



XXXI Congresso de  
Iniciação Científica  
----- Unicamp

2023



---

# ***Deep learning* aplicado à previsão de índices da bolsa de valores brasileira**

---

**Palavras chave:** Redes Neurais, Mercado Financeiro, Previsões de Séries Temporais.

**Bolsista:** Othavio Henrique de Jesus Ayres[IMECC-UNICAMP]

**Orientador:** JOÃO FLORINDO[IMECC-UNICAMP]

**Co-Orientadora:** SAMARA KIIHL[IMECC-UNICAMP]

**Palavras chave:** Redes Neurais, Mercado Financeiro, Previsões de Séries Temporais.

## **1 Introdução**

O sucesso de uma estratégia de investimento no mercado de ações negociadas em bolsas de valores é determinado pela sequência de decisões tomadas pelo investidor. Por isso, os investidores estão sempre atentos às cotações de preços, notícias, cenário político e econômico com o objetivo de detectar padrões que os permitam tomar as melhores decisões em cada situação. Sabe-se que atualmente grande parte das negociações que ocorrem nas principais bolsas de valores do mundo são executadas por sistemas automatizados de negociação (*Financial Trading Systems* em inglês) também chamados de robôs ou agentes de negociação. Esses sistemas utilizam técnicas de Inteligência Artificial para detectar padrões ocultos em tempo real a partir de dados de preços, volume, notícias e outras informações. Uma vez detectado um determinado padrão pelo agente, o sistema executa automaticamente a operação mais adequada para os objetivos do investimento naquele momento. Esses sistemas geralmente operam alta frequência e disputam por negócios que ofereçam ganhos muito pequenos. Porém, se essas oportunidades de negócios forem numerosas o suficiente eles podem acumular muitos ganhos [4]. Uma vantagem des-

ses sistemas é que eles podem detectar padrões e executar ordens em frações de segundo possibilitando ao investidor um melhor aproveitamento das oportunidades de negócios no mercado de bolsa de valores. Uma das abordagens em Inteligência Artificial utilizadas é o Aprendizado supervisionado (*Supervised Learning* em inglês). O modo de aprendizado supervisionado depende de dados de entrada  $\mathbf{x}^i$  junto com as suas respectivas saídas  $\mathbf{t}^i$  formando o chamado conjunto de treinamento:  $\{(\mathbf{x}^i, \mathbf{t}^i) : i = 1, 2, 3, \dots, p\}$ . Estes pares representam os dados de treinamento. Em problemas de classificação, como nos exemplos na área médica [1] as saídas são os rótulos das classes. O princípio básico deste modo de aprendizado consiste em treinar os pesos da rede tal que o erro gerado pelas diferenças entre as saídas produzidas a partir destas entradas e as saídas desejadas seja minimizado. Além disso, é importante garantir uma certa capacidade de generalização, que diz respeito à capacidade de reconhecer uma estrutura nos dados e produzir saídas adequadas para novos dados de entrada.

## 2 Metodologia

O erro da predição será calculado utilizando três métricas distintas: erro absoluto médio percentual (MAPE), erro quadrático médio (MSE) e erro médio absoluto (MAE). As equações (1), (2) e (3) indicam o valor de MAPE, MSE e MAE:

$$MAPE = \frac{1}{n} \sum_{t=1} \hat{n}(g_t - p_t)^2 \quad (1)$$

$$MSE = \frac{1}{n} \sum_{t=1} \hat{n} \left| 1 - \frac{g_t}{p_t} \right| \quad (2)$$

$$MAE = \frac{1}{n} \sum_{t=1} \hat{n}(g_t - p_t) \quad (3)$$

em que  $n$  é o número do  $n$ -ésimo dado de entrada e  $g_t$  e  $p_t$  são os valores real e previsto, respectivamente, no instante  $t$ .

