

Previsão do preço futuro do óleo de girassol: uma abordagem baseada no modelo CNN-Attention

Forecasting the future price of sunflower oil: an approach based on the CNN-Attention model

Leandro de Oliveira¹, José Airton Azevedo dos Santos²

RESUMO

O Mecanismo de Atenção (*Attention Mechanism*) tem sido, nos últimos anos, muito utilizado em diversas áreas de aplicação de aprendizagem profunda (*Deep Learning*). Neste contexto, este trabalho tem como objetivo propor a utilização da rede CNN-Attention para previsão do preço do óleo de girassol. A base de dados apresenta uma série mensal do preço do óleo de girassol no período entre Janeiro/1960 e Junho/2023, totalizando 762 observações. Modelos de previsão, baseados em Redes Neurais CNN (*Convolutional Neural Network*), CNN-Attention e LSTM (*Long Short Term Memory*) foram implementados na linguagem Python. Resultados obtidos, dos três modelos, foram comparados por meio das métricas R^2 (*Coefficient of Determination*), MAE (*Mean Absolute Error*), RSME (*Root Mean Squared Error*) e MAPE (*Mean Absolute Percent Error*). Verificou-se, para um horizonte de 13 meses, que o modelo CNN-Attention apresentou melhor desempenho, com relação as métricas R^2 , MAE, RSME e MAPE, que os modelos CNN e LSTM.

Palavras-chave: Mecanismo de atenção; Redes convolucionais; Keras; Python.

ABSTRACT

In recent years, the Attention Mechanism has been widely used in several areas of deep learning application. In this context, this work aims to propose the use of the CNN-Attention network to predict the price of sunflower oil. The database presents a monthly series of the price of sunflower oil in the period between January/1960 and June/2023, totaling 762 observations. Prediction models, based on CNN (Convolutional Neural Network), CNN-Attention and LSTM (Long Short Term Memory) Neural Networks were implemented in the Python language. Results obtained from the three models were compared using the metrics R^2 (Coefficient of Determination), MAE (Mean Absolute Error), RSME (Root Mean Squared Error) and MAPE (Mean Absolute Percent Error). It was verified, for a horizon of 13 months, that the CNN-Attention model presented better performance, in relation to the R^2 , MAE, RSME and MAPE metrics, than the CNN and LSTM models.

Keywords: Attention mechanism; Convolutional networks; Keras; Python.

¹ Discente. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio (PPGTCA - UTFPR). E-mail: leandroliveira@alunos.utfpr.edu.br. ORCID: <https://orcid.org/0000-0002-9027-1730>

² Doutor em Engenharia Elétrica. Docente do Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio (PPGTCA). Universidade Tecnológica Federal do Paraná (UTFPR). E-mail: airton@utfpr.edu.br. ORCID: <https://orcid.org/0000-0002-2568-5734>

1. INTRODUÇÃO

O girassol, nativo da América do Norte e Central, é uma cultura agrícola muito importante para o Brasil. Os girassóis são usados na fabricação de óleo de cozinha, farinhas e produtos de confeitaria. Podem ser utilizados também como matéria prima para produção de biodiesel. Seu bagaço é usado como matéria-prima na indústria de tortas, forrageiras, para alimentação do gado e em fertilizantes para a lavoura (MORPURGO, 2023; BARROS; JARDINE, 2021).

O girassol, quando comparado com a maioria das espécies cultivadas no Brasil, apresenta características agrônômicas importantes, como uma maior tolerância ao estresse hídrico e ao calor. A cultura, devido a essas características, torna-se uma importante alternativa econômica, principalmente nas regiões produtoras de grãos, em sistemas de sucessão de cultivos (RESENDE et al. 2022).

O óleo das sementes de girassol possui aplicações na indústria alimentícia, farmacêutica e cosmética. A semente de girassol, de maneira geral, possui de 45 a 65% de óleo em sua composição. Os principais componentes do óleo são: ácido palmítico (4%), ácido esteárico (1,47%), ácido oleico (49,02%), ácido linoleico (45,35%) e outros (0,11%). O óleo de sementes de girassol é considerado como um dos óleos de maior qualidade nutricional e organoléptica do mundo, trazendo benefícios para prevenção de doenças cardiovasculares e controle de níveis de colesterol no sangue (TORRES et al., 2021).

