

**Previsão do preço do leite no estado do Paraná: um estudo
de caso utilizando ATTENTION-BiLSTM-CNN**

***Milk price forecasting in the state of Paraná: a case study
using ATTENTION-BiLSTM-CNN***

***Predicción del precio de la leche en el estado de Paraná: un
estudio de caso utilizando ATTENTION-BiLSTM-CNN***

José Airton Azevedo dos Santos¹

¹Doutor e Mestre em Engenharia Elétrica pela Universidade Federal de Santa Catarina (UFSC). Professor do Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio (PPGTCA) da Universidade Tecnológica Federal do Paraná (UTFPR). **E-mail:** airton@utfpr.edu.br,
Orcid: <https://orcid.org/0000-0002-2568-5734>

Resumo: A necessidade de decisões estratégicas por parte do produtor de leite, dada a volatilidade do preço do produto, gera demanda por ferramentas de previsão que aprimorem a gestão. Neste contexto, este trabalho tem como objetivo propor a utilização da rede híbrida Attention-BiLSTM-CNN para previsão, no estado do Paraná, do preço médio pago ao produtor de leite (R\$/litro). A base de dados, disponibilizada pelo Instituto de Pesquisa Econômica Aplicada (Ipea), apresenta uma série histórica do preço médio pago ao produtor, abrangendo o período de janeiro de 2000 a julho de 2025. Modelos de previsão Attention-BiLSTM-CNN, BiLSTM-CNN, BiLSTM e SARIMA foram implementados na linguagem de programação Python. Os resultados de previsão dos quatro modelos foram comparados com base nas métricas MAE, RMSE e MAPE. Verificou-se, para um horizonte de 12 meses, que o modelo proposto apresentou o melhor desempenho de previsão (MAE=0.086 R\$/litro, RMSE=0.108 R\$/litro e MAPE=3%).

Palavras-chave: métricas; python; redes híbridas; séries temporais.

Abstract: Given the volatility of dairy prices, the need for strategic decisions by milk producers generates demand for forecasting tools to improve management. In this context, this work aims to propose the use of the Attention-BiLSTM-CNN hybrid network for forecasting the average price paid to milk producers (R\$/liter) in the state of Paraná. The database, provided by the Institute for Applied Economic Research (IPEA), presents a historical series of the average price paid to producers, covering the period from January 2000 to July 2025. Attention-BiLSTM-CNN, BiLSTM-CNN, BiLSTM, and SARIMA forecasting models were implemented in the Python programming language. Forecasting results from the four models were compared using the metrics: MAE, RMSE, and MAPE. For a 12-month horizon, the proposed model showed the best forecasting performance (MAE=0.086 R\$/liter, RMSE=0.108 R\$/liter, and MAPE=3%).

Keywords: metrics; python; hybrid networks; time series.

Resumen: Dada la volatilidad de los precios de los lácteos, la necesidad de decisiones estratégicas por parte de los productores de leche genera demanda de herramientas de pronóstico para mejorar la gestión. En este contexto, este trabajo tiene como objetivo proponer el uso de la red híbrida Attention-BiLSTM-CNN para pronosticar el precio promedio pagado a los productores de leche (R\$/litro) en el estado de Paraná. La base de datos, proporcionada por el Instituto de Investigación Económica Aplicada (IPEA), presenta una serie histórica del precio promedio pagado a los productores que abarca el período de enero de 2000 a julio de 2025. Los modelos de pronóstico Attention-BiLSTM-CNN, BiLSTM-CNN, BiLSTM y SARIMA se implementaron en el lenguaje de programación Python. Los resultados de pronóstico de los cuatro modelos se compararon utilizando las métricas: MAE, RMSE y MAPE. Para un horizonte de 12 meses, el modelo propuesto mostró el mejor desempeño de pronóstico (MAE=0,086 R\$/litro, RMSE=0,108 R\$/litro y MAPE=3%).

Palabras clave: métricas; python; redes híbridas; series temporales.

1 INTRODUÇÃO

O Brasil figura entre os principais produtores de leite do mundo, com atividades espalhadas por praticamente todos os municípios do país. Essa cadeia produtiva envolve uma vasta rede de propriedades rurais. Estas propriedades sustentam o emprego de milhões de trabalhadores e representam, para as famílias agricultoras, uma fonte vital de renda.

Estados como o Paraná, no contexto regional, destacam-se na produção leiteira nacional, impulsionados por volumes expressivos, padrões elevados de qualidade e práticas eficientes. O setor, em períodos recentes, tem demonstrado crescimento notável, o que reflete a resiliência e o potencial da atividade agropecuária brasileira.

A volatilidade do preço do leite, influenciada por fatores como variações climáticas, custos de insumos, perecibilidade dos produtos e flutuações na demanda, representa um desafio significativo para o setor. Essa instabilidade impacta diretamente as estratégias dos produtores, fomentando a necessidade de ferramentas de previsão baseadas em séries temporais. Modelos de previsão confiáveis auxiliam na programação de comercializações, no gerenciamento de estoques essenciais e na elaboração de acordos contratuais prévios, reduzindo, assim, os riscos associados a oscilações imprevisíveis no mercado.

