precisão, Recordação, Pontuação F1, k de Cohen	Theano, Scikit aprender, Python
[159]	Bolsa de Londres Intercâmbio	2007–2008	Limite o estado do livro de pedidos, negociações, ordens de compra/venda, exclusões de pedidos	—	—	CNN	Precisão, kapa	Café

The Rate of Publication Count in Topics

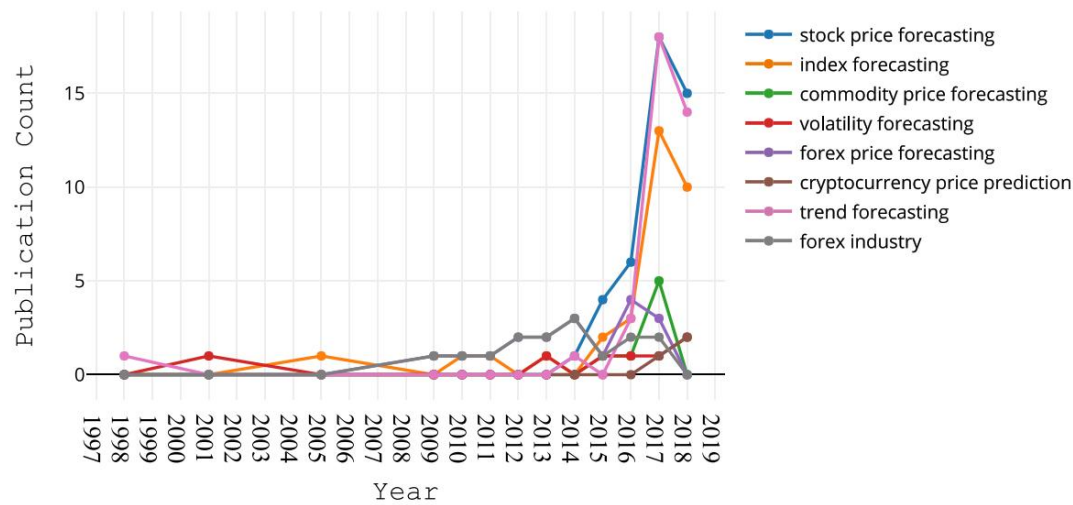


Fig. 6. Taxa de contagem de publicações em tópicos.

de Supercomputação, Sistemas de Apoio à Decisão, Systems, European Journal of Operational Research e IEEE Acesso. Os pesquisadores interessados também devem considerar as tendências nos últimos 3 anos, pois as tendências podem variar dependendo do áreas de implementação específicas.

A análise cuidadosa da Fig. 11 valida claramente a dominância de Modelos baseados em RNN (65 artigos) entre todos os outros para o modelo DL escolhas, seguido por DMLP (23 artigos) e CNN (20 artigos). o círculo interno representa todos os anos considerados, enquanto o círculo externo fornece apenas os estudos nos últimos 3 anos. Devemos notar que o RNN é um modelo geral com várias versões, incluindo LSTM e GRU. Para RNN, os pesquisadores preferem principalmente LSTM devido a sua fase de desenvolvimento de modelo relativamente simples; no entanto, outros tipos de RNN também são comuns. A Fig. 12 fornece um instantâneo de a distribuição do modelo RNN. Conforme mencionado acima, a LSTM teve a maior interesse entre todos com 58 artigos, enquanto Vanilla RNN e GRU tiveram 27 e 10 artigos, respectivamente. Portanto, fica claro

que LSTM é o modelo DL mais popular para séries temporais financeiras estudos de previsão e regressão.

Enquanto isso, DMLP e CNN eram geralmente preferidos para problemas de classificação. Como os dados de séries temporais geralmente consistem de componentes temporais, algum pré-processamento de dados pode ser necessário antes que a classificação real possa ocorrer. Daí, muitos essas implementações utilizam extração de recursos, técnicas de seleção e possíveis métodos de redução de dimensionalidade. Muitos pesquisadores usam principalmente DMLP devido ao fato de que sua rasa A versão MLP tem sido usada extensivamente antes e tem um histórico de sucesso para muitas aplicações financeiras diferentes, incluindo a previsão de séries temporais financeiras. Coerente com o nosso observações, a DMLP também foi mais preferida nas ações, índices, e previsão de tendência particular porque é, por definição, um problema de classificação com dois (tendência de alta ou tendência de baixa) e três instâncias de classe (tendência de alta, estacionária ou tendência de baixa).

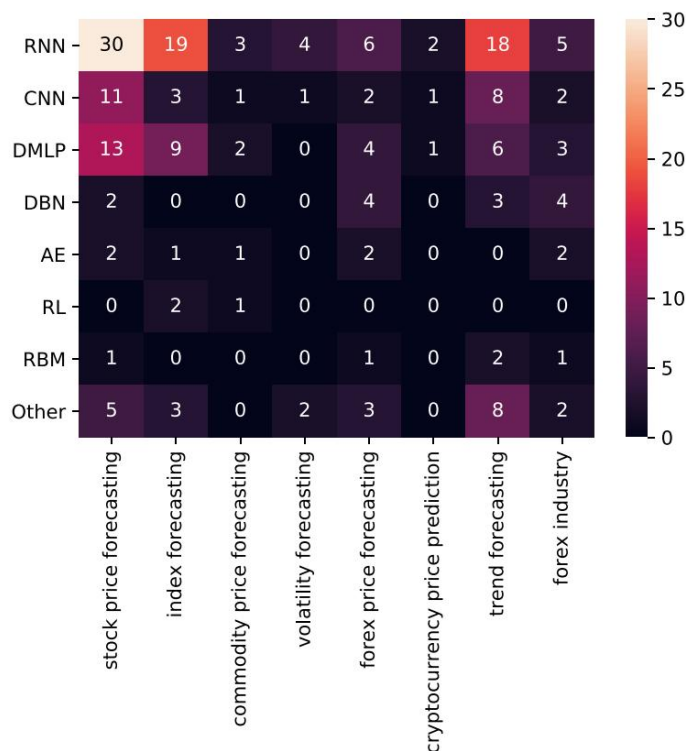


Fig. 7. Mapa de calor do modelo de tópico.

Além do DMLP, a CNN também é uma escolha popular para implementações de previsão de séries temporais financeiras do tipo classificação. A maioria desses estudos apareceu nos últimos 3 anos. Como mencionado anteriormente, para converter dados sequenciais variáveis no tempo em uma forma classificável mais estacionária, algum pré-processamento pode ser necessário. Embora existam algumas representações 1-D, a implementação 2-D para CNN é mais comum, principalmente herdada por meio de aplicativos de reconhecimento de imagem da CNN de implementações de visão computacional. Em alguns estudos [188,189,193,199,219], foram adaptadas transformações inovadoras de dados de séries temporais financeiras em uma representação semelhante a uma imagem, e desempenhos impressionantes foram alcançados. Como resultado, a CNN pode aumentar sua participação nas previsões de séries temporais financeiras nos próximos anos.

Como uma nota final, a Fig. 13 mostra quais frameworks e plataformas os pesquisadores e desenvolvedores usaram ao implementar seu trabalho. Nós tentamos o nosso melhor para extrair esta informação dos jornais. No entanto, devemos ter em mente que nem todas as publicações forneceram seu ambiente de desenvolvimento. Além disso, a maioria dos trabalhos não forneceu detalhes, impedindo-nos de um gráfico de comparação mais completo, ou seja, alguns pesquisadores afirmaram usar Python, mas nenhuma informação adicional foi fornecida, enquanto outros mencionaram o uso de Keras ou TensorFlow, fornecendo mais detalhes. Além disso, na seção "Outros", o uso do Pytorch aumentou no último ano, embora não seja visível no gráfico. Independentemente disso, as ferramentas relacionadas ao Python foram as tecnologias mais influentes por trás das implementações abordadas nesta pesquisa.

6. Discussão e questões em aberto

Do ponto de vista de aplicação, embora a previsão de séries temporais financeiras tenha um foco relativamente estreito, ou seja, as implementações foram baseadas principalmente na previsão de preços ou tendências, dependendo do modelo DL subjacente, existem modelos muito diferentes e versáteis na literatura. Devemos lembrar que mesmo

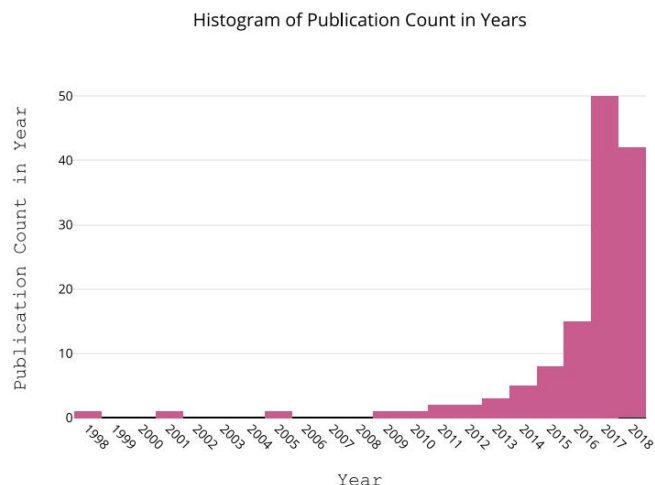


Fig. 8. Histograma da contagem de publicações em anos.

embora a previsão de séries temporais financeiras seja um subconjunto de estudos de séries temporais, devido às expectativas de lucro embutidas de modelos de previsão bem-sucedidos, existem algumas diferenças, de modo que uma precisão de previsão mais alta às vezes pode não refletir um modelo lucrativo. Portanto, a estrutura de risco e recompensa também deve ser levada em consideração. Neste ponto, tentaremos elaborar nossas observações sobre essas diferenças em vários projetos e implementações de modelos.

6.1. Modelos DL para previsão de séries temporais financeiras

De acordo com as estatísticas da publicação, o LSTM foi a escolha preferida da maioria dos pesquisadores para previsão de séries temporais financeiras. O LSTM e suas variações utilizaram dados variantes no tempo com representações incorporadas de feedback, resultando em desempenhos mais altos para implementações de previsão de séries temporais. Como a maioria dos dados financeiros, de uma forma ou de outra, inclui componentes dependentes do tempo, o LSTM foi a escolha natural em problemas de previsão de séries temporais financeiras. Enquanto isso, LSTM é um modelo DL especial derivado de uma família de classificadores mais geral, chamada RNN.

A análise cuidadosa da Fig. 11 ilustra a dominância de RNNs (que consistem principalmente de LSTM). De fato, mais da metade dos artigos publicados sobre previsão de séries temporais se enquadram na categoria de modelo RNN. Independentemente do tipo de problema, preço ou previsão de tendência, a natureza ordinal da representação de dados forçou os pesquisadores a considerar RNN, GRU e LSTM como preferências viáveis para suas escolhas de modelo. Assim, os modelos RNN foram escolhidos, pelo menos para benchmarking, em muitos estudos para comparação de desempenho com outros modelos desenvolvidos.

Enquanto isso, outros modelos também foram usados para problemas de previsão de séries temporais. Entre esses, o DMLP teve o maior interesse devido ao domínio de mercado de seu primo superficial (MLP) e sua ampla aceitação e longa história na sociedade ML. No entanto, há uma diferença fundamental em como os modelos baseados em DMLP e RNN foram usados para problemas de previsão de séries temporais financeiras.

O DMLP se encaixa bem para problemas de regressão e classificação. No entanto, em geral, a independência da ordem dos dados deve ser preservada para melhor utilizar a dinâmica interna de funcionamento dessas redes, embora alguns ajustes possam ser feitos através da configuração do algoritmo de aprendizado. Na maioria dos casos, os componentes de tendência dos dados precisam ser removidos da série temporal subjacente ou algumas transformações de dados podem ser necessárias para que os dados resultantes se tornem estacionários. Independentemente disso, algum pré-processamento cuidadoso pode ser necessário para que um modelo DMLP seja bem-sucedido. Em contraste,

The histogram of Publication Count in Publication Type

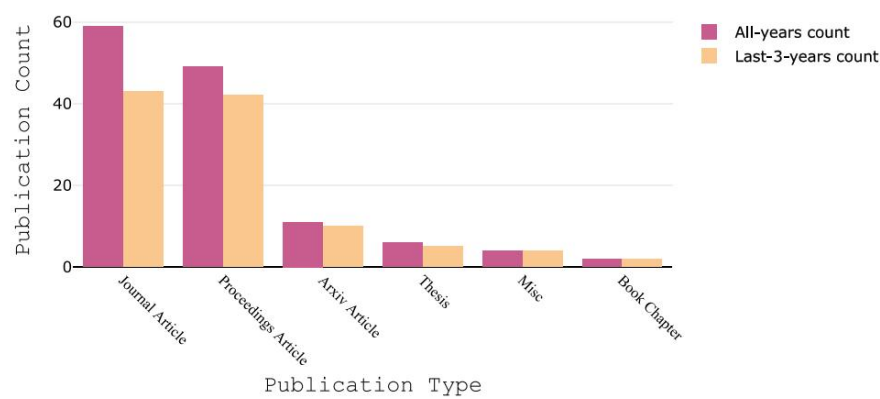


Fig. 9. O histograma da contagem de publicações em tipos de publicação.