Vários fatores, como as condições climáticas, pragas, doenças, instabilidades políticas, etc, podem afetar o preço do óleo de girassol. O estudo do comportamento de preços, de produtos agropecuários, pode gerar parâmetros de interesse para produtores, analistas de mercado, entre outros (LUCAS; JÚNIOR, 2013). Portanto, o mercado do girassol necessita de ferramentas que minimizem os riscos e as incertezas no processo de decisão.

Modelos de previsão, de séries temporais, podem ser utilizados na previsão de preços de ativos financeiros. Dentre estes modelos, de previsão de séries temporais, estão as Redes Neurais CNN e LSTM. As redes CNNs, inspiradas em processos biológicos, vem sendo aplicadas com sucesso na análise de imagens digitais e séries temporais (SANTOS, 2022). Já as redes LSTMs são muito utilizadas em tarefas que envolvam dados sequenciais, como as séries temporais (SANTOS; CHAUCOSKI, 2020).

As redes neurais, baseadas em mecanismos de atenção, têm atraído grande interesse em pesquisas de aprendizagem profunda. Qiu et al. (2020) utilizaram um mecanismo de

atenção, em conjunto com uma rede LSTM (*Long Short Term Memory*), para previsão do preço de ações das bases S&P 500 (*Standard's & Poor's 500*) e DJIA (*Dow Jones Industrial Average*). Mnih et al. (2014) utilizaram um mecanismo de atenção, em conjunto com um modelo RNN (*Recurrent Neural Network*), para classificar imagens. Já Bahdanau et al. (2014) aplicaram o mecanismo de atenção ao campo de processamento de linguagem natural.

Neste contexto, este trabalho tem como objetivo propor a utilização da rede CNN-Attention para previsão do preço do óleo de girassol.

2. MATERIAIS E MÉTODOS

Aborda-se, nesta seção, a metodologia utilizada neste trabalho, apresentando a base de dados, os métodos propostos para previsão do preço do óleo de girassol, bem como as métricas utilizadas.

Base de dados:

Utilizou-se, para previsão do preço do óleo de girassol (US\$/t) (Cotação Internacional), uma base de dados com 762 observações (Jan/1960 - Jun/2023) (IPEADATA, 2023). O *Boxplot* dos dados é apresentado na Figura 1.

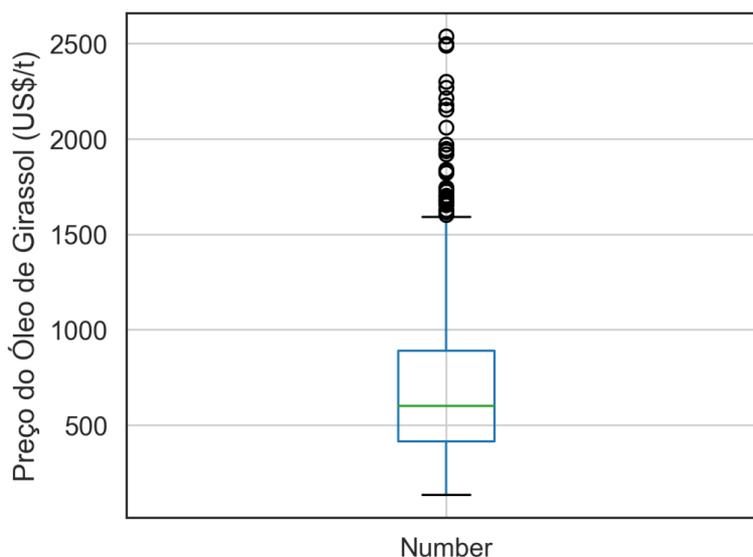


Figura 1. Boxplot dos dados do preço do óleo de girassol
A série histórica, do preço do óleo de girassol, é apresentada na Figura 2.