Embora a produção de leite tenha grande importância para o agro-negócio brasileiro, são poucos os trabalhos que utilizam redes neurais híbridas na previsão do preço do leite. Em grande parte, na previsão do preço do leite, são utilizadas abordagens tradicionais, como os modelos Arima e sua variante sazonal, o Seasonal Autoregressive Integrated Moving Average (Sarima). Essas técnicas clássicas, apesar de sua robustez e uso disseminado, muitas vezes apresentam limitações ao lidar com padrões não lineares complexos presentes nas séries temporais do leite. Portanto, as arquiteturas híbridas de *Deep Learning* tornam-se essenciais, sobretudo em períodos de forte instabilidade econômica, para superar as limitações dos modelos tradicionais.

Neste contexto, este trabalho tem como objetivo propor o uso da rede híbrida Attention-BiLSTM-CNN para a previsão do preço médio pago ao produtor de leite (R\$/litro) no estado do Paraná. A importância da arquitetura

híbrida Attention-BiLSTM-CNN na previsão do preço do leite reside em sua capacidade de combinar as melhores características de três modelos avançados de *Deep Learning*: o mecanismo *Attention* e as redes BiLSTM e CNN.

A Convolutional Neural Network (CNN) é essencial para extrair características locais e de curto prazo a partir de séries temporais. A *Bidirectional Long Short-Term Memory* (BiLSTM) lida com dependências temporais de longo prazo, capturando a tendência histórica e a sazonalidade dos preços. Já o Mecanismo de Atenção (*Attention Mechanism*) permite que o modelo foque nas partes mais relevantes dos dados de entrada.

Dessa forma, o estudo busca, ao analisar o desempenho da arquitetura híbrida (Attention-BiLSTM-CNN), contribuir para o avanço do estado da arte. Essa análise considera a elevada volatilidade e a complexidade não linear características do mercado lácteo, buscando superar algumas das limitações observadas nos modelos de previsão tradicionais.

2 MATERIAL E MÉTODOS

Nesta seção, apresenta-se uma breve descrição da rede Attention-BiLSTM-CNN utilizada neste trabalho, bem como as características da base de dados, as métricas de avaliação adotadas e uma síntese de trabalhos correlatos.

2.1 Trabalhos correlatos

Vários trabalhos relatam o uso de modelos de séries temporais para a previsão do preço do leite. Roschel, Simões e Godoy (2023), por exemplo, utilizaram modelos *Sarima* e *Multilayer Perceptron* (MLP) para a previsão, no Brasil, do preço pago aos produtores. Os autores observaram que o *Sarima* apresentou melhor desempenho de previsão. Concluíram que as arquiteturas de redes neurais mais simples, como a MLP, podem apresentar limitações em comparação a modelos estocásticos tradicionais para capturar a estrutura da série temporal láctea.

Bezerra Neto (2022) explorou o potencial da aprendizagem supervisionada para mitigar a incerteza do produtor quanto ao valor do leite recebido. Desenvolveu um modelo de regressão linear baseado em dados

de cooperativas agropecuárias. Embora o estudo tenha alcançado métricas de precisão satisfatórias, o autor ressalta que a abordagem apresentou limitações devido ao reduzido volume da amostra de dados e sugere que o uso de séries históricas mais extensas poderia melhorar o desempenho do modelo em aplicações futuras.

Vilela e Penedo (2024) investigaram, por meio da comparação da eficácia de modelos Arima e Sarima, o comportamento dos preços do leite na região do Triângulo Mineiro e Alto Paranaíba. O estudo confirmou a forte presença da sazonalidade no setor lácteo, demonstrando que a inclusão do componente sazonal no modelo Sarima $(0,1,1) (2,1,0)_{12}$ resultou em melhor desempenho. Os autores concluem que, embora os modelos autorregressivos sejam ferramentas viáveis para auxiliar o produtor na mitigação de riscos, a alta volatilidade da *commodity* permanece um desafio central para o planejamento das atividades agropecuárias.

Daneluz *et al.* (2018) analisaram, utilizando a metodologia de Box-Jenkins para identificar padrões de sazonalidade e tendência, o comportamento da série temporal de preços do leite em uma propriedade rural no Rio Grande do Sul. Os autores determinaram que o modelo Sarima $(1,1,0) (2,0,0)_{12}$ apresentou o melhor ajuste para a série de preços e destacaram que a utilização de previsões econômicas é uma ferramenta administrativa primordial para reduzir riscos e garantir a viabilidade estratégica da atividade leiteira, a curto e a longo prazo.

Já Kresova e Hess (2022) analisaram os determinantes dos preços regionais do leite cru na Rússia entre 2015 e 2019. Empregaram o algoritmo de aprendizado de máquina *Random Forest* (RF) com validação cruzada espaço-temporal. O estudo revelou que variáveis como a produção de leite para consumo humano, o nível de renda, o efetivo do rebanho e a densidade populacional são os principais fatores explicativos das variações de preço, evidenciando relações não lineares complexas. Os autores concluíram que o uso de técnicas de *Machine Learning* permite capturar autocorrelações espaciais e temporais que métodos tradicionais poderiam negligenciar. Além disso, ressaltaram que essas técnicas podem fornecer subsídios cruciais para a tomada de decisão estratégica de produtores e de empresas do setor de laticínios.