### 2.1 Dados

Como base de dados utilizaremos os principais índices da Bolsa de Valores de São Paulo (BM&FBOVESPA) que possuem valores entre o período de 1 de junho de 2021 até 1 de junho de 2023 (Com cerca de 2721 dias úteis entre estas duas datas). A BM&FBOVESPA reúne e disponibiliza uma série de informações sobre os negócios realizados em cada sessão em que se efetuam negócios com ativos listados, por sistema de negociação eletrônica, os chamados preg e entre elas estão os índices que mostram o comportamento de todo o mercado ou de segmentos específicos de mercado. Entre os índices os principais são:

- Ibovespa : Este índice é o valor atual em moeda corrente de uma carteira de ações,

e tem como propósito refletir não apenas as variações dos preços das ações, mas também o impacto da distribuição dos proventos, sendo considerado um indicador que avalia o retorno total de suas ações componentes.

- IBrX - Índice Brasil é um índice de preços que mede o retorno de uma carteira teórica composta por 100 ações selecionadas entre as mais negociadas na BM&FBOVESPA, em termos de número de negócios e volume financeiro. Essas ações são ponderadas na carteira do índice pelo seu respectivo número de ações disponíveis à negociação no mercado.

## 2.2 Modelos

### 2.2.1 LSTM

Redes neurais recorrentes (RNN, do inglês *Recurrent Neural Network*) são usadas para aprender padrões temporais de dados sequenciais, como dados de séries temporais. A RNN básica tem o chamado “desaparecimento do gradiente” como um problema para sequências de entrada longas. A LSTM foi proposta para resolver este problema e aprender melhor dependências de longo prazo. O modelo de Redes de Memórias de Curto e Longo Prazo (LSTM, do inglês *Long Short Term Memory*) adiciona portas e diferentes blocos de memória chamados células, indo assim além de um modelo RNN simples, permitindo maior retenção de padrões da sequência do que a arquitetura RNN padrão. A LSTM é uma arquitetura de rede recorrente que “lembra” valores em intervalos arbitrários. Em específico, a LSTM tem como objetivo classificar e analisar séries temporais que possuem um intervalo de tempo de duração desconhecida. As LSTMs também possuem uma estrutura de cadeia, mas o módulo de repetição possui uma estrutura diferente. Em vez de ter uma única camada de rede neural, existem quatro, interagindo de uma maneira alternativa.

### 2.2.2 CONVLSTM

As Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Network*) são uma classe de redes neurais profundas que possuem uma arquitetura focada na análise de dados multidimensionais e rápido aprendizado. Por conta disto, as CNNs são utilizadas em diversos modelos de reconhecimento de imagens e vídeos. As CNNs são versões alteradas dos perceptrons multicamadas, que buscam organizar padrões presentes nos dados de forma crescente em termos de complexidade. As CNNs possuem tal nome pois utilizam uma operação chamada convolução para tratar os dados de entrada. De maneira geral, uma CNN simples é formada por três tipos de camadas. O primeiro tipo corresponde à camada convolucional, que trata os dados na forma de tensor e os entrega na forma de *feature maps* (mapas de recursos ou características) para a próxima camada. O segundo é a camada de *pooling*, onde as dimensões dos dados são reduzidas por meio da introdução de uma operação de agrupamento. Essa camada recebe cada saída do mapa de recursos da

camada convolucional e prepara um mapa de características condensadas. Por fim, temos uma camada totalmente conectada que realiza a vetorização dos dados e sua predição final. A arquitetura ConvLSTM foi recentemente introduzida para formular a previsão de séries temporais como um problema de previsão de sequência espaço-temporal no qual tanto a entrada quanto a previsão alvo são sequências espaço-temporais. O ConvLSTM é conhecido por capturar padrões espaço-temporais de conjuntos de dados sequenciais de grande escala. A arquitetura do ConvLSTM é similar à do LSTM pois também é uma camada recorrente, mas todas as operações de multiplicação interna de matrizes que aconteceria em uma LSTM são substituídas por convoluções. Isto faz com que os dados que fluem por cada célula estejam na forma de uma imagem (3D), e não apenas como um vetor 1D, como em uma simples LSTM.