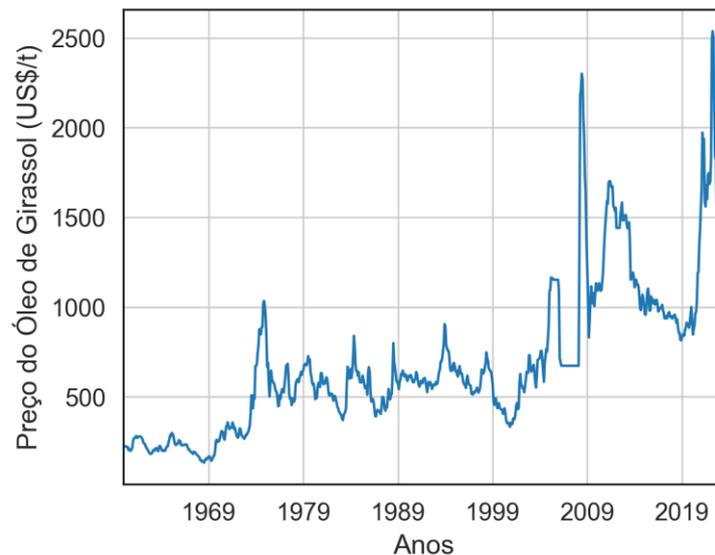


Figura 2. Série histórica do preço do óleo de girassol

Observou-se, da base de dados em 2022, um aumento no preço do óleo de girassol, nos primeiros meses da guerra entre a Rússia e a Ucrânia. Caindo, nos meses subsequentes, de 2487,8134 (US\$/t) em março/2022 para 1108.8604 (US\$/t) em junho/2023. A Rússia e a Ucrânia são um dos principais produtores e exportadores de óleo de girassol do mundo. O início da guerra na Ucrânia interrompeu inicialmente a produção e a exportação de óleo de girassol. A Ucrânia só aumentou sua oferta depois que um corredor, mediado pela ONU no Mar Negro, entrou em vigor em agosto/2022 (FORBES, 2022).

Rede CNN:

As redes CNNs, devido as suas vantagens, são muito utilizadas para reconhecimento de imagens. Sendo também muito eficazes para tratar sequências de tempo. As redes CNNs caracterizam-se por possuir camadas convolucionais. Estas camadas, para identificar características importantes, aplicam filtros aos dados de entrada. Esta abordagem permite que as redes CNNs capturem informações espaciais e temporais nos dados (LI et al., 2020; VARGAS, 2022, LAWAL, 2021; MANDELLI, 2023).

As redes CNNs são compostas, além da camada de entrada, por três camadas principais: camada convolucional, camada de *pooling* e uma camada totalmente conectada (Figura 3).

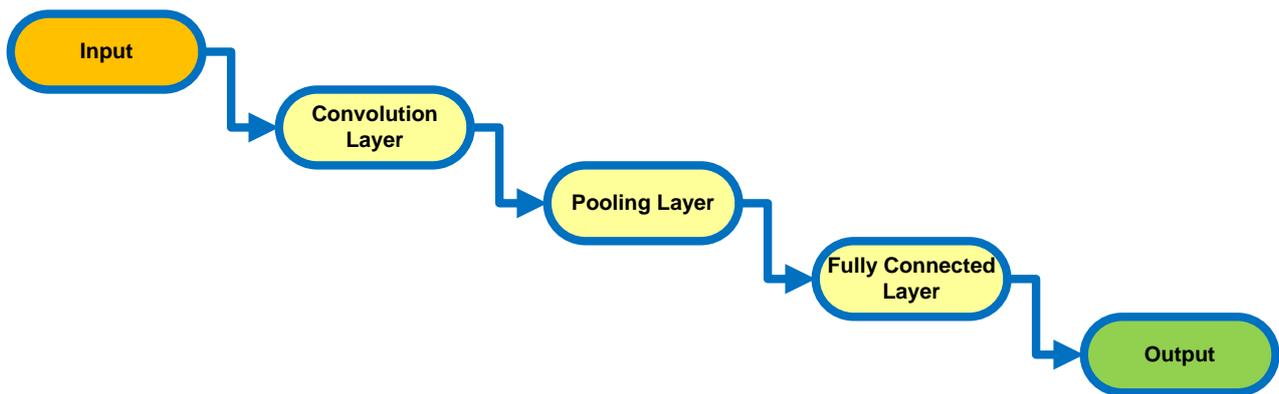


Figura 3. Estrutura da CNN

Rede LSTM:

A rede LSTM, proposta por Hochreiter e Scimdhuber (1997), é uma rede capaz de apreender seqüências de longo prazo. Esta rede é adequada para classificar, processar e prever séries temporais, com intervalos de tempo de duração desconhecida. A arquitetura de uma rede LSTM é baseada em células de memória (Figura 4). As portas de entrada (i_t), saída (o_t) e esquecimento (f_t) controlam o fluxo de informações na célula de memória.