2.2 Rede Attention-BiLSTM-CNN

A rede Attention-BiLSTM-CNN, proposta neste trabalho, é formada por um mecanismo de atenção e pelas redes BiLSTM e CNN.

2.2.1 Attention Mechanism

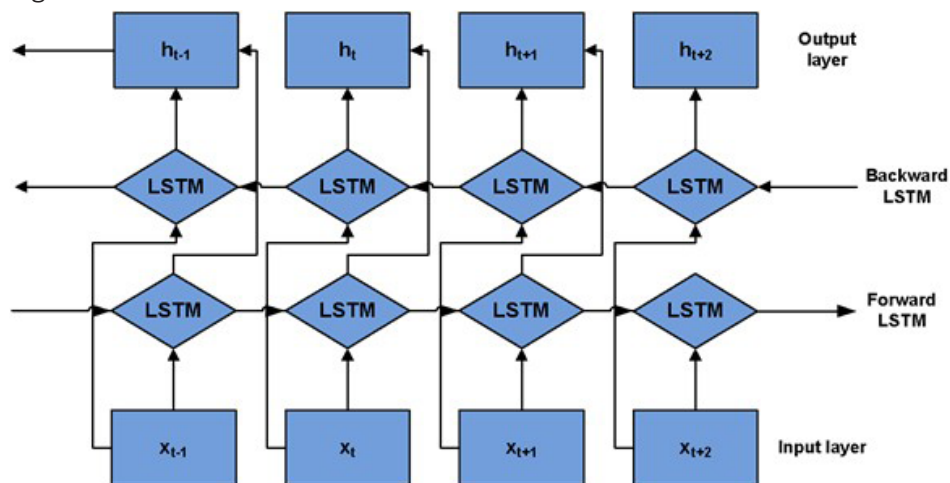
O mecanismo de atenção permite que um modelo de previsão priorize, por meio da atribuição de pesos, as partes mais importantes dos dados de entrada. Sua utilização faz com que o modelo foque nas partes dos dados mais importantes para previsão, ignorando as partes menos relevantes. Isso permite que o modelo realize previsões com maior precisão (Malingtan, 2023).

2.2.2 BiLSTM

A arquitetura BiLSTM, proposta originalmente por Schuster e Paliwal em 1997, representa uma evolução significativa em relação às redes LSTM tradicionais. Diferentemente das versões unidirecionais, a BiLSTM é composta por duas camadas intermediárias que processam a sequência de dados em direções opostas: uma camada *forward*, que percorre a sequência do início ao fim, e uma camada *backward*, que percorre do fim ao início. Ambas as direções produzem representações complementares, que são posteriormente combinadas e transmitidas à camada de saída (Figura 1).

Essa configuração bidirecional confere à rede a capacidade de considerar simultaneamente o contexto passado e futuro de uma sequência, o que resulta em um aprendizado mais rico e eficiente. É utilizada em tarefas que envolvem dados temporais ou sequenciais, como o processamento de linguagem natural, a previsão de séries temporais e o reconhecimento de padrões dinâmicos (Santos, 2022; Silva; Oliveira; Santos, 2023).

Figura 1 – Estrutura de uma rede BiLSTM

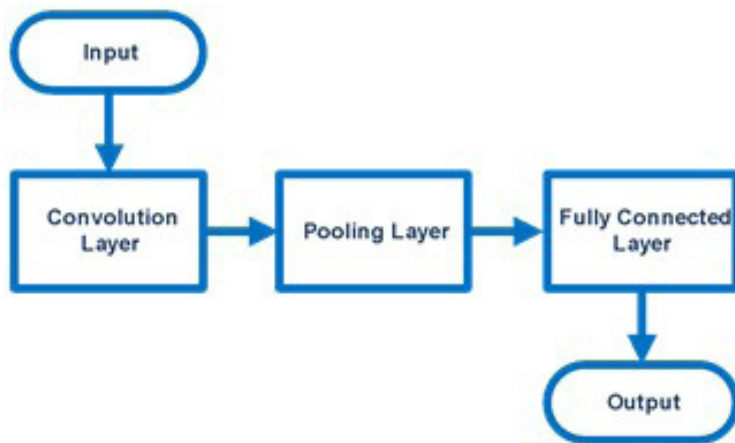


Fonte: Saadi *et al.* (2023).

CNN:

As *Convolutional Neural Networks* (CNN) são muito utilizadas no processamento de imagens. Sequências de dados, tais como séries temporais, são tratadas pela versão 1D das redes CNN. Estas redes são compostas basicamente por uma camada convolucional e uma camada de *pooling* (Figura 2). Em aplicações práticas, para séries temporais, esses dois tipos de camadas são empregados em conjunto: primeiro para extrair e condensar padrões e, depois, para simplificar a representação dos dados, antes de serem processados por uma camada final, geralmente totalmente conectada (Deng *et al.*, 2024; Li; Abdel-Aty; Yuan, 2019; Livieris; Pintelas; Pintelas, 2020; Lu *et al.*, 2020).