The histogram of Top Journals

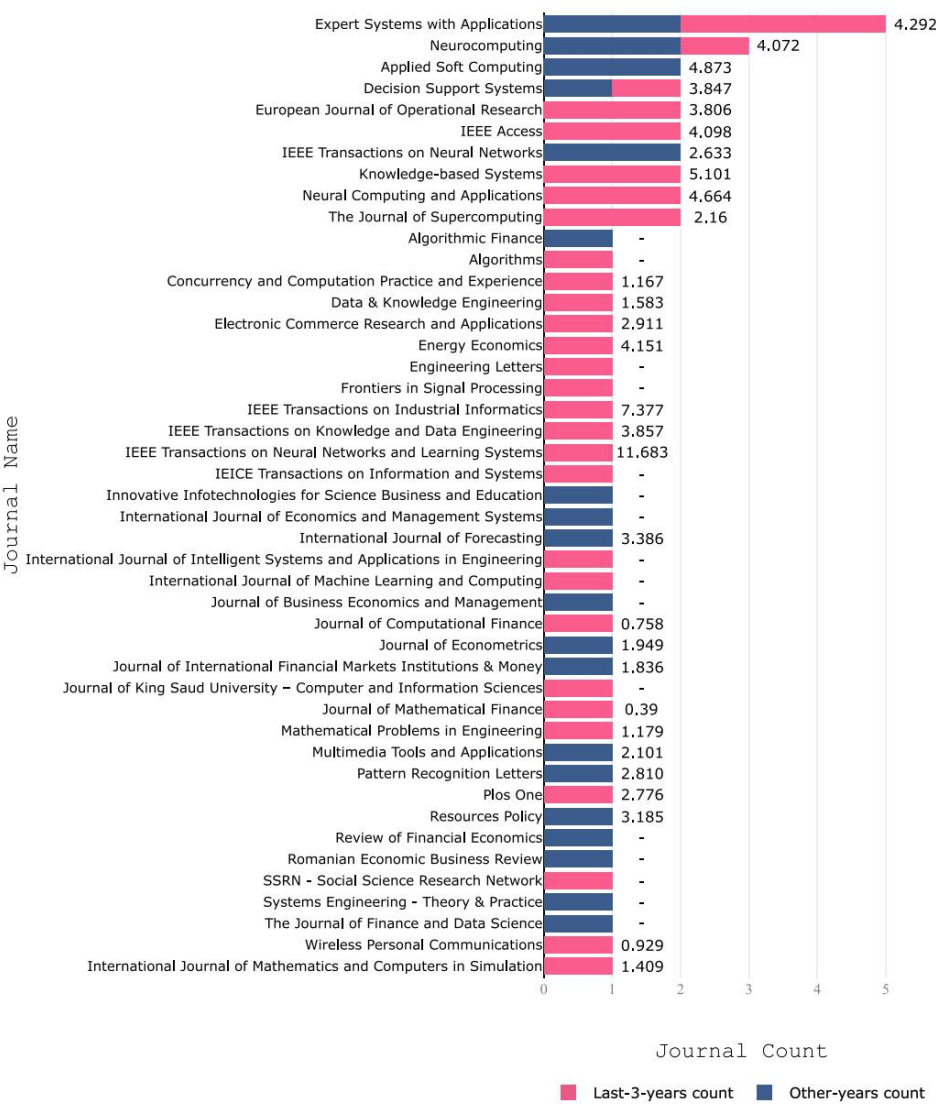


Fig. 10. Principais periódicos — os números correspondentes ao lado do gráfico de barras representam o fator de impacto dos periódicos.

Publication Count in Model Type

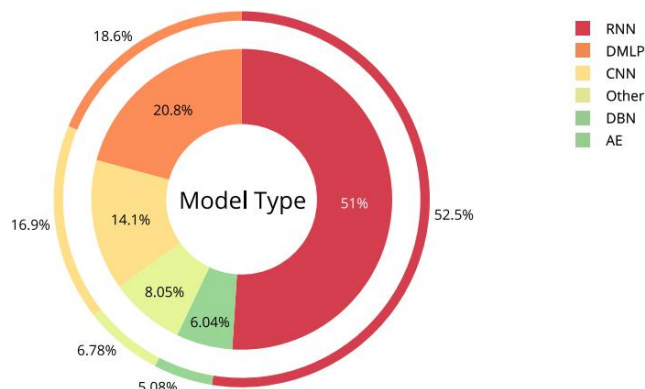


Fig. 11. O gráfico de pizza da contagem de publicações em tipos de modelo.

Distribution of RNN Models

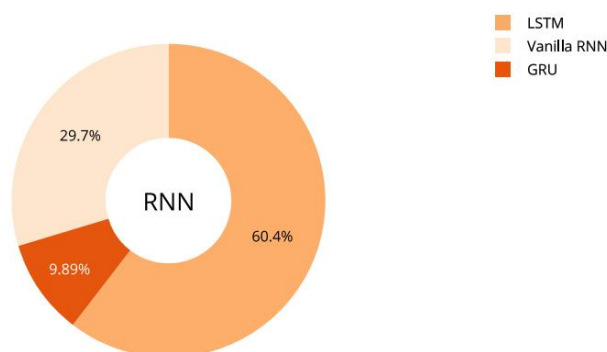


Fig. 12. Distribuição dos modelos RNN.

Frameworks

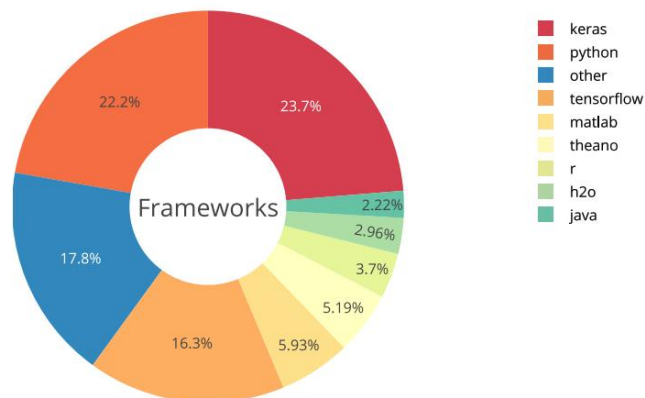


Fig. 13. Os Ambientes de Desenvolvimento Preferidos.

Modelos baseados em RNN podem trabalhar diretamente com dados que variam no tempo, tornando mais fácil para os pesquisadores desenvolverem modelos DL.

Como resultado, a maioria das implementações de DMLP tinha pré-processamento de dados embutido antes do estágio de aprendizado. No entanto, essa inconveniência não impediu que os pesquisadores utilizassem o DMLP e suas variações durante o processo de desenvolvimento do modelo. Em vez disso, muitas representações de dados versáteis foram tentadas para alcançar maior

desempenho geral de previsão. Uma combinação de parâmetros de análise fundamental e/ou técnica juntamente com outros recursos, como sentimento financeiro por meio de mineração de texto, foram incorporados a esses modelos. Na maioria dos estudos DMLP, o problema correspondente foi tratado como classificação, especialmente em modelos de previsão de tendência, enquanto os modelos baseados em RNN previram diretamente o próximo valor da série temporal. Ambas as abordagens tiveram algum sucesso em superar o benchmark subjacente; portanto, não é possível reivindicar a superioridade de um tipo de modelo sobre outro. No entanto, como regra geral, os pesquisadores preferem modelos baseados em RNN para regressão de séries temporais e DMLP para classificação de tendências (ou identificação de pontos de compra e venda).

Outro modelo que ganhou popularidade recentemente é a CNN. A CNN também funciona melhor para problemas de classificação e, ao contrário dos modelos baseados em RNN, é mais adequada para representações de dados não variáveis no tempo ou estáticas. Os comentários feitos sobre o DMLP também são válidos principalmente para a CNN. Além disso, ao contrário do DMLP, a CNN exige principalmente a localidade dentro da representação de dados para resultados de classificação de melhor desempenho. Uma área de implementação particular da CNN são os problemas de reconhecimento de objetos baseados em imagens. Nos últimos anos, os modelos baseados em CNN dominaram esse campo, superando com folga todos os outros modelos. Enquanto isso, a maioria dos dados financeiros varia no tempo e pode não ser fácil implementar a CNN diretamente para aplicativos financeiros. No entanto, em alguns estudos recentes, vários grupos de pesquisa independentes seguiram uma transformação inovadora de dados financeiros 1-D variantes no tempo em dados 2-D, principalmente estacionários, semelhantes a imagens, para que pudessem utilizar o poder da CNN por meio de filtragem adaptativa e dimensionalidade implícita. redução. Assim, com essa abordagem, eles foram capazes de desenvolver modelos de sucesso.

Há também uma tendência crescente de usar implementações de negociação algorítmica financeira baseada em RL profundo; estes estão principalmente associados a vários modelos baseados em agentes, onde diferentes agentes interagem e aprendem com suas interações. Este campo tem ainda mais oportunidades para oferecer com avanços na análise de sentimentos financeiros por meio de mineração de texto para capturar a psicologia do investidor; como resultado, as finanças comportamentais podem se beneficiar desses estudos específicos associados a modelos de aprendizado baseados em RL acoplados a estudos baseados em agentes.

Outros modelos, incluindo DBN, AE e RBM, também foram usados por vários pesquisadores, e desempenhos superiores foram relatados em alguns de seus trabalhos. No entanto, leitores interessados devem verificar esses estudos caso a caso para ver como eles foram modelados a partir da representação de dados e perspectivas de aprendizagem.

6.2. Discussões sobre recursos selecionados

Independentemente do problema de previsão subjacente, os dados brutos das séries temporais quase sempre foram incorporados direta ou indiretamente dentro do vetor de características, o que é particularmente válido para modelos baseados em RNN. No entanto, na maioria dos outros tipos de modelos, outros recursos também foram incluídos. Os recursos de análise fundamentalista e análise técnica estavam entre as escolhas mais favoráveis para estudos de previsão de ações/índices.

Enquanto isso, nos últimos anos, a mineração de textos financeiros ganhou atenção especial, principalmente para extrair o sentimento do investidor/comerciante. O fluxo de streaming de notícias financeiras, tweets, declarações e blogs permitiu que os pesquisadores construíssem modelos de previsão e avaliação melhores e mais versáteis, integrando dados numéricos e textuais. A metodologia geral envolve extrair a análise de sentimento financeiro por meio de mineração de texto e combinar essas informações com dados de análise fundamental/técnica para obter um melhor desempenho geral. É lógico supor que essa tendência continuará com a integração de técnicas mais avançadas de texto e PNL.

6.3. Discussões sobre os tipos de ativos previstos

Embora a previsão de preços forex seja sempre popular entre pesquisadores e profissionais, a previsão de ações/índices sempre teve o maior interesse entre todos os grupos de ativos. Independentemente do preço/previsão de tendências e modelos de negociação de algo foram principalmente incorporados com esses estudos de previsão.

Atualmente, a pesquisa de previsão de séries temporais financeiras sobre criptomoedas é um tema quente. A previsão de preços de criptomoedas tem crescente demanda da comunidade financeira. Porque isso tópico é relativamente novo, podemos ver mais estudos e implementações em um futuro próximo devido às altas expectativas e recompensas.

Houve também uma série de publicações sobre preços de commodities pesquisas de previsão, particularmente para prever o preço do petróleo. Óleo a previsão de preços é crucial devido ao seu tremendo efeito sobre o mundo atividades econômicas e planejamento. Enquanto isso, o ouro é considerado um investimento seguro e quase todos os investidores, ao mesmo tempo, consideram alocar parte de sua carteira para investimentos. Em tempos de incerteza política, muitas pessoas se voltam ao ouro para proteger suas economias. Embora não tenhamos encontrado um notável estudo para a previsão do preço do ouro, devido ao seu histórico importância, pode haver oportunidades nesta área em anos para venha.