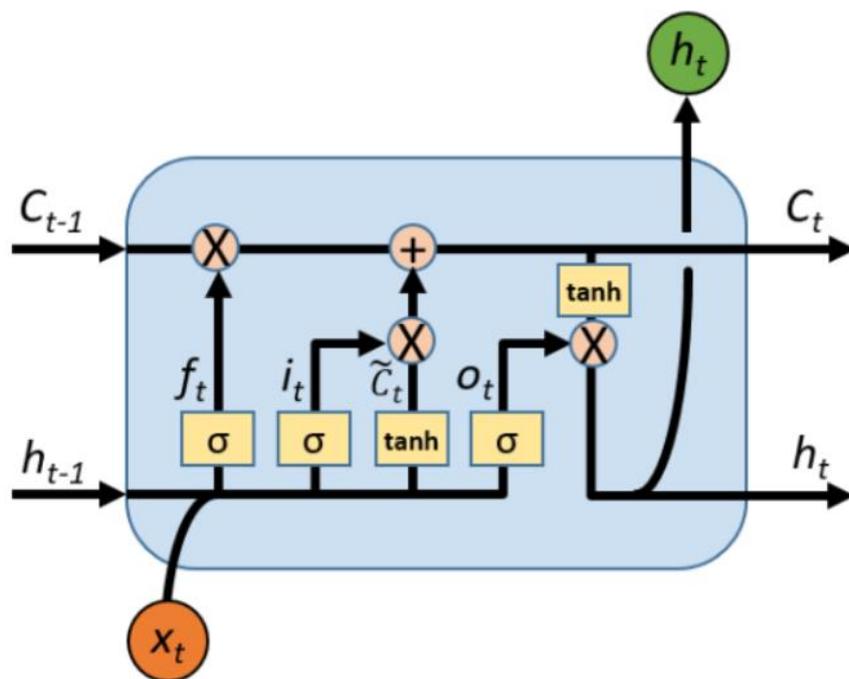


Figura 4. Célula de memória
Fonte: Sun et al. (2018).

As equações da rede LSTM são (SHEWALKAR et al., 2019):

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad 1$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad 2$$

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad 3$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad 4$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad 5$$

$$h_t = o_t * \tanh(C_t) \quad 6$$

Onde x_t e h_t são os vetores de entrada e saída, w_f , w_i , w_c e w_o são os pesos, b_f , b_i , b_c e b_o os bias, σ a função de ativação *sigmoid* e *tanh* a função de ativação tangente hiperbólica.

Mecanismo de atenção (*Attention Mechanism*)

Segundo Malingtan (2023), mecanismos de atenção, no aprendizado profundo, são usados para ajudar o modelo a focar nas partes mais relevantes dos dados. Em muitos problemas, os dados podem ser muito grandes e complexos e pode ser difícil para o modelo processar todos eles. Os mecanismos de atenção permitem que o modelo se concentre seletivamente nas partes dos dados, que são mais importantes para fazer uma previsão e ignore as partes menos relevantes. Isso pode ajudar o modelo a fazer previsões mais precisas e a funcionar com mais eficiência. Detalhes sobre o mecanismo de atenção podem ser encontrados em Saed (2022).

Etapas do trabalho

Inicialmente, neste trabalho realizou-se a análise dos dados de entrada (preço do óleo de girassol). Na sequência, na etapa de treinamento e validação, foram selecionados, por meio das métricas R^2 , MAE, RMSE e MAPE, os melhores modelos LSTM, CNN e CNN-Attention. Por fim, na etapa de teste, testaram-se os modelos para os dados que não participaram da etapa de treinamento e validação (Junho/2022 à Junho/2023).

Método de fragmentação e critério de parada

Neste trabalho, para divisão do conjunto de dados, em treinamento e validação, utilizou-se o método de *Houldout*. A base de dados foi dividida em 80% dos dados para treinamento e 20% para validação.