Figura 2 – Estrutura da CNN



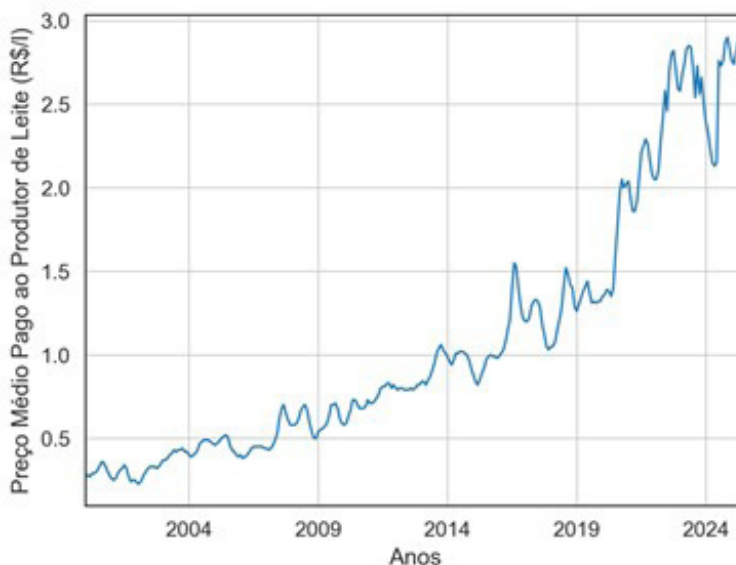
Fonte: Alzubaidi *et al.* (2021).

2.3 Base de dados

Para a previsão do preço médio pago ao produtor de leite (R\$/litro), no estado do Paraná, utilizou-se uma base de dados com 307 meses (janeiro de 2000 a julho de 2025), obtida do Instituto de Pesquisa Econômica Aplicada (Ipea) (Ipea, 2025).

Na Figura 3, é apresentada a série histórica do preço médio pago ao produtor paranaense. Esta série é marcada por uma tendência nominal positiva de longo prazo, mas com volatilidade elevada e crescente, culminando em picos e quedas acentuados na fase mais recente, o que evidencia a complexidade e o risco da atividade leiteira no Paraná.

Figura 3 – Série histórica do preço médio pago ao produtor



Fonte: Ipea (2025).

A característica mais marcante desta série histórica é a forte e consistente tendência de alta nominal do preço do leite ao longo de todo o período. O preço passa de R\$ 0,27/litro no início de 2000 para atingir um pico de R\$ 2,9/litro em novembro de 2024.

A base de dados, obtida do Ipea, foi dividida nos conjuntos de treinamento, validação e teste. O conjunto de treinamento ficou com os 206 primeiros dados, o conjunto de validação foi formado pelos seguintes 89 dados e o conjunto de teste foi formado pelos 12 dados restantes. Após esse processo, os dados foram normalizados no intervalo [0,1]. A normalização foi realizada por meio da função `MinMaxScaler()` da biblioteca `Scikit-learn`. Observa-se, para evitar vazamento de informações, que o processo de normalização foi aplicado apenas ao conjunto de treinamento.

2.4 Métricas

Neste trabalho, os modelos foram avaliados pelas seguintes métricas (Cankurt; Subasi, 2015; Oliveira; Santos, 2024; Santos; Polo; Santos, 2024):

- *Mean Absolute Percent Error* (MAE): indica a média do erro absoluto entre o valor previsto e o valor real;
- *Root Mean Squared Error* (RMSE): indica o desvio padrão da amostra da diferença entre o valor previsto e o valor real; w
- *Mean Absolute Percent Error* (Mape): mede o tamanho do erro em termos percentuais.

As equações das métricas MAE, RMSE e Mape são apresentadas na Tabela 1.

Tabela 1 – Equações- Métricas

Métrica	Equação	Valor Perfeito
MAE	$\frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $	MAE=0
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i)^2)}$	RMSE=0
MAPE	$\frac{1}{n} \sum_{i=1}^n \left \frac{(y_i - \hat{y}_i)}{y_i} \right \times 100$	MAPE=0

Fonte: Santos, Polo e Santos (2024).

Onde: y_i é o valor real do período i , $\hat{y}_i$ é a previsão para o período i 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	Resultado
Média (R\$/litro)	1,079
Mínimo (R\$/litro)	0,23
Máximo (R\$/litro)	2,9
Desvio Padrão (R\$/litro)	0,76
Coeficiente de Variação (%)	70,4

Fonte: autores (2025).

Observa-se, a partir dos dados apresentados na Tabela 2, que o preço médio pago ao produtor de leite ficou em média de 1,079 R\$/litro no período em estudo, apresentando consumos mínimo e máximo de 0,23 R\$/litro e 2,9 R\$/litro, respectivamente. Observa-se também que a média de 1,079 R\$/litro, dada a enorme dispersão dos dados, não pode ser usada como um valor típico ou confiável para fins de planejamento. O Coeficiente de Variação (CV) de 70,4% é a métrica mais importante, indicando que a volatilidade do preço do leite é alta para o período analisado.