6.4. Questões em aberto e trabalhos futuros

Apesar da motivação geral para a previsão de séries temporais financeiras permanecer praticamente inalterada, os meios para atingir os objetivos financeiros variam dependendo das escolhas e do trade-off entre técnicas tradicionais e modelos recém-desenvolvidos. Porque nosso foco fundamental está na aplicação de DL para tempo financeiro estudos de série, tentaremos avaliar o estado atual da pesquisa e extrapolar isso para o futuro.

6.4.1. Escolhas de modelos para o futuro

A predominância de modelos baseados em RNN para previsão de preço/tendência provavelmente não desaparecerá tão cedo, principalmente devido à sua fácil adaptação à maioria dos problemas de previsão de ativos. Enquanto isso, algumas versões aprimoradas do LSTM ou RNN original modelos, geralmente integrados com sistemas híbridos de aprendizagem, são agora se tornando mais comum. Os leitores devem verificar individualmente estudos e avaliar seus desempenhos para ver qual deles se encaixa melhor para suas necessidades particulares e requisitos de domínio.

Observamos um interesse crescente em implementações 2-D CNN de problemas de previsão financeira convertendo o tempo series em um tipo de dados semelhante a uma imagem. Esta metodologia inovadora parece funcionar de forma bastante satisfatória e oferece oportunidades promissoras. Mais estudos desse tipo provavelmente continuarão no futuro próximo.

Atualmente, novos modelos são gerados a partir de modelos mais antigos através da modificação ou aprimoramento de modelos existentes para que melhores desempenhos possam ser alcançados. Tais topologias incluem Generative Adversarial Network (GAN) e redes Capsule. Eles têm sido usados em vários estudos não financeiros; porém, financeiro a previsão de séries temporais ainda não foi investigada para esses modelos. Como tal, pode haver oportunidades interessantes tanto de pesquisa e pontos de vista práticos.

Outro modelo de DL que tem sido investigado minuciosamente é Gráfico CNN. Os gráficos podem ser usados para representar portfólios, redes de comunidades financeiras, dados de análise fundamental, etc. Mesmo que algoritmos de grafos possam ser aplicados diretamente a tais configurações, diferentes representações gráficas também podem ser implementadas para problemas de previsão de séries temporais. Sem muito trabalho foi feito sobre este tópico em particular; no entanto, através do gráfico

representações de dados de séries temporais e implementação de algoritmos de análise de gráficos ou implementação de CNN através desses gráficos são entre as possibilidades que os pesquisadores podem escolher.

Como nota final para modelos futuros, acreditamos que modelos baseados em agentes e RL profundos oferecem grandes oportunidades para pesquisadores. Os algoritmos HFT e os sistemas de consultoria robótica dependem muito sistemas automatizados de negociação algorítmica que podem decidir o que comprar e quando comprar sem qualquer intervenção humana. Esses modelos acima mencionados podem se encaixar muito bem em tais desafios ambientes. A ascensão das máquinas também levará a uma corrida armamentista tecnológica (e algorítmica) entre as fintechs e quant fundos para serem os melhores em sua busca interminável por "alcançar alfa". Novas pesquisas nessas áreas podem ser exatamente o que é necessário.

6.4.2. Projeções futuras para previsão de séries temporais financeiras

Muito provavelmente, para o futuro previsível, a previsão de séries temporais financeiras terá uma estreita cooperação de pesquisa com outras áreas de aplicação financeira, como negociação algorítmica e gestão de portfólio, como era o caso antes. No entanto, mudanças nas características dos dados disponíveis e na introdução de novas classes de ativos podem não apenas alterar as estratégias de previsão de desenvolvedores, mas também forçar os desenvolvedores a buscar novos ou alternativos técnicas para melhor se adaptar a esses novos desafios de trabalho condições. Além disso, métricas como Pontuação de capacidade de probabilidade ranqueada contínua (CRPS) [221] para avaliar distribuições de probabilidade podem ser incluídos para uma análise mais completa.

Uma tendência crescente, não apenas para previsão de séries temporais financeiras, mas para todos os sistemas inteligentes de suporte à decisão, é a interação humano-computador e a pesquisa de PNL. Dentro desse campo, o texto mineração e análise de sentimento financeiro são de particular importância para a previsão de séries temporais financeiras. As finanças comportamentais podem beneficiar dos novos avanços nestes domínios.

Para utilizar o poder da mineração de texto, pesquisadores começaram desenvolvendo novas representações de dados, como Stock2Vec [204], que pode ser útil para combinar dados textuais e numéricos para melhores modelos de previsão. Além disso, o ensemble baseado em PNL modelos que integram semântica de dados com dados de séries temporais podem aumentar a precisão dos modelos existentes.

Uma área que pode se beneficiar significativamente da interligação mercados financeiros é um modelo de negociação de arbitragem estatística automatizada desenvolvimento. Ele já foi usado em mercados de forex e commodities antes. Além disso, muitos profissionais atualmente buscam oportunidades de arbitragem nos mercados de criptomoedas [220] devido à o grande número de moedas disponíveis em vários mercados. Rupturas de preços, alta volatilidade e variações de spread de compra e venda causar oportunidades de arbitragem em diferentes plataformas. Alguns oportunistas desenvolvem modelos de software que podem rastrear esses preços anomalias para a materialização instantânea dos lucros. Além disso, é possível construir pares de carteiras de negociação em diferentes ativos classes usando modelos apropriados. É possível que os modelos DL podem aprender (ou prever) essas oportunidades com mais rapidez e eficiência do que os sistemas clássicos baseados em regras. Isso também beneficiará Estudos de HFT, que estão constantemente em busca de soluções mais rápidas e algoritmos de negociação eficientes e sistemas embarcados com latência mínima. Para conseguir isso, a Unidade de Processamento Gráfico (GPU) - ou Soluções de hardware baseadas em Field Programmable Gate Array (FPGA) incorporados com modelos DL podem ser utilizados. Há uma falta de pesquisa realizada neste aspecto de hardware do tempo financeiro previsão de séries e negociação algorítmica. Dado que existe poder computacional suficiente disponível, vale a pena investigar a possibilidade de algoritmos melhores, porque as recompensas são altas.

6.5. Respostas às nossas perguntas iniciais de pesquisa

Agora estamos prontos para responder às nossas questões de pesquisa inicialmente formuladas. Com base em nossas observações, nossas respostas a essas perguntas são as seguintes:

- Quais modelos de DL são usados para previsão de séries temporais financeiras no?
Resposta: Os modelos baseados em RNN (particularmente LSTM) são os modelos mais comumente usados. Enquanto isso, CNN e DMLP têm sido usados extensivamente em implementações do tipo classificação (como classificação de tendência), desde que o processamento de dados apropriado seja aplicado aos dados brutos.
- Como o desempenho dos modelos DL se compara ao de seus equivalentes tradicionais de aprendizado de máquina?

Resposta: Na maioria dos estudos, os modelos DL foram melhores que os ML. No entanto, também houve muitos casos em que seus desempenhos foram comparáveis. Houve até dois estudos particulares [82,175] em que os modelos ML tiveram um desempenho melhor do que os modelos DL. Enquanto isso, a preferência de implementações de DL sobre modelos de ML está crescendo. Avanços no poder de computação, disponibilidade de big data, desempenho superior, recursos implícitos de aprendizado de recursos e ambiente de desenvolvimento de modelo amigável para modelos DL estão entre os principais motivos para essa migração.

Uma questão importante que pode valer a pena mencionar é a possibilidade do viés de publicação de DL sobre modelos de ML. Como o DL é mais recente que o ML, uma implementação de DL bem-sucedida publicada pode atrair mais público do que um modelo de ML bem-sucedido comparável. Assim, os pesquisadores implicitamente podem ter uma motivação adicional para desenvolver modelos DL.

No entanto, esta é provavelmente uma preocupação válida para todas as publicações acadêmicas, independentemente da área de estudo [222].

Enquanto isso, nesta pesquisa, nosso objetivo foi extrair os estudos de DL publicados para previsão financeira sem nenhuma suposição prévia, para que o leitor possa decidir qual modelo funciona melhor para ele por meio de suas próprias experiências.

- Qual é a direção futura da pesquisa DL para previsão de séries temporais financeiras?
Resposta: Modelos híbridos baseados em NLP, semântica e mineração de texto combinados com dados de séries temporais podem ser mais comuns em um futuro próximo.

7. Conclusões

A previsão de séries temporais financeiras é muito popular entre os pesquisadores de ML há mais de 40 anos. A comunidade financeira teve um novo impulso ultimamente com a introdução de implementações de EAD para pesquisa de previsão financeira, e muitas novas publicações surgiram nesse sentido. Em nossa pesquisa, queríamos revisar os estudos existentes para fornecer um instantâneo do status atual da pesquisa de implementações de DL para previsão de séries temporais financeiras. Agrupamos os estudos de acordo com suas classes de ativos pretendidas juntamente com o modelo de DL preferido associado ao problema. Nossas descobertas indicam que, embora a previsão financeira tenha um longo histórico de pesquisa, o interesse geral na comunidade de EAD está aumentando por meio da utilização de novos modelos de EAD; portanto, existem muitas oportunidades para os pesquisadores.

Declaração de interesse concorrente

Nenhum autor associado a este artigo divulgou quaisquer conflitos potenciais ou pertinentes que possam ser percebidos como conflito iminente com este trabalho. Para declarações de divulgação completas, consulte <https://doi.org/10.1016/j.asoc.2020.106181>.

Reconhecimento

Este trabalho é apoiado por Scientific and Technological Research of the Turkey (TUBITAK) outorga nº 215E248.

Glossário

Algoritmo de gradiente adaptativo **AdaGrad** . 3

Estimativa de Momento Adaptativo **ADAM** . 3, 4, 6, 16

Codificador de carro **AE** . 3, 4, 6-9, 11, 13-16, 23

Inteligência Artificial **IA** . 2

Amostragem de Importância Recozida **AIS** . 6

AMEX Bolsa de Valores Americana. 10

Rede Neural Artificial **ANN** . 2–6, 9, 11, 13, 15–17

AR Retorno Ativo. 10

AR Autoregressivo. 10

Heterocedasticidade condicional autorregressiva **ARCH** . 15

ARIMA Autoregressive Integrated Moving Average. 8, 9, 15

Média Móvel Autoregressiva **ARMA** . 12, 14, 15

Área **AUC** sob a curva. 10

Área **AUROC** sob as características operacionais do receptor. 17, 19

B&H Compre e Retenha. 11

BELM Basic Extreme Learning Machine. 19

Unidade Recorrente com Porta Bidirecional **Bi-GRU** . 18, 19

Bi-LSTM LSTM bidirecional. 11–13

Índice da Bolsa de Valores de Istambul **BIST** . 9-11, 13

Bolsa de Valores Brasileira da **Bovespa** . 11

Retropropagação BPTT através do tempo. 4

BSE Bolsa de Valores de Bombaim. 14

Índice de Canal de Commodities **CCI** . 12

CD Divergência Contrastiva. 6

CDAX German Stock Market Index Calculado pela Deutsche Boerse. 10

Redes de Crença Profunda de Valor Contínuo **CDBN** . 14

Granulação difusa **de CDBN-FG** com redes de crença profunda de valor contínuo. 16