Como critério de parada utilizou-se o método conhecido como Método de Parada Antecipada (*Earling Stopping*) com o parâmetro *patience*=66. Este método, segundo Silva (2010) ajuda a projetar uma rede neural com bom poder de generalização. O parâmetro *patience* indica o número de épocas, após a qual nenhuma melhoria foi observada.

Métricas:

Os modelos, implementados neste trabalho, foram avaliados pelos parâmetros: Coeficiente de Determinação (R^2), Raiz Quadrada do Erro Quadrático Médio (RMSE), Erro Médio Absoluto (MAE) e Erro Percentual Absoluto Médio (MAPE). Tem-se uma previsão perfeita com $R^2=1$ e $RMSE=MAE=MAPE=0$ (PINHEIRO et al., 2020; BASTIANI et al., 2018; CANKURT; SUBASI, 2015).

Na Tabela 1 apresentam-se as equações das métricas MAE, RMSE, MAPE e R^2 .

Tabela 1. Equações MAE, RMSE, MAPE e R^2 .

Sigla	Equação
MAE	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) $
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i)^2)}$
MAPE	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)/y_i \times 100$
R^2	$\frac{(\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2)}{(\sum_{i=1}^n (y_i - \bar{y}_i)^2)}$

Onde: y_i é o valor real do período i , \hat{y}_i é a previsão para o período i , \bar{y}_i o valor médio e n é o número de observações.

3. RESULTADOS E DISCUSSÃO

Inicialmente, neste trabalho, realizou-se uma análise descritiva dos dados (Tabela 2).

Tabela 2. Análise descritiva.

Parâmetro analisado	Preço
Média (US\$/t)	688,21
Mediana (US\$/t)	600,55
Mínimo (US\$/t)	132,72
Máximo (US\$/t)	2536,62
Desvio Padrão (US\$/t)	418,81
Coeficiente de Variação (%)	60,7%

Pode-se observar, dos dados apresentados na Tabela 2, que o preço ficou, para o período em estudo, em média de 688,21 US\$/t. Apresentando, neste período, preços mínimo e máximo de 132,72 e 2536,62 US\$/t, respectivamente. Observa-se também um alto coeficiente de variação dos dados (60,7%) (PIMENTEL, 2009).

Treinamento e Validação

Realizou-se neste trabalho, para encontrar os melhores modelos, vários ajustes nos hiperparâmetros das redes CNN, CNN-Attention e LSTM. A Tabela 3 apresenta a melhor combinação de parâmetros encontrada.

Tabela 3. Parâmetros das redes.

	Parâmetros	Valor
Convolution Layer	Convolution layer filters	64
	Convolution layer kernel size	5
	Convolution layer activation function	Relu
	Convolution layer padding	Same
Pooling Layer	Pooling layer pool size	1
	Pooling layer padding	Same
	Pooling layer activation function	Relu
Training Parameters	Epoch	400
	Loss function	Relu
	Optimizer	Adam
	Batch size	10
	Time step	12
	Learning rate	0,00007

LSTM Layer	Activation Function	Relu
	Numbers of hidden units	16

Na Figura 5, apresenta-se a curva de aprendizagem para o preço do óleo de girassol. Observa-se, por meio desta figura, boa estabilidade na convergência das curvas de treino e validação.