3.1 Sarima

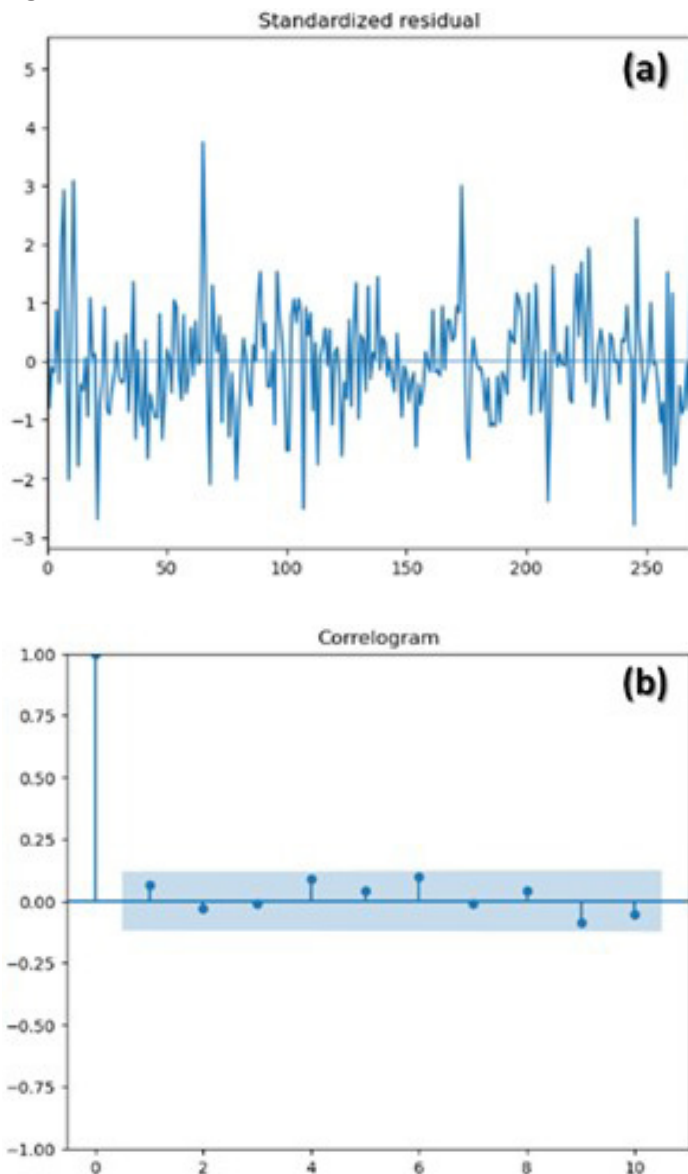
Inicialmente, identificou-se, de maneira automática, por meio do pacote pmdarima, o modelo Sarima (2,1,2)(4,2,1)₁₂. Este modelo obteve, entre os modelos testados, o menor valor do critério de AKAIKE (AIC).

A suposição fundamental, para modelos Sarima (e Arima, em geral), é que os resíduos constituam um ruído branco: média zero, variância constante (homocedasticidade) e ausência de autocorrelação. O gráfico de resíduos padronizados (*Standardized Residual*) (Figura 4a) exibe uma distribuição aleatória ao longo do tempo, com valores predominantemente contidos na faixa de ± 2 (com raros picos até ± 3), sem tendências, padrões sazonais ou variações na dispersão. Isso indica que a variância é constante e que os erros são independentes.

Já o correlograma (*Correlogram*) (Figura 4b) demonstra que todos os coeficientes de autocorrelação se encontram dentro do intervalo de confiança de 95% (área azul sombreada), sem nenhum pico significativo. Isso confirma a ausência de autocorrelação residual e reforça a validade da

hipótese de ruído branco. Esses resultados, combinados com a ausência de estrutura sistemática nos resíduos, validam a adequação do modelo Sarima para a série analisada.

Figura 4 – Gráficos do modelo Sarima



Fonte: Elaborado pelos autores (2025).

3.2 ATTENTION-BILSTM-CNN

3.2.1 Treinamento e Validação

Na etapa de treinamento e validação de modelos de *Machine Learning*, o foco principal reside na identificação, por meio da otimização de hiperparâmetros, dos melhores modelos preditivos. Neste trabalho, para realizar a otimização dos hiperparâmetros, utilizou-se a ferramenta Optuna. Essa ferramenta, baseada em técnicas bayesianas, automatiza a busca eficiente por combinações ideais de hiperparâmetros, como taxa de aprendizado, número de camadas ocultas, tamanho do *batch*, taxa de *dropout*, etc. (Deng et al., 2024). Na Tabela 3 são apresentados os melhores hiperparâmetros encontrados por meio do Optuna.