Rede Neural Convolutiva **CNN** . 1–3, 5, 8–21, 23–25

CRBM Máquina Boltzman Restrita Contínua. 14

Pontuação de Probabilidade Classificada Contínua **CRPS** . 24

Bolsa de Valores **CSE** Colombo. 9

CSI China Securities Index. 9, 11, 12, 14, 15

Precisão de direção **DA** . 16

26	OB Sezer, MU Gudelek e AM Ozbayoglu / Applied Soft Computing Journal 90 (2020) 106181
DAX O índice de ações alemão. 15-17	Heteroscedas condicionais auto-regressivas generalizadas GARCH ticidade. 12, 13, 15, 16, 27
DBN Deep Belief Network. 1, 3, 6, 8, 9, 12-14, 16-19, 23	GASVR GA com um SVR. 11, 13, 14, 26
DE Evolução Diferencial. 14	Árvores reforçadas com gradiente GBT . 8, 9
Rede Neural de Feedforward Profundo DFNN. 8, 10, 11, 13, 15	GDAX Global Digital Asset Exchange. 14
Modelo Gerativo Neural Profundo do DGM . 18, 19	Modelo Linear Generalizado GLM . 13
Índice DJI Dow Jones. 19	Modelo Linear Generalizado GML . 11
DJIA Dow Jones Industrial Average. 11, 13, 14, 19	GP Programação Genética. 2, 14
DL Aprendizagem Profunda. 1–6, 8–13, 15, 18–21, 23–25	GPA A Abordagem do Processo Gaussiano. 3, 4
DLR Deep Learning Representação. 19	GRU Unidade Recorrente Fechada. 4, 8-10, 15-21
DMLP Deep Multilayer Perceptron. 3-5, 11, 14, 17, 18, 20, 21, 23, 25	Pesquisa de Grade GS . 3–8
Rede Neural Profunda DNN . 3, 5, 8–16, 18, 19	Índice de Preços de Commodities GSPC S&P500. 13
DOW30 Dow Jones Industrial Average 30. 11, 12, 14	Rede de Atenção Híbrida HAN . 19
Programação Dinâmica DP . 7	HAR Processo Autorregressivo Heterogêneo. 11, 13, 26
Precisão da previsão de direção do DPA . 10	HAR-GASVR HAR com GASVR. 13, 15
DRL Deep Reinforcement Learning. 2, 3, 8	Negociação de alta frequência HFT . 8, 12, 19, 24
DRSE Deep Random Subspace Ensembles. 18	HIT Taxa de acerto. 13
Rede Neural Profunda e Ampla DWNN . 8, 9	HMAE Heterocedasticity Ajustado MAE. 15
Algoritmo Evolutivo EA . 11, 12	Modelo de Markov Oculto HMM . 18
Computação Evolutiva EC . 2	HMRPSO Versão modificada do PSO. 12, 14
EGARCH Exponencial GARCH. 11	HMSE MSE ajustado para heterocedasticidade. 15
ELM Extreme Learning Machine. 8, 9, 19	Taxa de acerto de RH . 10, 15, 16
Média Móvel Exponencial EMA . 12	Índice de Ações HS China Shanghai Shenzhen. 14, 20
EMD2FNN Decomposição e fatoração em modo empírico Rede neural baseada em máquina. 16, 17	Índice HSI Hong Kong Hang Seng. 11, 13, 14, 16, 17
Fundo negociado em bolsa ETF . 14, 16, 17	IBOVESPA Indice Bolsa de Valores de Sao Paulo. 17
Rede Neural Totalmente Conectada FCNN . 20	Coefficiente de Informação IC . 10
FDDR Fuzzy Deep Direct Reinforcement Learning. 11–14	Índice de Informações de RI . 10
Rede neural feedforward FFnn . 6, 9, 11, 14, 15	Índice da Bolsa de Valores de Istambul ISE . 13, 16
Otimização de pesquisa FHSO Firefly Harmony. 11, 15, 16	Índice Composto IXIC NASDAQ. 13
Rede neural de link funcional FLANN . 12	KELM Kernel Extreme Learning Machine. 19
Rede Neural Totalmente Conectada FNN . 4, 12, 16, 17	KL-Divergência Kullback Leibler Divergência. 6
Índice da Bolsa de Valores FTSE London Financial Times. 11, 14-17	KOSPI O Índice de Preços de Ações Composto da Coreia. 8, 9, 11, 13-15
Algoritmo Genético AG . 2, 11, 17, 26	Bolsa de Valores KSE Coreia. 17
Campo Angular GAF Gramian. 16, 17	LAR Preditor de auto-regressão linear. 12
GAN Generative Adversarial Network. 10, 24	Alocação de Dirichlet Latente LDA . 12, 14, 18, 20
GAN-FD GAN para minimizar a perda de erro de previsão e perda de previsão de direção. 9	Dados do Livro de Ordens de Limite LOB . 20
	Sistema de Informação Neurofuzzy Localmente Recorrente do LRNFIS. 11, 13, 16
	Memória de curto prazo LSTM . 1–5, 8–21, 24, 25

Convergência e divergência da média móvel MACD . 17, 19	Análise de Componentes Principais PCA . 9, 10, 13, 19
MAD Desvio Absoluto Médio. 9	Coefficiente de Correlação de PCC Pearson. 16
Erro Absoluto Médio MAE . 9, 10, 13–17, 20	PCD Percentual de Direção Correta. 10
Mapeamento de média móvel MAM . 16, 17	PLR Representação Linear por Partes. 10
Erro de Porcentagem Absoluta Média MAPE . 9, 10, 12–17, 20	Rede Neural Probabilística PNN . 16, 17
MASE Desvio Padrão Médio. 13	Oscilador de preço percentual PPOSC . 12
MC Monte Carlo. 7	Rede PSN Psi-Sigma. 14–16
Coefficiente de Correlação MCC Matthew. 12, 14, 19	Otimização de enxame de partículas PSO . 12, 14
Análise Discriminante Multilinear MDA . 10	R² Correlação quadrada, correlação múltipla de regressão não linear ção. 10
Rebaixamento Máximo MDD . 9	Rede neural de função de base radial RBF . 8, 9, 13, 16, 17
Processo de Decisão MDP Markov. 7	Máquina Boltzmann restrita RBM . 3, 5, 6, 8, 14, 16-19, 23
Informações Mútuas do MI . 9	Link Funcional Computacionalmente Eficiente Recorrente RCEFLANN Rede neural. 12, 14
Aprendizado de máquina de ML . 1–4, 8, 11, 13, 15, 17, 21, 25	RCNN CNN recorrente. 10, 12
MLP Multilayer Perceptron. 5, 8-11, 13-17, 20, 21	Unidade linear retificada ReLU . 3, 6
Modelo MM Markov. 15	RF Floresta Aleatória. 8–10, 12, 13
MODRL Aprendizagem por Reforço Profundo Multiobjetivo. 14	RHE Híbrido Recorrente Elman. 11, 13, 16
MoE Mistura de Especialistas. 11, 13	RL Aprendizagem por reforço. 7, 11-14, 23, 24
Algoritmo Evolutivo Multiobjetivo MOEA . 2	Rede de densidade de mistura recorrente RMDN . 13, 16, 27
Mudança de regime MRS Markov. 12	RMDN-GARCH RMDN com GARCH. 13, 15
Pesquisa manual do MS . 3–8	Erro quadrático médio RMSE . 9, 10, 12-17, 20, 27
Erro quadrático médio do MSE . 6, 9, 10, 12-17	Propagação quadrática média RMSProp . 3–6
Erro de previsão quadrática média MSFE . 13	Erro Relativo da Raiz Média Quadrada RMSRE . 10
Erro de previsão quadrática média do MSPE . 10	Rede Neural Recorrente RNN . 2–5, 8–21, 23–25
NASDAQ National Association of Securities Dealers Automated Cotações. 8–12, 14, 17, 19	RS Pesquisa Aleatória. 3–8
NIFTY Bolsa de Valores Nacional da Índia. 11, 14, 19	Erro ao quadrado relativo do RSE . 10
NIKKEI Tokyo Nikkei Index. 11, 14, 16	Índice de força relativa do RSI . 12, 17
Processamento de Linguagem Natural em PNL . 4, 23-25	RW Caminhada Aleatória. 12, 14, 15
Erro Absoluto Médio Normalizado NMAE . 15, 16	Índice S&P500 Standard's & Poor's 500. 8–17, 19
Erro quadrado médio normalizado NMSE . 9, 15, 16	Índice Composto SCI SSE. 10
Rede Neural NN . 3, 5, 8, 13, 15, 17-19	Autoencoders SDAE Stacked Denoising. 12, 14
norma-RMSE RMSE normalizado. 9	Memória de frequência de estado SFM . 8, 9
NSE National Stock Exchange da Índia. 9	Descida de gradiente estocástico SGD . 3–6
NYMEX Bolsa Mercantil de Nova York. 12, 14	SLP Perceptron de camada única. 17
NYSE Bolsa de Valores de Nova York. 9–11, 17, 19	Erro de Porcentagem Absoluta Média Simétrica SMAPE . 10
OCHL Aberto, Fechado, Alto, Baixo. 10, 11, 13	Otimização Global Baseada em Modelo Sequencial SMBGO . 3, 4
OCHLV Abrir, Fechar, Alto, Baixo, Volume. 8–17, 19	ESPIÃO SPDR S&P 500 ETF. 17
Bolsa de Valores OMX Estocolmo. 11, 13, 14, 16, 17	Razão SR Sharpe. 10, 11, 13, 14, 16

Rede Neural Recorrente Empilhada **SRNN** . 8, 9

SSE Bolsa de Valores de Xangai. 9–11, 14, 17

Composto da Bolsa de Valores de Xangai **SSEC** . 17

Desvio Padrão **STD** . 11, 13, 14

Máquina de vetores de suporte **SVM** . 9, 10, 16, 17, 19

Regressor de vetor de suporte **SVR** . 11, 12, 14, 16, 26

Índice composto da bolsa de valores de **SZSE** Shenzhen. 11, 13, 14, 17

Índice de ações ponderado por capitalização **TAIEX** Taiwan. 11, 14

Pacote da Biblioteca de Análise Técnica **TALIB** . 14, 17

TAQ Comércio e Cotação. 10

TAR Threshold Autoregressive. 11, 13, 16

TD Diferença Temporal. 7

Rede neural de atraso de tempo **TDNN** . 16, 17

TF-IDF Termo Frequência Inversa do Documento. 19

TGRU GRU de dois fluxos. 18

THEIL-U Coeficiente de desigualdade de Theil. 14, 16

TR Retorno Total. 10, 13

Abordagem do Estimador **Parzen** Estruturado em Árvore TSPEA. 3, 4

TUNINDEX Índice do mercado de ações da Tunísia. 13

TWSE Bolsa de Valores de Taiwan. 12, 14

Modelo de correção de erros de vetor **VEC** . 11, 13, 16

Índice de Volatilidade **VIX** S&P500. 11, 13, 15

VXD Dow Jones Industrial Average Volatility Index. 11, 13, 15

Índice de Volatilidade **VXN** NASDAQ100. 11, 13, 15

Taxa de acerto ponderada **WHR** . 15, 16

Faixa Percentual de **William%R** Williams. 12

Regressão de série temporal multicanal ponderada **WMTR** . 10

Transformadas **WT** Wavelet. 14

WTI West Texas Intermediate. 14

Aumento de gradiente **XGBoost** eXtreme. 17

Referências

[1] Ahmet Murat Ozbayoglu, Mehmet Ugur Gudelek, Omer Berat Sezer, Aprendizado profundo para aplicações financeiras: uma pesquisa, 2020, arXiv:2002.05786.

[2] Rafik A. Aliev, Bijan Fazlollahi, Rashad R. Aliev, Soft computing and its applications in business and economics, in: Studies in Fuzziness and Soft Computing, 2004.

[3] Ludmila Dymowa, Soft Computing em Economia e Finanças, Springer Berlin Heidelberg, 2011.

[4] Boris Kovalerchuk, Evgenii Vityaev, Data Mining in Finance: Advances in Relational and Hybrid Methods, Kluwer Academic Publishers, Norwell, MA, EUA, 2000.

[5] Anthony Brabazon, Michael O'Neill (Eds.), Natural Computing in Computational Finance, Springer Berlin Heidelberg, 2008.

[6] Arash Bahrammirzaee, Uma pesquisa comparativa de aplicações de inteligência artificial em finanças: redes neurais artificiais, sistema especialista e sistemas inteligentes híbridos, Computação Neural. Aplic. 19 (8) (2010) 1165-1195.

[7] D. Zhang, L. Zhou, Descobrindo pepitas de ouro: Mineração de dados em aplicação financeira, IEEE Trans. Sistema Homem Cibernético. C 34 (4) (2004) 513-522.

[8] Assunção Mochón, David Quintana, Yago Sáez, Pedro Isasi Viñuela, Soft computing técnicas aplicadas às finanças, Appl. Intel. 29 (2007) 111-115.

[9] Sendhil Mullainathan, Jann Spiess, Aprendizado de máquina: Uma abordagem econométrica aplicada, J. Econ. Perspectiva. 31 (2) (2017) 87-106.

[10] Shu-Heng Chen (Ed.), Algoritmos Genéticos e Programação Genética em Finanças Computacionais, Springer EUA, 2002.

[11] Mãe. Guadalupe Castillo Tapia, Carlos A. Coello Coello, Aplicações de algoritmos evolucionários multiobjetivos em economia e finanças: uma pesquisa, em: 2007 IEEE Congress on Evolutionary Computation, IEEE, 2007.

[12] Antonin Ponsich, Antonio Lopez Jaimes, Carlos A. Coello Coello, Uma pesquisa sobre algoritmos evolutivos multiobjetivos para a solução do problema de otimização de portfólio e outras aplicações financeiras e econômicas, IEEE Trans. Evoluir Computar. 17 (3) (2013) 321-344.