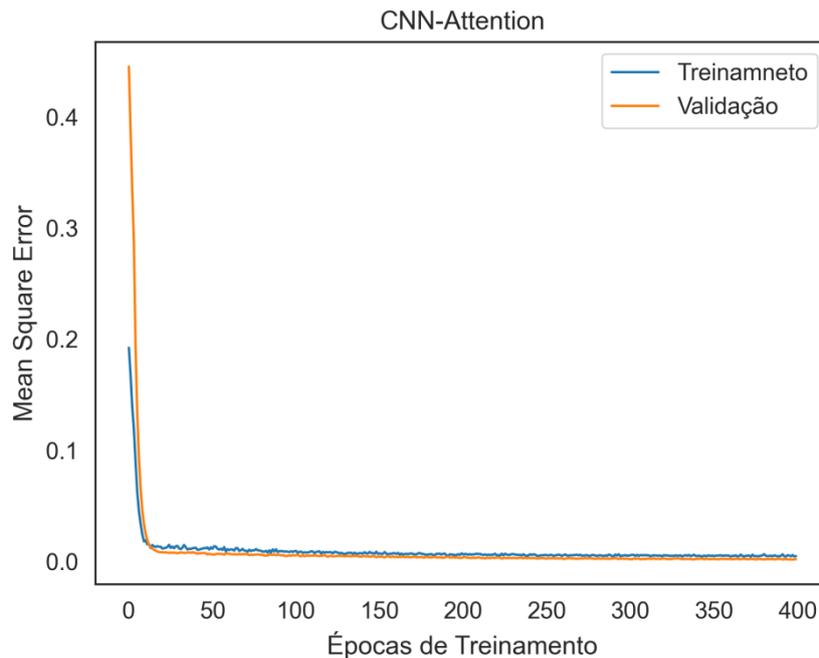


Figura 5. Curvas de aprendizagem de treino e validação

Na Figura 6 apresentam-se, como exemplo, os resultados da previsão, de treino e validação, para o modelo CNN-Attention.

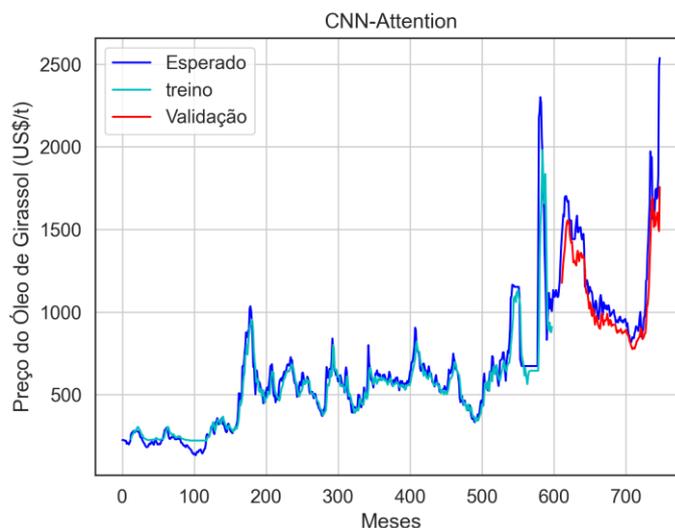


Figura 6. Rede CNN-Attention – treino e validação

Na Tabela 4 apresentam-se os dados, observados e preditos, pelos modelos de previsão, para os 13 meses que não participaram da etapa de treino e validação (conjunto de teste).

Tabela 4. Resultados das previsões dos treze meses - (US\$/t).

Mês	Observado	CNN-Attention	CNN	LSTM
jun/22	2154,40	2030,70	1836,95	2241,70
jul/22	1839,85	1890,92	1667,80	2127,14
ago/22	1823,44	1808,40	1431,69	2090,28
set/22	1578,28	1791,31	1233,60	2002,68
out/22	1627,13	1630,86	1093,08	1824,94
nov/22	1617,58	1597,48	951,84	1676,51
dez/22	1512,78	1553,79	909,81	1534,28
jan/23	1488,49	1601,65	819,05	1493,92
fev/23	1428,09	1477,82	765,67	1406,57
mar/23	1282,80	1334,40	658,03	1296,46
abr/23	1267,78	1178,09	613,81	1191,25
mai/23	1129,57	1125,48	558,29	1107,96
jun/23	1108,86	1081,17	528,38	1048,07

Na Figura 7 apresentam-se, na forma gráfica, os resultados das previsões para o conjunto de teste.