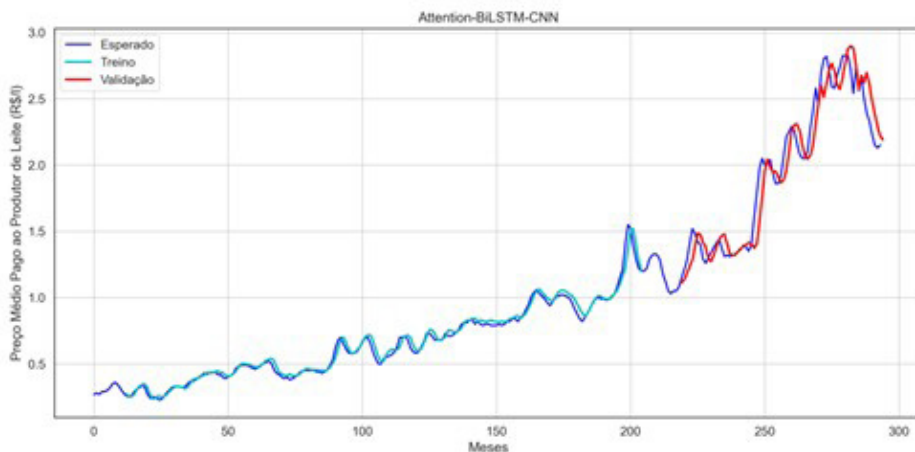
Tabela 3 – Parâmetros das redes

Hiperparâmetro	Intervalo de Busca	Valores Ótimos
<i>Epochs</i>	[50,500]	300
<i>Batch</i>	[16,64]	16
<i>LSTM activation function</i>	[tanh,relu]	tanh
<i>Convolution layer kernel size</i>	[3,7]	4
<i>Convolution layer filters</i>	[8,128]	64
<i>Numbers of hidden units</i>	[16,128]	128
<i>Pooling layer pool size</i>	[1,2]	2
<i>Dropout</i>	[0.2,0.5]	0.43
<i>Learning rate</i>	[0.00001,0.1]	0.00167
<i>Optimizer</i>	[RMSprop, Adam]	Adam

Fonte: Elaborado pelos autores (2025).

Na Figura 5, são apresentados, como exemplo, os resultados da predição, de treino e de validação para o modelo Attention-BiLSTM-CNN.

Figura 5 – Rede Attention-BiLSTM-CNN- Treino e validação



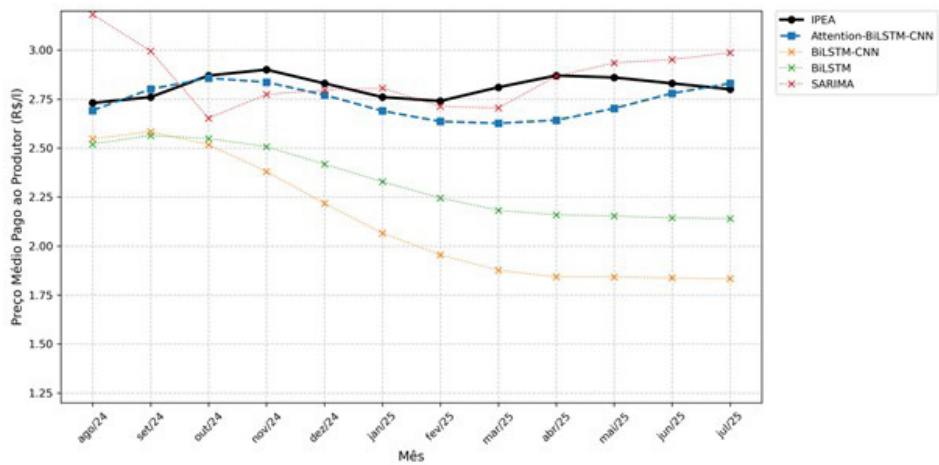
Fonte: Elaborado pelos autores (2025).

Observa-se, na Figura 5, que a curva de treinamento (treino) está praticamente sobreposta à curva dos valores esperados (reais). Isso sugere que o modelo capturou de forma satisfatória a relação entre os dados de entrada e os preços históricos do leite. O modelo também conseguiu prever com sucesso, no conjunto de validação, a aceleração e os picos de preço (como o aumento acentuado após o mês 250), bem como as quedas sazonais e temporárias.

3.2.2 Teste

A Figura 6 apresenta, para o conjunto de teste (dados não vistos), os resultados da previsão do preço médio pago ao produtor de leite (R\$/litro), abrangendo o período de agosto de 2024 a julho de 2025. O gráfico compara o preço real (Ipea) às previsões das quatro diferentes arquiteturas de aprendizado de máquina. Os modelos Attention-BiLSTM-CNN e Sarima são os que apresentam a melhor generalização e precisão no conjunto de teste. Eles conseguem prever o padrão geral de queda e de posterior recuperação do preço, seguindo de perto a curva do Ipea. Já os modelos BiLSTM-CNN e BiLSTM apresentam os piores resultados, demonstrando uma tendência de subestimar o preço real.

Figura 6 – Previsões- Conjunto de teste



Fonte: Elaborado pelos autores (2025).

A Tabela 4 apresenta os valores das métricas de erro dos quatro modelos no conjunto de teste (agosto de 2024 a julho de 2025), que são cruciais para quantificar o desempenho de cada arquitetura na previsão do preço do leite.

Tabela 4 – Métricas- Conjunto de teste

Modelos	MAE (R\$/litro)	RMSE (R\$/litro)	Mape (%)
Attention-BiLSTM-CNN	0,086	0,108	3,00
BiLSTM-CNN	0,688	0,754	24,41
BiLSTM	0,487	0,520	17,29
SARIMA	0,136	0,181	4,86

Fonte: Elaborado pelos autores (2025).

Observa-se, a partir dos dados da Tabela 4, que o modelo Attention-BiLSTM-CNN apresenta o melhor desempenho. Seu MAE de 0,086 R\$/litro indica que, ao longo dos 12 meses previstos, o erro médio absoluto entre o preço previsto e o preço real foi de 8,6 centavos por litro. Os Mapes de 3% para o Attention-BiLSTM-CNN e de 4,86% para o Sarima são considerados muito bons para séries temporais econômicas voláteis, como o preço de *commodities*.