### 3 Resultados obtidos

Para validação dos modelos testamos ambos os modelos em condições diversas

Figura 1: Tabela com as métricas de erro da CONVLSTM

Medida de erro	Erro de treino	Erro de teste	Nº de iterações
MSE	0.03509	0.17872	300
MAE	0.11209	0.0579	300
MAPE	1.27%	1.00%	300
MSE	0.01323	0.00244	900
MAE	0.06759	0.16982	900
MAPE	1.46%	0.99%	900

Figura 2: Tabela com as métricas de erro da LSTM

Medida de erro	Erro de treino	Erro de teste	Nº de iterações
MSE	0.00528	0.00228	300
MAE	0.07561	3.02949	300
MAPE	1.61%	0.06%	300
MSE	0.00479	0.00266	900
MAE	0.04553	0.0647	900
MAPE	1.14%	0.08%	900

Nota-se que pelo grau de complexidade de cada modelo a arquitetura CONVLSTM possui mais parâmetros para serem treinados:

Figura 3: Tabela de Complexidade de cada modelo

Modelo	Nº Parametros (10 <sup>3</sup> )	MACs (10 <sup>6</sup> )	Floaps(10 <sup>6</sup> )
LSTM(100->50 Days)	6.898	23.45	46.9
CONVLSTM(100->50 Days)	232	208.205	416.41

## 4 Conclusão

Em geral, a arquitetura CONVLSTM mostrou-se capaz de alcançar o desempenho do modelo utilizando da arquitetura LSTM, o que indica que a arquitetura CONVLSTM tem potencial para ser aplicada em problemas de predição de valores de ações. Entretanto, vale ressaltar que o custo computacional da arquitetura CONVLSTM é consistentemente mais alto que o custo computacional da arquitetura LSTM, o que afeta a duração de seu treinamento de maneira direta, portanto para casos práticos a arquitetura LSTM é uma escolha melhor. Todavia, vale ressaltar que o mercado de ações possui valores bastante caóticos quando analisado em curtos intervalos de tempo, e nota-se a influência desta imprevisibilidade dentro de ambos os modelos. Esses resultados sugerem o desenvolvimento de estudos futuros aprofundando as possíveis interpretações dos dados de entrada para utilização na CONVLSTM, ou até mesmo informações externas que poderão ser acrescentadas neste novo formato de entrada que podem resultar melhores resultados para predição.

## Referências

- [1] Aristotelis Tsirigos. Nicolas Coudray. *Deep learning links histology, molecular signatures and prognosis in cancer*. *Nat Cancer*, 1:755–757, 27-July 2020.
- [2] Alexandre Xavier. *An introduction to CONVLstm*. *NeuronioBR*, 25-março 2019.
- [3] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, 31(7):1235–1270, 07 2019.
- [4] Ha Young Kim Si Woon Lee. *Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation*. *Elsevier*, 161(113704):01–20, 8-Julho 2020.
- [5] Alan Marcus Zvi Bodie, Alex Kane. *Investments*. Irwin/McGraw-Hill, 10th edition, 9-Setembro 2013.
- [6] Uday Pratap Singh Gourav Kumar, Sanjeev Jain. *Stock Market Forecasting Using Computational Intelligence: A Survey*. *EpringerLink*, 28(11831):1886–1784, February-15 2020.
- [7] Governo Brasileiro. Índices representativos do mercado de ações, Abril 2022.
- [8] Jiahao Su, Wonmin Byeon, Furong Huang, Jan Kautz, and Animashree Anandkumar. Convolutional tensor-train LSTM for spatio-temporal learning. *CoRR*, abs/2002.09131, 2020.
- [9] Xingjian SHI, Zhouong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun WOO. Convolutional lstm network: A machine learning approach for precipitation nowcasting. Curran Associates, Inc., 2015.