[13] Ruben Aguilar-Rivera, Manuel Valenzuela-Rendon, JJ Rodriguez-Ortiz, algoritmos genéticos e abordagens darwinianas em aplicações financeiras: uma pesquisa, Expert Syst. Aplic. 42 (21) (2015) 7684-7697.

[14] Roy Rada, Sistemas especialistas e computação evolucionária para investimentos financeiros: Uma revisão, Expert Syst. Aplic. 34 (4) (2008) 2232-2240.

[15] Yuhong Li, Weihua Ma, Applications of Artificial Neural Networks in Financial Economics : A survey, in: 2010 International Symposium on Computational Intelligence and Design, IEEE, 2010.

[16] Michal Tkáč, Robert Verner, Redes neurais artificiais nos negócios: Duas décadas de pesquisa, Appl. Computação Suave. 38 (2016) 788-804.

[17] B. Elmsili, B. Outtaj, Aplicações de redes neurais artificiais em economia e pesquisa de gestão: Uma revisão exploratória da literatura, em: 2018 4th International Conference on Optimization and Applications, ICOA, 2018, pp. 1–6.

[18] Marc-André Mittermayer, Gerhard F. Knolmayer, Text Mining Systems for Market Response to News: A Survey, 2006.

[19] Leela Mitra, Gautam Mitra, Applications of news analytics in finance: A review, in: The Handbook of News Analytics in Finance, John Wiley & Sons, Ltd., 2012, pp. 1–39.

[20] Arman Khadjeh Nassirtoussi, Saeed Aghazoborgi, Teh Ying Wah, David Chek Ling Ngo, Mineração de texto para previsão de mercado: Uma revisão sistemática, Expert Syst. Aplic. 41 (16) (2014) 7653-7670.

[21] Colm Kearney, Sha Liu, Sentimento textual em finanças: Uma pesquisa de métodos e modelos, Int. Rev. Finanças. Anal. 33 (2014) 171-185.

[22] B. Shravan Kumar, Vadlamani Ravi, Um levantamento das aplicações de mineração de texto no domínio financeiro, Knowl.-Based Syst. 114 (2016) 128-147.

[23] Frank Z. Xing, Erik Cambria, Roy E. Welsch, Previsão financeira baseada em linguagem natural: uma pesquisa, Artif. Intel. Rev. 50 (1) (2017) 49–73.

[24] Bruce J. Vanstone, Clarence Tan, A survey of the application of soft computing to investment and financial trading, in: Brian C. Lovell, Duncan A. Campbell, Clinton B. Fookes, Anthony J. Maeder (Eds.), Proceedings of the Eighth Australian and New Zealand Intelligent Information Systems Conference, ANZIS 2003, The Australian Pattern Recognition Society, 2003, pp. 211–216, Copyright The Australian Pattern Sociedade de Reconhecimento 2003. Todos os direitos reservados. Permissão garantida.

[25] Ehsan Hajizadeh, H. Davari Ardakani, Jamal Shahrabi, Application of Data Mining Techniques in Stock Markets: A Survey, 2010.

[26] Binoy B. Nair, VP Mohandas, Aplicações de inteligência artificial na previsão financeira – uma pesquisa e alguns resultados empíricos, Intell. Decis. Tecnol. 9 (2) (2014) 99-140.

[27] Rodolfo C. Cavalcante, Rodrigo C. Brasileiro, Victor L.F. Souza, Jarley P. Nobrega, Adriano LI Oliveira, Inteligência computacional e mercados financeiros: Uma pesquisa e direções futuras, Expert Syst. Aplic. 55 (2016) 194-211.

[28] Bjoern Krollner, Bruce J. Vanstone, Gavin R. Finnie, Previsão de tempo financeiro com técnicas de aprendizado de máquina: uma pesquisa, em: ESANN, 2010.

- [29] PD Yoo, MH Kim, T. Jan, Técnicas de aprendizado de máquina e uso de informações de eventos para previsão do mercado de ações: Uma pesquisa e avaliação, em: Conferência Internacional sobre Inteligência Computacional para Modelagem, Controle e Automação e Conferência Internacional sobre Agentes Inteligentes, Tecnologias Web e Comércio na Internet, CIMCA-IAWTIC'06, vol. 2, 2005, pp. 835-841.
- [30] G. Preethi, B. Santhi, técnicas de previsão do mercado de ações: uma pesquisa, J. Teor. Aplic. Inf. Tecnol. 46 (2012) 24-30.
- [31] George S. Atsalakis, Kimon P. Valavanis, Levantamento de técnicas de previsão do mercado de ações – Parte II: Métodos de computação soft, Expert Syst. Aplic. 36 (3) (2009) 5932-5941.
- [32] Amitava Chatterjee, O. Felix Ayadi, Bryan E. Boone, Rede neural artificial e os mercados financeiros: Uma pesquisa, Manag. Finanças 26 (12) (2000) 32–45.
- [33] R. Katarya, A. Mahajan, Uma pesquisa de técnicas de rede neural na análise de tendências de mercado, em: 2017 International Conference on Intelligent Sustainable Systems, ICISS, 2017, pp. 873-877.
- [34] Yong Hu, Kang Liu, Xiangzhou Zhang, Lijun Su, EWT Ngai, Mei Liu, Aplicação de computação evolucionária para descoberta de regras em negociação algorítmica de ações: Uma revisão da literatura, Appl. Computação Suave. 36 (2015) 534-551.
- [35] Wei Huang, KK Lai, Y. Nakamori, Shouyang Wang, Previsão de taxas de câmbio com redes neurais artificiais: Uma revisão, Int. J. Inf. Technol. Determinado. Mak. 03 (01) (2004) 145–165.
- [36] Dadabada Pradeepkumar, Vadlamani Ravi, Híbridos de computação suave para previsão de taxa de câmbio: Uma revisão abrangente, Comput. Operar. Res. 99 (2018) 262-284.
- [37] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Aprendizado profundo, Natureza 521 (7553) (2015) 436-444.
- [38] Jürgen Schmidhuber, Deep learning em redes neurais: uma visão geral, Neural Netw. 61 (2015) 85-117.
- [39] Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016, <http://www.deeplearningbook.org>.
- [40] George Cybenko, Aproximação por superposições de um sigmoidal função, matemática. Sistemas de Sinais de Controle 2 (4) (1989) 303-314.
- [41] Barry L. Kalman, Stan C. Kwasny, Por que tanh: escolhendo uma função sigmoidal, em: [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, vol. 4, IEEE, 1992, pp. 578-581.
- [42] Vinod Nair, Geoffrey E. Hinton, Unidades lineares retificadas melhoram máquinas Boltzmann restritas, em: Proceedings of the 27th International Conference on Machine Learning, ICML-10, 2010, pp. 807–814.
- [43] Andrew L. Maas, Awni Y. Hannun, Andrew Y. Ng, não linearidades de retificadores melhoram modelos acústicos de redes neurais, em: Proc. ICML, vol. 30, 2013, pág. 3.
- [44] Prajit Ramachandran, Barret Zoph, Quoc V. Le, Procurando funções de ativação, 2017, pré-impressão arXiv arXiv:1710.05941.
- [45] Li Deng, Dong Yu, et al., Aprendizado profunda: métodos e aplicações, Encontrado. Processo de Sinalização de Tendências. 7 (3–4) (2014) 197–387.
- [46] Matt W. Gardner, SR Dorling, Redes neurais artificiais (o perceptron multicamada)—uma revisão de aplicações nas ciências atmosféricas, Atmos. Aproximadamente. 32 (14–15) (1998) 2627–2636.
- [47] Herbert Robbins, Sutton Monro, Um método de aproximação estocástica, Ann. Matemática. Estado. (1951) 400-407.
- [48] Ilya Sutskever, James Martens, George Dahl, Geoffrey Hinton, Sobre a importância da inicialização e do impulso no aprendizado profundo, em: Conferência Internacional sobre Aprendizado de Máquina, 2013, pp. 1139–1147.
- [49] John Duchi, Elad Hazan, Yoram Singer, métodos subgradientes adaptativos para aprendizagem online e otimização estocástica, J. Mach. Aprender. Res. 12 (julho) (2011) 2121-2159.
- [50] Tijmen Tieleman, Geoffrey Hinton, Palestra 6.5-RMSProp: Divida o gradiente por uma média corrente de sua magnitude recente, COURSERA: Neural Netw. Mach. Aprender. 4 (2) (2012) 26–31.
- [51] Diederik P. Kingma, Jimmy Ba, Adam: Um método para otimização estocástica, 2014, pré-impressão arXiv arXiv:1412.6980.
- [52] Yoshua Bengio, Patrice Simard, Paolo Frasconi, et al., Aprender dependências de longo prazo com gradiente descendente é difícil, IEEE Trans. Rede Neural. 5 (2) (1994) 157-166.
- [53] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Aprofundando-se em retificadores: superando o desempenho de nível humano na classificação de imagenet, em: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034.
- [54] Xavier Glorot, Yoshua Bengio, Compreendendo a dificuldade de treinar redes neurais de feedforward profundo, em: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010, pp. 249-256.
- [55] James S. Bergstra, Rémi Bardenet, Yoshua Bengio, Balázs Kégl, Algoritmos para otimização de hiperparâmetros, em: Advances in Neural Information Processing Systems, 2011, pp. 2546–2554.
- [56] James Bergstra, Yoshua Bengio, Busca aleatória para otimização de hiperparâmetros, J. Mach. Aprender. Res. 13 (fevereiro) (2012) 281-305.
- [57] Razvan Pascanu, Tomas Mikolov, Yoshua Bengio, Sobre a dificuldade de treinar redes neurais recorrentes, em: Conferência Internacional sobre Aprendizado de Máquina, 2013, pp. 1310-1318.
- [58] Sepp Hochreiter, Jürgen Schmidhuber, Long curto prazo memory, Neural Comput. 9 (8) (1997) 1735-1780.
- [59] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, etc. e tradução automática, 2016, pré-impressão arXiv arXiv:1609.08144.
- [60] Klaus Greff, Rupesh K. Srivastava, Jan Koutník, Bas R. Steunebrink, Jürgen Schmidhuber, LSTM: A search space odyssey, IEEE Trans. Rede Neural. Aprender. Sistema 28 (10) (2016) 2222-2232.
- [61] Nils Reimers, Iryna Gurevych, Optimal hyperparameters for deep lstm networks for sequence labeling tasks, 2017, arXiv preprint arXiv:1707. 06799.
- [62] Shuiwang Ji, Wei Xu, Ming Yang, Kai Yu, redes neurais convolucionais 3D para reconhecimento de ação humana, IEEE Trans. Pattern Anal. Mach. Intel. 35 (1), 221-231.
- [63] Christian Szegedy, Alexander Toshev, Dumitru Erhan, Rede neural profunda funciona para detecção de objetos, em: Advances in Neural Information Processing Systems, 2013, pp. 2553–2561.
- [64] Jonathan Long, Evan Shelhamer, Trevor Darrell, Redes totalmente convolucionais para segmentação semântica, em: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
- [65] Xueheng Qiu, Le Zhang, Ye Ren, P. Suganthan, Gehan Amaratunga, Ensemble deep learning para regressão e previsão de tempo, em: 2014 IEEE Symposium on Computational Intelligence in Ensemble Learning, CIEL, 2014, pp. 1–6.
- [66] Rafael Hrasko, André GC Pacheco, Renato A. Krohling, Previsão de séries temporais usando máquinas de Boltzmann restritas e retropropagação, Procedia Comput. Sci. 55 (2015) 990-999.
- [67] Ruslan Salakhutdinov, Andriy Mnih, Geoffrey Hinton, máquinas Boltzmann restritas para filtragem colaborativa, em: Proceedings of the 24th International Conference on Machine Learning, ACM, 2007, pp. 791–798.
- [68] Yoshua Bengio, Aprendizado profunda de representações para aprendizagem não supervisionada e transferência, em: Anais do Workshop ICML sobre Aprendizagem Não Supervisionada e Transferência, 2012, pp. 17–36.
- [69] Abdel-rahman Mohamed, George Dahl, Geoffrey Hinton, Redes de crenças profundas para reconhecimento de telefone, em: Nips Workshop on Deep Learning for Speech Recognition and Related Applications, vol. 1, Vancouver, Canadá, 2009, pp. 39.
- [70] Honglak Lee, Roger Grosse, Rajesh Ranganath, Andrew Y. Ng, Redes de crenças profundas convolucionais para aprendizado não supervisionado escalável de representações hierárquicas, em: Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 609–616.
- [71] Laurens Van Der Maaten, Aprendendo uma incorporação paramétrica preservando a estrutura local, em: Inteligência Artificial e Estatística, 2009, pp. 384–391.
- [72] Chengwei Yao, Gencai Chen, Adaptação de hiperparâmetros para máquinas Boltzmann restritas baseadas em energia livre, em: 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC, vol. 2, IEEE, 2016, pp. 243-248.
- [73] Miguel A. Carreira-Perpinan, Geoffrey E. Hinton, On contrastive divergence learning., in: Aistats, vol. 10, Citeseer, 2005, pp. 33–40.
- [74] Prasanna Tamilselvan, Pingfeng Wang, Diagnóstico de falha usando classificação de estado de saúde baseada em aprendizado de crenças profundas, Reliab. Eng. Sistema Saf. 115 (2013) 124-135.
- [75] Geoffrey E. Hinton, Simon Osindero, Yee-Whye Teh, Um algoritmo de aprendizado rápido para redes de crenças profundas, Computação Neural. 18 (7) (2006) 1527-1554.
- [76] Qinxue Meng, Daniel Catchpole, David Skillicorn, Paul J. Kennedy, Autoencoder relacional para extração de recursos, em: 2017 International Joint Conference on Neural Networks, IJCNN, IEEE, 2017, pp. 364–371.
- [77] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, Pierre-Antoine Manzagol, Extraído e compondo recursos robustos com denoising autoencoders, em: Proceedings of the 25th International Conference on Machine Learning, ACM, 2008, pp. 1096–1103.
- [78] Richard S. Sutton, Andrew G. Barto, Introduction to Reinforcement Learning, volume 135, MIT press Cambridge, 1998.
- [79] Duy Nguyen-Tuong, Jan Peters, Aprendizado de modelo para controle de robô: uma pesquisa, Cognição Processo. 12 (4) (2011) 319-340.
- [80] Eunsuk Chong, Chulwoo Han, Frank C. Park, Redes de aprendizagem profunda para análise e previsão do mercado de ações: metodologia, representações de dados e estudos de caso, Expert Syst. Aplic. 83 (2017) 187-205.
- [81] Kai Chen, Yi Zhou, Fangyan Dai, Um método baseado em LSTM para previsão de retorno de ações: Um estudo de caso do mercado de ações da China, em: 2015 IEEE International Conference on Big Data, Big Data, IEEE, 2015.
- [82] Eva Dezsi, Ioan Alin Nistor, O aprendizado de máquina profundo pode ser mais esperto que o mercado? Uma comparação entre modelagem econométrica e memória de curto prazo, Rom. Eco. Ônibus. Rev. 11 (4.1) (2016) 54–73.