Nota-se também, a partir dos dados apresentados na Tabela 4, que os modelos BiLSTM-CNN e BiLSTM apresentam erros muito altos para uso prático (Mape=24,41% e Mape=17,29%). Já o Sarima, método estatístico clássico, apresentou um desempenho consistente, com um Mape de 4,86% e um MAE de 0,136 R\$/litro. Embora seja um resultado satisfatório para uma abordagem linear, ele foi superado pela arquitetura híbrida com camada convolucional (Attention-BiLSTM-CNN), que alcançou Mape de 3%. Isso indica o melhor desempenho, na captura de não linearidades complexas, da rede híbrida Attention-BiLSTM-CNN.

3.3 Contextualização dos resultados

A aplicação de *Deep Learning* para a previsão de *commodities* agrícolas, no cenário internacional recente, tem convergido para o uso de arquiteturas híbridas e de mecanismos de atenção. Tal movimento valida a abordagem adotada nesta pesquisa. Por exemplo, Yin *et al.* (2020) utilizaram um modelo *Attention*, em conjunto com a rede LSTM, para previsão de preços de hortaliças. O modelo proposto foi aplicado a cinco culturas agrícolas: repolho, rabanete, cebola, pimenta e alho. Concluíram, neste trabalho, que o modelo Attention-LSTM, com relação ao modelo LSTM, apresentou um melhor desempenho de previsão.

Já Gu *et al.* (2022) avançaram na discussão ao destacar a superioridade de modelos LSTM equipados com mecanismos de *Attention* (como o DIA-LSTM). Essa arquitetura, segundo os autores, é mais robusta para capturar dependências temporais dinâmicas. Estes exemplos se alinham diretamente aos resultados obtidos neste estudo, no qual a inclusão do módulo de atenção melhorou o desempenho do modelo BiLSTM-CNN.

3.4 Limitações do Trabalho

Este trabalho apresenta as seguintes limitações:

- Abordagem univariada: o estudo utilizou uma abordagem univariada. Isso significa que o modelo aprende apenas com o histórico do próprio preço do leite.

- Delimitação geográfica: o trabalho é um estudo de caso focado exclusivamente no estado do Paraná. Portanto, os resultados e o modelo treinado não podem ser generalizados automaticamente para todo o Brasil sem novos testes.
- Comparação metodológica: o estudo compara o modelo proposto (Attention-BiLSTM-CNN) apenas com variações de redes neurais e com o modelo estatístico clássico Sarima. Não foi comparado o desempenho da rede Attention-BiLSTM-CNN com outros algoritmos de *Machine Learning*, como o Prophet (do Facebook), o XGBoost ou modelos *Transformers* puros.

4 CONSIDERAÇÕES FINAIS

O trabalho propôs o uso da rede neural híbrida Attention-BiLSTM-CNN para a previsão do preço médio do leite (R\$/litro) no estado do Paraná. Esta proposta se justifica pela extrema volatilidade e complexidade inerentes a essa série temporal. A análise descritiva dos dados revelou um coeficiente de variação de 70,4% e uma grande amplitude de preço (de R\$ 0,23/litro a R\$ 2,9/litro) ao longo dos 307 meses analisados (janeiro de 2000 a julho de 2025). Isso evidenciou a necessidade de um algoritmo de previsão avançado. Frente a este cenário de volatilidade de preços, que afeta diretamente as decisões do produtor, a pesquisa teve como objetivo primário fornecer uma ferramenta de suporte ao processo decisório. Para comprovar sua eficácia, o desempenho da rede híbrida Attention-BiLSTM-CNN foi comparado ao das arquiteturas BiLSTM-CNN, BiLSTM e Sarima.

Os resultados das previsões dos quatro modelos foram avaliados no conjunto de teste (de outubro de 2023 a julho de 2025), utilizando as métricas MAE, RMSE e Mape. O modelo proposto, Attention-BiLSTM-CNN, apresentou o melhor desempenho de previsão, com erros absolutos e percentuais baixos: MAE de 0,086 R\$/litro e Mape de 3%. Esse bom desempenho foi atribuído à combinação do mecanismo de atenção com as redes BiLSTM, que modela as dependências de longo prazo, e CNN, essencial para a extração de características locais e a captura da volatilidade de curto prazo.

A comparação com o método estatístico clássico reforça os bons resultados obtidos pela abordagem proposta. O modelo Sarima, amplamente

utilizado na literatura, apresentou desempenho consistente, com Mape de 4,86% e MAE de 0,136 R\$/litro. Esses resultados confirmam a utilidade de modelos lineares para capturar a sazonalidade básica. Contudo, a superioridade das métricas do modelo híbrido evidencia que a inclusão de camadas convolucionais e de mecanismos de atenção permite uma melhor captura das não linearidades e da volatilidade do mercado.

Em síntese, o modelo híbrido Attention-BiLSTM-CNN demonstrou ser uma ferramenta eficiente para a previsão do preço do leite. Portanto, este modelo pode ser utilizado pelo produtor paranaense para apoio ao planejamento, à análise de cenários de investimento e ao embasamento da tomada de decisão.

REFERÊNCIAS

ALZUBAIDI, L. *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, [s.l.], v. 8, 2021.

BEZERRA NETO, J. A. *Predição do preço do leite utilizando dados de cooperativas*. 2022. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Centro de Engenharia Elétrica e Informática, Universidade Federal de Campina Grande (UFCG). Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/37739>. Acesso em: 10 mai. 2025.