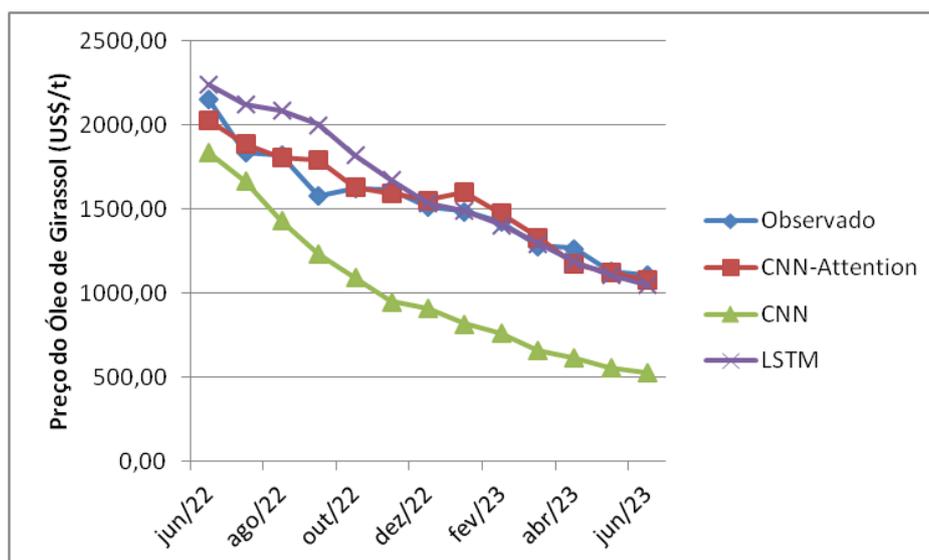


Figura 7. Previsões – conjunto de teste

Na Tabela 5 apresentam-se, para o conjunto de teste, os resultados das métricas MAE, RMSE, MAPE e R².

Tabela 5. Métricas – Conjunto de teste

Modelo	MAE (USD/kg)	RMSE (USD/kg)	MAPE (%)	R ²
CNN-Attention	61,81	84,3	4	0,92
CNN	522,4	545,1	36,1	-2,61
LSTM	118,74	174,4	7,3	0,63

Observa-se, dos resultados apresentados Figura 7 e na Tabela 5, que o modelo CNN-Attention obteve o melhor resultado. Este modelo, comparado com os modelos CNN e LSTM, apresentou um R² mais próximo de 1 (0,92) e erros menores (MAE, RSME e MAPE).

4. CONSIDERAÇÕES FINAIS

Buscou-se neste trabalho, a partir da dinâmica do passado, determinar o comportamento futuro do preço do óleo de girassol. Para tanto, utilizou-se uma série histórica, do preço do óleo de girassol, no período de janeiro/1960 à junho/2023, fornecida pelo Instituto de Pesquisa Econômica Aplicada (IPEA). Modelos de previsão CNN, CNN-Attention e LSTM, implementados na linguagem Python, passaram pelas Etapas de Treinamento, Validação e Teste,

Inicialmente, realizou-se o Treinamento e a Validação das redes CNN, CNN-Attention e LSTM. Na sequência, observou-se para o conjunto de teste (13 meses que não participaram das etapas de treinamento e validação (Jun/2022 à Jun/2023)), que a diferença entre valores reais e preditos, para a rede CNN-Attention, foi muito pequena, o que demonstra a boa capacidade de generalização desta rede. O modelo CNN-Attention, proposto neste trabalho, apresentou um MAPE de 4% e um R² próximo de 1 (0,92).

Por fim, pode-se concluir que a rede CNN-Attention pode auxiliar, produtores e compradores, na minimização dos riscos e incertezas relacionados com o mercado de óleo de girassol.