- [83] AJP Samarawickrama, TGI Fernando, Uma abordagem de rede neural recorrente na previsão de preços diários de ações uma aplicação ao mercado de ações do Sri Lanka , em: 2017 IEEE International Conference on Industrial and Information Systems, ICIIS, IEEE, 2017.
- [84] M. Hiransha, EA Gopalakrishnan, Vijay Krishna Menon, KP Soman, previsão do mercado de ações Nse usando modelos de aprendizado profundo, *Procedia Comput. Sci.* 132 (2018) 1351-1362.
- [85] Sreelekshmy Selvin, R. Vinayakumar, EA Gopalakrishnan, Vijay Krishna Menon, KP Soman, Previsão de preço de ações usando LSTM, rnn e modelo de janela deslizante CNN, em: 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI, IEEE , 2017.
- [86] Sang Il Lee, Seong Joon Yoo, portfólio baseado em limiar: o papel do limiar e suas aplicações, *J. Supercomput.* (2018).
- [87] Xiumin Li, Lin Yang, Fangzheng Xue, Hongjun Zhou, Previsão de série temporal do preço das ações usando redes de crenças profundas com plasticidade intrínseca, em: 2017 29th Chinese Control and Decision Conference, CCDC, IEEE, 2017.
- [88] Lin Chen, Zhilin Qiao, Minggang Wang, Chao Wang, Ruijin Du, Harry Eu gene Stanley, Qual algoritmo de inteligência artificial melhor prevê o mercado de ações chinês? Acesso IEEE 6 (2018) 48625–48633.
- [89] Christopher Krauss, Xuan Anh Do, Nicolas Huck, Redes neurais profundas, árvores impulsionadas por gradiente, florestas aleatórias: arbitragem estatística no S&P 500, *European J. Oper. Res.* 259 (2) (2017) 689-702.
- [90] Rohitash Chandra, Shelvin Chand, Avaliação de arquiteturas de redes neurais co-evolucionárias para previsão de séries temporais com aplicação móvel em finanças, *Appl. Computação Suave.* 49 (2016) 462-473.
- [91] Shuanglong Liu, Chao Zhang, Jinwen Ma, CNN-LSTM Modelo de rede neural para análise de estratégia quantitativa em mercados de ações, em: *Processamento de informações neurais*, Springer International Publishing, 2017, pp. 198–206.
- [92] JB Heaton, NG Polson, JH Witte, *Aprendizado profundo em finanças*, 2016, arxiv: 1602.06561.
- [93] Gilberto Batres-Estrada, *Deep Learning for Multivariate Financial Time Series (tese de mestrado)*, KTH, Mathematical Statistics, 2015.
- [94] Zhaozheng Yuan, Ruixun Zhang, Xiuli Shao, Rede neural profunda e ampla funciona em vários conjuntos de dados temporais com correlação, em: *Proceedings of the 2018 International Conference on Computing and Data Engineering - ICCDE 2018*, ACM Press, 2018.
- [95] Liheng Zhang, Charu Aggarwal, Guo-Jun Qi, Previsão de preço de ações via descoberta de padrões de negociação multifrequência, em: *Anais da 23ª Conferência Internacional ACM SIGKDD sobre Descoberta de Conhecimento e Mineração de Dados - KDD17*, ACM Press, 2017.
- [96] Masaya Abe, Hideki Nakayama, Deep learning para previsão de retornos de ações na seção transversal, em: *Advances in Knowledge Discovery and Data Mining*, Springer International Publishing, 2018, pp. 273–284.
- [97] Guanhao Feng, Jingyu He, Nicholas G. Polson, *Aprendizado profundo para previsão retornos de ativos*, 2018, arxiv:1804.09314.
- [98] Jianqing Fan, Lingzhou Xue, Jiawei Yao, Previsão suficiente usando modelos fatoriais, *SSRN Electron.J.* (2014).
- [99] Mathias Kraus, Stefan Feuerriegel, Suporte à decisão de divulgações financeiras com redes neurais profundas e aprendizado de transferência, *Decis. Sistema de suporte* 104 (2017) 38–48.
- [100] Shotaro Minami, Previsão de preço de ações com eventos de ação corporativa usando LSTM-RNN, *J. Math. Finanças* 08 (01) (2018) 58–63.
- [101] Xiaolin Zhang, Ying Tan, Deep stock ranker: A LSTM neural network model for stock selection, in: *Data Mining and Big Data*, Springer International Publishing, 2018, pp. 614–623.
- [102] Qun Zhuge, Lingyu Xu, Gaowei Zhang, LSTM Neural Network with Emotional Analysis for Prediction of Stock Price, 2017.
- [103] Ryo Akita, Akira Yoshihara, Takashi Matsubara, Kuniaki Uehara, Deep learning para previsão de ações usando informações numéricas e textuais, em: 2016 IEEE/ACIS 15th International Conference on Computer and Information Science, ICIS, IEEE, 2016.
- [104] A. Ozbayoglu, análise técnica baseada em neurais na previsão do mercado de ações, em: *Sistemas de engenharia inteligentes através de redes neurais artificiais*, Volume 17, ASME, 2007, pp. 261-266.
- [105] Kaustubh Khare, Omkar Darekar, Prafull Gupta, VZ Attar, Previsão de preços de ações de curto prazo usando aprendizado profundo, em: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology, RTEICT, IEEE, 2017.
- [106] Xingyu Zhou, Zhisong Pan, Guyu Hu, Siqi Tang, Cheng Zhao, Previsão do mercado de ações em dados de alta frequência usando redes adversárias generativas, *Math. Probl. Eng.* 2018 (2018) 1–11.
- [107] Ritika Singh, Shashi Srivastava, Previsão de ações usando aprendizado profundo, *Multimedia Tools Appl.* 76 (18) (2016) 18569-18584.
- [108] Sercan Karaoglu, Ugur Arpacı, Uma abordagem de aprendizado profundo para otimização da detecção sistemática de sinais em sistemas de negociação financeira com big data, *Int. J. Intel. Sistema Aplic. Eng.* (2017) 31–36, Edição Especial.
- [109] Bo Zhou, Deep learning e a seção transversal de retornos de ações: redes neurais combinando preço e informações fundamentais, *SSRN Electron. J.* (2018).
- [110] Narek Abroyan e, *Redes neurais para classificação de risco do mercado financeiro*, *Frente. Processo de Sinal.* 1 (2) (2017).
- [111] Google, *sistema e método para fundos gerenciados por computador superam o desempenho* Referências, Patente dos EUA, 2016.
- [112] Dat Thanh Tran, Martin Magris, Juho Kannianen, Moncef Gabbouj, Alexandros Iosifidis, representação de tensor em dados financeiros de alta frequência para previsão de mudança de preço, em: 2017 IEEE Symposium Series on Computational Intelligence, SSCI, IEEE, 2017.
- [113] Guanhao Feng, Nicholas G. Polson, Jianeng Xu, Deep factor alpha, 2018, arxiv:1805.01104.
- [114] Xiao Ding, Yue Zhang, Ting Liu, Junwen Duan, Deep learning para previsão de ações orientada a eventos, em: *Proceedings of the 24th International Conference on Artificial Intelligence, IJCAI'15*, AAAI Press, 2015, pp. 2327–2333 .
- [115] Manuel R. Vargas, Beatriz SLP de Lima, Alexandre G. Evsukoff, Deep learning para previsão do mercado de ações a partir de artigos de notícias financeiras, em: 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, CIVEMSA, IEEE , 2017.
- [116] Che-Yu Lee, Von-Wun Soo, Preveja o preço das ações com notícias financeiras com base em redes neurais convolucionais recorrentes, em: 2017 Conference on Technologies and Applications of Artificial Intelligence, TAAI, IEEE, 2017.
- [117] Hitoshi Iwasaki, Ying Chen, Precificação de ativos de sentimento de tópico com aprendizado supervisionado DNN, *SSRN Electron. J.* (2018).
- [118] Sushree Das, Ranjan Kumar Behera, Mukesh Kumar, Santanu Kumar Rath, análise de sentimento em tempo real de dados de streaming do Twitter para previsão de ações, *Procedia Comput. Sci.* 132 (2018) 956-964.
- [119] Jiahong Li, Hui Bu, Junjie Wu, Previsão do mercado de ações sensível ao sentimento: um método de aprendizado profundo, em: 2017 International Conference on Service Systems and Service Management, IEEE, 2017.
- [120] Zhongshengz, Medindo o Índice de Crise Financeira para Alerta de Risco através da Análise de Rede Social, 2018.
- [121] Janderson B. Nascimento, Marco Cristo, The impact of structured event embeddings on scalable stock forecasting models, in: *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web - WebMedia15*, ACM Press, 2015.
- [122] Songqiao Han, Xiaoling Hao, Hailiang Huang, Uma abordagem de extração de eventos para análise de negócios de notícias chinesas online, *Electron. Comer. Res. Aplic.* 28 (2018) 244-260.
- [123] Wei Bao, Jun Yue, Yulei Rao, Uma estrutura de aprendizado profundo para séries temporais financeiras usando autoencoders empilhados e memória de longo prazo, *PLoS One* 12 (7) (2017) e0180944.
- [124] AK Parida, R. Bisoi, PK Dash, Chebyshev funções polinomiais baseadas em sistema de informação neuro-fuzzy recorrente localmente para previsão de dados de mercado financeiro e de energia, *J. Finance Data Sci.* 2 (3) (2016) 202-223.
- [125] Thomas Fischer, Christopher Krauss, *Aprendizado profundo com redes de memória de longo prazo para previsões do mercado financeiro*, *European J. Oper. Res.* 270 (2) (2018) 654-669.
- [126] Philip Wiedegren, Previsão de ativos financeiros baseada em aprendizado profundo, (tese de mestrado), KTH, Estatísticas matemáticas, 2017.
- [127] Anastasia Borovykh, Sander Bohte, Cornelis W. Oosterlee, Redes neurais convolucionais dilatadas para previsão de séries temporais, *J. Comput. Finanças* (2018).
- [128] Khaled A. Althelaya, El-Sayed M. El-Alfy, Salahadin Mohammed, Avaliação do LSTM bidirecional para previsão do mercado de ações de curto e longo prazo , em: 2018 9th International Conference on Information and Communication Systems, ICICS, IEEE, 2018.
- [129] Alexiei Dingli, Karl Sant Fournier, Previsão de séries temporais financeiras – Uma abordagem de aprendizado profundo, *Int. J. Mach. Aprender. Computar.* 7 (5) (2017) 118–122.
- [130] Arit Kumar Rout, PK Dash, Rajashree Dash, Ranjeeta Bisoi, Previsão de séries temporais financeiras usando uma rede neural recorrente de baixa complexidade e abordagem de aprendizado evolutivo, *J. King Saud Univ. Computar. Inf. Sci.* 29 (4) (2017) 536-552.
- [131] Gyeen Jeong, Ha Young Kim, Melhorando as decisões de negociação financeira usando Q-learning profundo: Prevendo o número de ações, estratégias de ação e aprendizado de transferência, *Expert Syst. Aplic.* 117 (2019) 125-138.
- [132] Yujin Baek, Ha Young Kim, Modaguet: Uma nova estrutura de previsão para o valor do índice do mercado de ações com um módulo LSTM de prevenção de overfitting e um módulo LSTM de previsão, *Expert Syst. Aplic.* 113 (2018) 457-480.
- [133] Magnus Hansson, Sobre previsão de retorno de ações com redes LSTM, 2017.
- [134] Aaron Elliot, Cheng Hua Hsu, Previsão de séries temporais: Previsão de ações preço, 2017, arxiv:1710.05751.
- [135] Zhixi Li, Vincent Tam, Combining the real-time wavelet denoising e long-short-term-memory neural network for predizendo índices de ações, em: 2017 IEEE Symposium Series on Computational Intelligence, SSCI, IEEE, 2017.
- [136] Sima Siami-Namini, Akbar Siami Namin, Previsão de séries temporais econômicas e financeiras: ARIMA vs. LSTM, 2018, arxiv:1803.06386.