CANKURT, S.; SUBASI, A. Comparasion of linear regression and neural network models forecasting tourist arrivals to Turkey. *Eurasian Journal of Science & Engineering*, [s.l.], v. 1, n. 1, 2015.

DANELUZ, M.; BATISTA, H.; CANEVER, M.; MENEZES, G. *Previsão do preço e produção do leite utilizando um Modelo SARIMA: um estudo de caso*. In: ENCONTRO DE ECONOMIA GAÚCHA, 9., 2018, Porto Alegre. *Anais [...]*. Porto Alegre: PUCRS, 2018. Disponível em: <https://editora.pucrs.br/edipucrs/acessolivre/anais/encontro-de-economia-gaucha/assets/edicoes/2018/arquivos/86.pdf>. Acesso em: 10 maio 2025.

DENG, Z.; ZHOU, D.; KANG, Z.; DONG, H. Deep learning-based dynamic forecasting method and application for ultra-deep fractured reservoir production. *Frontiers in Energy Research*, [s.l.], v. 12, 2024. DOI:

<https://doi.org/10.3389/fenrg.2024.1369606>

GU, Y. H.; JIN, D.; YIN, H.; ZHENG, R.; PIAO, X.; YOO, S. J. Forecasting agricultural commodity prices using dual input attention LSTM. *Agriculture*, [S. l.], v. 12, n. 2, 2022.

IPEA. *Preço médio recebido pelo agricultor: leite (litro - Paraná)*. Disponível em: <https://http://www.ipeadata.gov.br/Default.aspx>. Acesso em: 10 out. 2025.

KRESOVA, S.; HESS, S. Identifying the determinants of regional raw milk prices in Russia using machine learning. *Agriculture*, [s.l.], v. 12, n. 7, 2022. DOI: <https://doi.org/10.3390/agriculture12071006>

LI, P.; ABDEL-ATY, M.; YUAN, J. Real-time crash risk prediction on arterials based on LSTM-CNN. *Accident Analysis and Prevention*, [s.l.], v. 135, 2019.

LIVIERIS, I. E.; PINTELAS, E.; PINTELAS, P. A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, [s.l.], v. 32, 2020.

LU, W.; LI, J.; LI, Y.; SUN, A.; WANG, J. A CNN-LSTM-based model to forecast stock prices. *Complexity*, [s.l.], 2020.

MALINGTAN, N. Attention mechanism in deep learning. *Scaler Topics*, [s.l.], 2023. Disponível em: <https://www.scaler.com/topics/deep-learning/attention-mechanism-deep-learning/>. Acesso em: 6 dez. 2023.

OLIVEIRA, L.; SANTOS, J. A. A. Previsão do preço futuro do óleo de girassol: uma abordagem baseada no modelo CNN-Attention. *CEREUS*, Gurupi, v. 16, n. 2, 2024.

ROSCHER, M. S.; SIMÕES, A. R. P.; GODOY, A. P. J. Análise e previsão do preço do leite pago aos produtores de leite no Brasil. In: Workshop de Pós-Graduação em Zootecnia e Ciência Animal do Estado de Mato Grosso do Sul, 7., 2023, Campo Grande. *Anais [...]*. Campo Grande: UFGD; UFMS; UEMS; UCDB, 2023. Disponível em: <https://ocs.ufgd.edu.br/index.php?conferenc e=pgzootecnia&schedConf=viiworkshop&page=paper&op=view&path%5B%5D=2212>. Acesso em: 10 mai. 2025.

SAADI, M. N.; BOUTRAA, N.; READJATI, A.; MOHAMED, B. A novel method for bearing fault diagnosis based on BiLSTM neural networks. *The International Journal of Advanced Manufacturing Technology*, v. 125, n. 3-4, 2023.

SANTOS, J. A. A.; POLO, A. N. B.; SANTOS, C. O. Análise Comparativa de Modelos LSTM e GRU Para Previsão da Velocidade do Vento. *Revista FSA*, Teresina, v. 21, n. 5, 2024.

SANTOS, J. A. Previsão do preço do café arábica: uma aplicação de redes neurais CNN-BLSTM. *Research, Society and Development*, [s.l.], v. 11, n. 3, e3511326101, 2022.

SCHUSTER, M.; PALIWAL, K. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, v. 45, n. 11, p. 2673–2681, 1997.

SILVA, G. B.; OLIVEIRA, L.; SANTOS, J. A. A. BiLSTM-CNN network applied to the forecast of electricity consumption in agroindustry of the southwest region of Goiás. *Revista CEREUS*, Gurupi, v. 15, n. 3, 2023.

VILELA, E. H. P.; PENEDO, A. S. T. Previsão do comportamento dos preços do leite usando modelos ARIMA e SARIMA. *Revista Estação Científica*, Juiz de Fora, v. 18, n. 31, 2024. DOI: <https://doi.org/10.5281/zenodo.13625559>

YIN, H.; JIN, D.; GU, Y. H.; PARK, C. J.; HAN, S. K.; YOO, S. J. STL-ATTLSTM: vegetable price forecasting using STL and attention mechanism-based LSTM. *Agriculture*, [s.l.], v. 10, n. 12, 2020.