REFERÊNCIAS

- BAHDANAU, D.; CHO, K.; BENGIO, Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. **Computer Science**, 2014.
- BARROS, T. D.; JARDINE, J. G. **Girassol**. 2021. Disponível em: <<https://www.embrapa.br/agencia-de-informacao-tecnologica/tematicas/agroenergia/biodiesel/materias-primas/girassol>>. Acesso em: 04 jan. 2024.
- BASTIANI, M.; SANTOS, J. A. A.; SCHMIDT, C. A P.; SEPULVEDA, G. P. L. Application of data mining algorithms in the management of the broiler production. **Geintec**. v. 8, 2018.
- CANKURT, S.; SUBASI, A. Comparasion of linear regression and neural network models forecasting tourist arrivals to turkey. **Eurasian Journal of Science & Engineering**, 2015.
- FORBES. **Preços do óleo de girassol caem com retomada da oferta da Ucrânia**. 2022. Disponível em: <<https://forbes.com.br/forbesagro/2022/09/precos-do-oleo-de-girassol-caem-com-retomada-da-oferta-da-ucrania/>>. Acesso em: 04 jan. 2024.
- IPEADATA. **Commodities-óleo de girassol-cotação**. 2023. Disponível em: <internacionalhttp://www.ipeadata.gov.br/Default.aspx>. Acesso em: 04 nov. 2023.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **MIT press**, v. 9, 1997.
- LAWAL, A.; REHMAN, S; ALEMS, L. M.; ALAM, M. Wind Speed Prediction Using Hybrid 1D CNN and BLSTM Network. **IEEE Acess**, v.9, 2021.
- LI, P.; ABDEL-ATY, M.; YUAN, J. Real-time crash risk prediction on arterials based on LSTM-CNN. **Accident Analysis and Prevention**. v. 135, 2020.
- LUCAS, L. S.; JÚNIOR, A. G. S. Análise da volatilidade de preços do óleo de girassol no Brasil de 1960 a 2011. **Revista de Política Agrícola**, 2013.
- MANDELLI, P. **O que são Redes Neurais Convolucionais (CNNs)**. 2023. Disponível em: <<https://domineia.com/redes-neurais-convolucionais/>>. Acesso em: 04 out 2023.
- MALINGTAN, N. **Attention mechanism in deep learning**. 2023. Disponível em: <<https://www.scaler.com/topics/deep-learning/attention-mechanism-deep-learning/>>. Acesso em: 06 dez. 2023.
- MNIH, V.; CHO, K. BENGIO, Y. Recurrent Models of Visual Attention[J]. **Advances in neural information processing systems**, 2014.
- MORPURGO, C. R. **Plantação de girassol bate recorde de 46.393 hectares em 2023**. 2023. Disponível em: <<https://revistacultivar.com.br/noticias/plantacao-de-girassol-bate-recorde-de-46393-hectares-em-2023>>. Acesso em: 12 ago. 2023.
- PIMENTEL, F. **Curso de estatística experimental**. Piracicaba: ESALQ, 2009.
- PINHEIRO, T. C.; SANTOS, J. A. A.; PASA, L. A. Gestão da produção de frangos de corte por meio de redes neurais artificiais. **Revista Holos**. vol. 2, 2020.

QIU, J.; WANG, B.; ZHOU, C. Forecasting stock prices with long-short term memory neural network based on attention mechanism. **Plos One**, 2020.

RESENDE, J. C. F.; NOBRTE, M. C. R.; FREIRE, R. F.; MENDES, J. O.; SANTOS, D.; MOREIRA, D. A. Girassol: produção em Minas Gerais na safra ou safrinha. **EPAMIG**, 2022.

SANTOS, J. A. Previsão do preço do café arábica: uma aplicação de redes neurais CNN-BLSTM. **Research, Society and Development**, v. 11, 2022.

SANTOS, J. A. A.; CHAUCOSKI, Y. Previsão do Consumo de Energia Elétrica na Região Sudeste: Um Estudo de Caso Usando SARIMA e LSTM. **Revista Cereus**, v. 4, 2020.

SHEWALKAR, A.; NYAVANANDI, D.; LUDWIG, S. A. Performance evaluation of deep neural networks applied to speech recognition; RNN, LSTM and GRU. **JAICR**, 2019.

SAEED, M. **Adding a Custom Attention Layer to a Recurrent Neural Network in Keras**. 2022. Disponível em: <<https://machinelearningmastery.com/adding-a-custom-attention-layer-to-recurrent-neural-network-in-keras/>>. Acesso em: 19 mar. 2023.

SILVA, M. P. **Aplicação de redes neurais artificiais no diagnóstico de falha de turbina a gás**. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.

SUN, Q.; JANKOVIC, M. V.; BALLY, L. Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network, Belgrade, 2018. Disponível em: <https://arxiv.org/abs/1809.03817>. Acesso em: 29 jan. 2023.

TORRES, S. B.; QUEIROZ, A. L. F. G.; SANTOS, A. N. A. Óleo de girassol (*Helianthus annuus* L.) como cicatrizante de feridas em idosos diabéticos. **Brazilian Journal of Health Review**, 2021.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. **Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres**. 2022. Disponível em: <<http://sibgrapi.sid.inpe.br/col/sid.inpe.br/sibgrapi/2016/09.12.15.44/doc/um-estudo-sobre.pdf>>. Acesso em: 19 mar. 2023.