- [137] Tsung-Jung Hsieh, Hsiao-Fen Hsiao, Wei-Chang Yeh, Previsão de mercados de ações usando transformadas wavelet e redes neurais recorrentes: Um sistema integrado baseado no algoritmo de colônia de abelhas artificial, *Appl. Computação Suave* . 11 (2) (2011) 2510-2525.
- [138] Luna M. Zhang, Redes neurais profundas genéticas usando diferentes funções de ativação para mineração de dados financeiros, em: 2015 IEEE International Conference on Big Data, Big Data, IEEE, 2015.
- [139] Stelios D. Bekiros, modismos irracionais, emulação de memória de curto prazo e previsibilidade de ativos, *Rev. Financial Econ.* 22 (4) (2013) 213-219.
- [140] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, Victor Chang, Uma abordagem inovadora de rede neural para previsão do mercado de ações, *J. Supercomputador*. (2018).
- [141] Yue Deng, Feng Bao, Youyong Kong, Zhiqian Ren, Qionghai Dai, Aprendizado de reforço direto profundo para representação e negociação de sinais financeiros, *IEEE Trans. Neural Netw. Learn. Syst.* 28 (3) (2017) 653-664.
- [142] Bing Yang, Zi-Jia Gong, Wenqi Yang, previsão do índice do mercado de ações usando o conjunto de rede neural profunda, em: 2017 36th Chinese Control Conference, CCC, IEEE, 2017.
- [143] Oussama Lachiheb, Mohamed Salah Gouider, Um projeto de rede neural profunda hierárquica para previsão de retornos de ações, *Procedia Comput. Sci.* 126 (2018) 264-272.
- [144] Bang Xiang Yong, Mohd Rozaini Abdul Rahim, Ahmad Shahidan Abdullah, A stock market trading system using deep neural network, in: *Communications in Computer and Information Science*, Springer Singapore, 2017, pp. 356-364.
- [145] Serdar Yümlü, Fikret S. Gürgen, Nesrin Okay, Uma comparação de modelos neurais globais, recorrentes e suavizados por partes para a previsão da bolsa de valores de Istambul (ISE), *Pattern Recognit. Lett.* 26 (13) (2005) 2093-2103.
- [146] Hongju Yan, Hongbing Ouyang, Previsão de séries temporais financeiras com base em aprendizado profundo, *Wirel. Pers. Comun.* 102 (2) (2017) 683-700.
- [147] Takahashi, Long memory and predictability in financial markets, in: *Annual Conference of the Japanese Society for Artificial Intelligence*, 2017, http://dx.doi.org/10.11517/pjsai.JSAI2017.0_2D3OS19a1.
- [148] Melike Bildirici, Elçin A. Alp, Özgür Ö. Ersin, modelo de rede neural de cointegração TAR : Uma análise empírica de taxas de câmbio e retornos de ações, *Expert Syst. Aplic.* 37 (1) (2010) 2-11.
- [149] Ioannis Psaradellis, Georgios Sermpinis, Modelagem e negociação dos índices de volatilidade implícita dos EUA. Evidência dos índices VIX, VXN e VXD, *Int. J. Previsão*. 32 (4) (2016) 1268-1283.
- [150] Jiaqi Chen, Wenbo Wu, Michael Tindall, Hedge Fund Return Prediction and Fund Selection: A Machine Learning Approach, Occasional Paper 16-4, Federal Reserve Bank of Dallas, 2016.
- [151] Marios Mourelatos, Christos Alexakos, Thomas Amorgianiotis, Spiridon Likothanassis, modelagem e negociação de índices financeiros utilizando técnicas de aprendizado profundo: O caso de uso de grande capitalização ATHENS SE FTSE/ASE, em: 2018 Innovations in Intelligent Systems and Applications, INISTA, IEEE, 2018.
- [152] Yuzhou Chen, Junji Wu, Hui Bu, Incorporação e previsão do mercado de ações: Um método de aprendizado profundo, em: 2018 15th International Conference on Service Systems and Service Management, ICSSSM, IEEE, 2018.
- [153] Weiyu Si, Jinke Li, Peng Ding, Ruonan Rao, Uma abordagem de aprendizado de reforço profundo multiobjetivo para negociação intradiária de futuros de índices de ações, em: 2017 10th International Symposium on Computational Intelligence and Design, ISCID, IEEE, 2017.
- [154] Weiling Chen, Chai Kiat Yeo, Chiew Tong Lau, Bu Sung Lee, Aproveitando as notícias de mídia social para prever o movimento do índice de ações usando RNN-boost, *Data Knowl. Eng.* (2018).
- [155] Matthew Francis Dixon, Diego Klabjan, Jin Hoon Bang, Previsão de mercados financeiros baseada em classificação usando redes neurais profundas, *SSRN Electron. J.* (2016).
- [156] Fernando Sánchez Lasheras, Francisco Javier de Cos Juez, Ana Suárez Sánchez, Alicia Krzemieć, Pedro Riesgo Fernández, Previsão do preço à vista do cobre COMEX por meio de redes neurais e modelos ARIMA, *Resour. Política* 45 (2015) 37-43.
- [157] Yang Zhao, Jianping Li, Lean Yu, Uma abordagem de conjunto de aprendizado profundo para previsão de preços de petróleo bruto, *Energy Econ.* 66 (2017) 9-16.
- [158] Yanhui Chen, Kaijian He, Geoffrey KF Tso, Previsão de preços do petróleo bruto: um modelo baseado em aprendizado profundo, *Procedia Comput. Sci.* 122 (2017) 300-307.
- [159] Jonathan Doering, Michael Fairbank, Sheri Markose, Redes neurais convolucionais aplicadas à previsão de microestrutura de mercado de alta frequência, em: 2017 9th Computer Science and Electronic Engineering, CEEC, IEEE, 2017.
- [160] P. Tino, C. Schittenkopf, G. Dorffner, negociação de volatilidade financeira usando redes neurais recorrentes, *IEEE Trans. Rede Neural.* 12 (4) (2001) 865-874.
- [161] Ruoxuan Xiong, Eric P. Nichols, Yuan Shen, Deep learning stock volatility with Google domestic trends, 2015, [arxiv:1512.04916](https://arxiv.org/abs/1512.04916).
- [162] Yu-Long Zhou, Ren-Jie Han, Qian Xu, Wei-Ke Zhang, redes de memória de longo prazo para previsão de volatilidade CSI300 com volume de pesquisa baidu, 2018, [arxiv:1805.11954](https://arxiv.org/abs/1805.11954).
- [163] Ha Young Kim, Chang Hyun Won, Previsão da volatilidade do índice de preços de ações: Um modelo híbrido integrando LSTM com vários modelos do tipo GARCH, *Expert Syst. Aplic.* 103 (2018) 25-37.
- [164] Nikolay Nikolaev, Peter Tino, Evgueni Smirnov, Aprendizado de variância de série dependente do tempo com redes de densidade de mistura recorrente, *Neurocomputing* 122 (2013) 501-512.
- [165] Campbell R. Harvey, Previsões de crescimento econômico dos mercados de títulos e ações, *Finanças. Anal. J.* 45 (5) (1989) 38-45.
- [166] Daniele Bianchi, Matthias Büchner, Andrea Tamoni, prêmios de risco de títulos com aprendizado de máquina, *SSRN Electron. J.* (2018).
- [167] Venketas Warren, tamanho do mercado Forex: uma vantagem dos traders, 2019.
- [168] Ren Zhang, Furao Shen, Jinxi Zhao, Um modelo com granulação difusa e redes de crença profunda para previsão de taxas de câmbio, em: 2014 International Joint Conference on Neural Networks, IJCNN, IEEE, 2014.
- [169] Jing Chao, Furao Shen, Jinxi Zhao, Previsão de taxa de câmbio com redes de crença profunda, em: The 2011 International Joint Conference on Neural Networks, IEEE, 2011.
- [170] Jing Zheng, Xiao Fu, Guijun Zhang, Pesquisa sobre previsão cambial baseado na rede de crenças profundas, *Neural Comput. Aplic.* (2017).
- [171] Furao Shen, Jing Chao, Jinxi Zhao, Previsão da taxa de câmbio usando redes de crenças profundas e método de gradiente conjugado, *Neurocomputing* 167 (2015) 243-253.
- [172] Hua Shen, Xun Liang, Um modelo de previsão de séries temporais baseado em algoritmo integrado de aprendizado profundo com autoencoders empilhados e SVR para previsão de FX, em: ICANN, 2016.
- [173] Georgios Sermpinis, Jason Laws, Andreas Karathanasopoulos, Christian L. Dunis, Previsão e negociação da taxa de câmbio EUR/USD com expressão gênica e redes neurais PSI sigma, *Expert Syst. Aplic.* 39 (10) (2012) 8865-8877.
- [174] Georgios Sermpinis, Christian Dunis, Jason Laws, Charalampos Stasinakis, Previsão e negociação da taxa de câmbio EUR/USD com combinação de rede neural estocástica e alavancagem variável no tempo, *Decis. Sistema de suporte* 54 (1) (2012) 316-329.
- [175] Georgios Sermpinis, Charalampos Stasinakis, Christian Dunis, combinações de redes neurais estocásticas e genéticas na negociação e efeitos de alavancagem de variação de tempo híbridos, *J. Int. Financiar Marca. Inst. Dinheiro* 30 (2014) 21-54.
- [176] SUN Bo, XIE Chi, previsão da taxa de câmbio RMB no contexto da crise financeira, *Syst. Eng. Teoria Pratica.* 29 (12) (2009) 53-64.
- [177] Nijole Maknickienė e Algirdas Maknickas, Sistema de previsão do mercado financeiro com rede neural evoluto e método dephli, *J. Bus. Eco. Gerenciar* 14 (2) (2013) 403-413.
- [178] Nijole Maknickiene, Aleksandras Vytautas Rutkauskas, Algirdas Maknickas, Investigation of Financial Market Prediction by Recurrent Neural Network, 2014.
- [179] Luca Di Persio, Oleksandr Honchar, Abordagem de redes neurais artificiais para a previsão de movimentos de preços do mercado de ações, *Int. J. Eco. Gerenciar Sistema* (1) (2016) 158-162.
- [180] Jerzy Korczak, Marcin Hernes, Deep learning para previsão de séries temporais financeiras no sistema a-trader, em: Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, IEEE, 2017.
- [181] Gonçalo Duarte Lima Freire Lopes, Deep Learning for Market Forecasts, 2018.
- [182] Sean McNally, Jason Roche, Simon Caton, Predicting the price of bitcoin using machine learning, in: 2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing, PDP, IEEE, 2018.
- [183] Sanjiv Das, Karthik Mokashi, Robbie Culkin, Os mercados são realmente eficientes? Experimentos usando algoritmos de aprendizado profundo para previsão de movimento do mercado, *Algoritmos* 11 (9) (2018) 138.
- [184] Ariel Navon, Yosi Keller, Previsão de séries temporais financeiras usando aprendizado profundo, 2017, [arxiv:1711.04174](https://arxiv.org/abs/1711.04174).
- [185] EW Saad, DV Prokhorov, DC Wunsch, Estudo comparativo de previsão de tendência de ações usando atraso de tempo, redes neurais recorrentes e probabilísticas, *IEEE Trans. Rede Neural.* 9 (6) (1998) 1456-1470.
- [186] Luca Di Persio, Oleksandr Honchar, Abordagem de redes neurais recorrentes para a previsão financeira de ativos do Google, *Int. J. Matemática. Computar. Simul.* 11 (2017) 713.
- [187] Guizhu Shen, Qingping Tan, Haoyu Zhang, Ping Zeng, Jianjun Xu, Aprendizado profundo com redes de unidades recorrentes fechadas para previsões de sequências financeiras, *Procedia Comput. Sci.* 131 (2018) 895-903.
- [188] Jou-Fan Chen, Wei-Lun Chen, Chun-Ping Huang, Szu-Hao Huang, An-Pin Chen, Análise de dados de séries temporais financeiras usando redes neurais convolucionais profundas, em: 2016 7th International Conference on Cloud Computing and Big Data, CCBDD, IEEE, 2016.
- [189] Omer Berat Sezer, Ahmet Murat Ozbayoglu, Financial trading model with stock bar chart image series with deep convolutional neural networks, 2019, [arXiv:1903.04610](https://arxiv.org/abs/1903.04610).

- [190] Feng Zhou, Hao min Zhou, Zhihua Yang, Lihua Yang, EMD2FNN: Uma estratégia que combina decomposição de modo empírico e rede neural baseada em máquina de fatoração para previsão de tendências do mercado de ações, *Expert Syst. Aplic.* 115 (2019) 136-151.
- [191] Kristina Ausmees, Slobodan Milovanovic, Fredrik Wrede, Afshin Zafari, *Domar redes de crenças profundas*, 2017.
- [192] Kamran Raza, Previsão do desempenho do mercado de ações usando técnicas de aprendizado de máquina, em: 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies, ICIEECT, IEEE, 2017.
- [193] Omer Berat Sezer, Murat Ozbayoglu, Erdogan Dogdu, Um sistema de negociação de ações baseado em rede neural profunda baseado em parâmetros de análise técnica otimizados evolutivos, *Procedia Comput. Sci.* 114 (2017) 473-480.
- [194] Qiubin Liang, Wenge Rong, Jiayi Zhang, Jingshuang Liu, Zhang Xiong, previsão de tendência do mercado de ações baseado em máquina restrita Boltzmann, em: 2017 International Joint Conference on Neural Networks, IJCNN, IEEE, 2017.
- [195] Luigi Troiano, Elena Mejuto Villa, Vincenzo Loia, Replicando uma estratégia de negociação por meio de LSTM para aplicações do setor financeiro, *IEEE Trans. Ind. Inf.* 14 (7) (2018) 3226-3234.
- [196] David MQ Nelson, Adriano CM Pereira, Renato A. de Oliveira, Previsão do movimento dos preços dos mercados de ações com redes neurais LSTM, em: 2017 International Joint Conference on Neural Networks, IJCNN, IEEE, 2017.
- [197] Yuan Song, Yingnian Wu, Previsão de tendência de ações: com base na máquina Métodos de Aprendizagem (Tese de Mestrado), 2018.
- [198] M. Ugur Gudelek, S. Arda Boluk, A. Murat Ozbayoglu, A deep learning based stock trading model with 2-D CNN trend detection, em: 2017 IEEE Symposium Series on Computational Intelligence, SSCI, IEEE, 2017.
- [199] Omer Berat Sezer, Ahmet Murat Ozbayoglu, negociação financeira algorítmica com redes neurais convolucionais profundas: abordagem de conversão de séries temporais para imagens, *Appl. Computação Suave*. 70 (2018) 525-538.
- [200] Hakan Gunduz, Yusuf Yaslari, Zehra Cataltepe, previsão Intraday de Borsa Istanbul usando redes neurais convolucionais e correlações de recursos, *Knowl.-Based Syst.* 137 (2017) 138-148.
- [201] Yifu Huang, Kai Huang, Yang Wang, Hao Zhang, Jihong Guan, Shuigeng Zhou, Explorando o humor do Twitter para impulsionar a previsão de tendências financeiras com base em modelos de rede profunda, em: *Intelligent Computing Methodologies*, Springer International Publishing, 2016, pp. 449-460.
- [202] Yangtuo Peng, Hui Jiang, Alavancar notícias financeiras para prever movimentos de preços de ações usando incorporação de palavras e redes neurais profundas, em: *Anais da Conferência de 2016 do Capítulo Norte-Americano da Associação de Linguística Computacional: Tecnologias de Linguagem Humana, Associação para Linguística Computacional*, 2016.
- [203] Huy D. Huynh, L. Minh Dang, Duc Duong, Um novo modelo para previsão de movimentos de preços de ações usando rede neural profunda, em: *Proceedings of the Eighth International Symposium on Information and Communication Technology - SoICT 2017*, ACM Press, 2017.
- [204] L. Minh Dang, Abolghasem Sadeghi-Niaraki, Huy D. Huynh, Kyungbok Min, Hyeonjoon Moon, Abordagem de aprendizado profundo para previsão de tendências de estoque de curto prazo com base em rede de unidade recorrente fechada de dois fluxos, *Acesso IEEE* (2018) 1.
- [205] Ishan Verma, Lipika Dey, Hardik Meisheri, Detectando, quantificando e acessando o impacto de eventos de notícias nos índices de ações indianos, em: *Proceedings of the International Conference on Web Intelligence - WI17*, ACM Press, 2017.
- [206] Leonardo dos Santos Pinheiro, Mark Dras, Previsão do mercado de ações com aprendizado profundo: Um modelo de linguagem neural baseado em caracteres para negociação baseada em eventos, em: *Proceedings of the Australasian Language Technology Association Workshop 2017*, 2017, pp. 6-15.
- [207] Jordan Prosky, Xingyou Song, Andrew Tan, Michael Zhao, Previsibilidade de sentimento para ações, *CoRR abs/1712.05785* (2017).
- [208] Yang Liu, Qingguo Zeng, Huanrui Yang, Adrian Carrio, Previsão do movimento dos preços das ações a partir de notícias financeiras com aprendizado profundo e incorporação de gráficos de conhecimento, em: *Gestão de Conhecimento e Aquisição de Sistemas Inteligentes*, Springer International Publishing, 2018, pp. 102-113.
- [209] Akira Yoshihara, Kazuki Fujikawa, Kazuhiro Seki, Kuniaki Uehara, Predizendo tendências do mercado de ações por redes neurais profundas recorrentes, em: *Lecture Notes in Computer Science*, Springer International Publishing, 2014, pp. 759-769.
- [210] Lei Shi, Zhiyang Teng, Le Wang, Yue Zhang, Alexander Binder, DeepClue: Interpretação visual da previsão de estoque profunda baseada em texto, *IEEE Trans. Conhecimento Eng. de Dados* (2018) 1.
- [211] Xi Zhang, Yunjia Zhang, Senzhang Wang, Yuntao Yao, Binxing Fang, Philip S. Yu, Melhorando a previsão do mercado de ações via fusão de informações heterogêneas, *Sistema baseado em conhecimento* 143 (2018) 236-247.
- [212] Ziniu Hu, Weiqing Liu, Jiang Bian, Xuanzhe Liu, Tie-Yan Liu, Ouvindo sussurros caóticos: Uma estrutura de aprendizado profundo para previsão de tendências de ações orientada para notícias, em: *Anais da Décima Primeira Conferência Internacional da ACM sobre Pesquisa na Web e Mineração de Dados, WSDM '18*, ACM, Nova York, NY, EUA, 2018, pp. 261-269.
- [213] Qili Wang, Wei Xu, Han Zheng, Combinando a sabedoria das multidões e a análise técnica para a previsão do mercado financeiro usando conjuntos de subespaços aleatórios profundos, *Neurocomputing* 299 (2018) 51-61.
- [214] Takashi Matsubara, Ryo Akita, Kuniaki Uehara, Previsão de preço de ações por modelo generativo neural profundo de artigos de notícias, *IEICE Trans. Inf. Sistema* E101.D (4) (2018) 901-908.
- [215] Xiaodong Li, Jingjing Cao, Zhaoqing Pan, Análise de impacto de mercado por meio de arquiteturas de aprendizado profundo, *Computação neural. Aplic.* (2018).
- [216] Avraam Tsantekidis, Nikolaos Passalis, Anastasios Tefas, Juho Kannianen, Moncef Gabbouj, Alexandros Iosifidis, Using Deep Learning to Detect Price Change Indications in Financial Markets, in: 2017 25th European Signal Processing Conference, EUSIPCO, IEEE, 2017.
- [217] Justin Sirignano, Rama Cont, Características universais da formação de preços nos mercados financeiros: perspectivas do aprendizado profundo, *SSRN Electron. J.* (2018).
- [218] Przemyslaw Buczkowski, Predizendo tendências de ações com base em recomendações de especialistas usando redes neurais gru/LSTM, em: *Lecture Notes in Computer Science*, Springer International Publishing, 2017, pp. 708-717.
- [219] Avraam Tsantekidis, Nikolaos Passalis, Anastasios Tefas, Juho Kannianen, Moncef Gabbouj, Alexandros Iosifidis, Forecasting stock prices from the limit order book using convolutional neural networks, em: 2017 IEEE 19th Conference on Business Informatics, CBI, IEEE, 2017.
- [220] Thomas G. Thomas Günter Fischer, Christopher Krauss, Alexander Deint, arbitragem estatística em mercados de criptomoedas, *J. Risk Financial Manag.* 12 (2019).
- [221] T. Gneiting, AE Raftery, regras de pontuação estritamente adequadas, previsão e estimativa, *J. Amer. Estatista. Associação* 102 (477) (2007) 359-378.
- [222] Hannah R. Rothstein, Alex J. Sutton, Michael Borenstein, Publication Bias in Meta-Analysis – Prevention, Assessment and Adjustment, John Wiley & Sons, Ltd, 2005.



Omer Berat Sezer recebeu o diploma de bacharel em Engenharia Elétrica e Eletrônica (EEE) pela Middle East Technical University (METU) NCC, Ancara, em 2009; Mestrado em EEE pela METU, Ancara, em 2013 com ênfase em redes de computadores. Ele é formado em Engenharia da Computação pela TOBB ETU com Ph.D. licenciatura, Ancara, em 2018, com ênfase em machine learning e deep learning. Ele trabalhou como engenheiro de software e pesquisador no TUBITAK Space Technologies Research Institute, em Ancara, por dez anos.



Ele também trabalha agora como engenheiro de software no setor automotivo na Alemanha. Seus interesses de pesquisa são aprendizado de máquina, inteligência artificial, análise de dados de séries temporais, inteligência computacional, visão de máquina e engenharia de software.



Mehmet Ugur Gudelek recebeu bacharelado em Engenharia Elétrica pela Middle East Technical University (METU), Ancara, Turquia, em 2015 com ênfase em computadores e mestrado em Engenharia de Computação pela TOBB University of Economics and Technology (TOBB ETU), Ancara, Turquia, em 2019. Atualmente é Doutor em Engenharia da Computação. candidato na TOBB ETU.

Ele também é assistente de ensino e pesquisa no TOBB ETU. Seus interesses de pesquisa são aprendizado de máquina, aprendizado profundo e análise de dados de séries temporais.



Ahmet Murat Ozbayoglu formou-se no Departamento de Engenharia Elétrica da METU, Ankara, Turquia, em 1991. Ele obteve seu mestrado e doutorado em Engenharia de Sistemas na Universidade de Ciência e Tecnologia do Missouri em 1993 e 1996, respectivamente. Depois, trabalhou na Beyond Inc, St. Peters, MO como engenheiro e consultor de desenvolvimento de produtos. Em 2001, ingressou na MEMC Electronic Materials Inc., como engenheiro de software. Em 2005, ingressou no Departamento de Engenharia da Computação da TOBB University of Economics and Technology, Ancara.

Seus interesses de pesquisa incluem aprendizado de máquina, reconhecimento de padrões, big data, negociação algorítmica, inteligência computacional, visão de máquina.