

## **Previsão de variáveis econômicas com aprendizado de máquina: previsão com redes neurais**

### **Economic variables forecasting with machine learning: forecasting with neural networks**

DOI: 10.55905/oelv21n9-042

Recebimento dos originais: 04/08/2023

Aceitação para publicação: 04/09/2023

#### **Christiano Pereira Amaral**

Graduado em Economia

Instituição: Universidade do Estado do Rio de Janeiro (UERJ)

Endereço: Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro - RJ

E-mail: amaral.christianopa@gmail.com

#### **Daiane Rodrigues dos Santos**

Doutora em Engenharia Elétrica

Instituição: Universidade do Estado do Rio de Janeiro (UERJ)

Endereço: Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro - RJ

E-mail: daianesantoseco@gmail.com

#### **Elcyon Rocha Lima**

Doutor em Economia

Instituição: Universidade do Estado do Rio de Janeiro (UERJ)

Endereço: Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro - RJ

E-mail: elcyon.lima@gmail.com

#### **Tuanny Barcellos**

Mestre em Engenharia de Produção

Instituição: Universidade do Estado do Rio de Janeiro (UERJ)

Endereço: Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro - RJ

E-mail: tuanybarcellos@id.uff.br

### **RESUMO**

Muitas decisões tomadas em diferentes setores da economia dependem da previsão de séries macroeconômicas. Portanto, a aplicação de técnicas sofisticadas, como redes neurais, tem se mostrado uma estratégia viável para aumentar a precisão das previsões nesses casos e tornar as decisões mais congruentes. Este estudo teve como objetivos a implementação e avaliação de métodos de aprendizado de máquina, com foco em redes neurais, para a previsão de variáveis econômicas selecionadas. Para alcançar este objetivo foram construídos modelos de aprendizado de máquina lstm (long-short term memory). As variáveis selecionadas para a análise foram: a taxa de juros, o pib (sendo utilizado o

produto industrial mensal para o seu cálculo mensal) e a inflação. A série temporal utilizada para a aplicação do modelo está no formato mensal dos períodos de 01/01/2002 até 01/03/2023. Após o treinamento e estruturação do modelo, calculou-se os valores dos erros (mae, mape e mse) do resultado de cada um. Cabe destacar que se testou diversas redes neurais lstm e observou-se que os modelos com múltiplas camadas de entradas possuem maior aderência e geram precisões mais assertivas.

**Palavras-chave:** aprendizado de máquina, redes neurais, variáveis econômicas.

## ABSTRACT

Many decisions taken in different sectors of the economy depend on the forecast of the macroeconomic series. Therefore, applying sophisticated techniques, such as neural networks, has proven to be a viable strategy to increase forecast accuracy in these cases and make decisions more congruent. This study aimed to implement and evaluate machine learning methods, focusing on neural networks, for forecasting selected economic variables. Lstm (long-short term memory) machine learning models were built to achieve this goal. The variables selected for the analysis were: interest rate, gdp (using the monthly industrial product for its monthly calculation), and inflation. The time series used to apply the model is monthly from 01/01/2002 to 03/01/2023. After training and structuring the model, the error values (mae, mape, and mse) of the result of each one was calculated. It should be noted that several lstm neural networks were tested, and it was observed that the models with multiple layers of inputs have greater adherence and generate more assertive precisions.

**Keywords:** machine learning, neural network, economic variables.

## 1 INTRODUÇÃO

A variação, ou volatilidade, dos índices macroeconômicos, como a inflação, é algo natural na economia de um país. Isso ocorre, pois, a constante mudança nas ofertas e demandas impacta diretamente no poder de compra da população. Sendo assim, uma das principais razões para que se implemente a análise e previsões desses índices, é a possibilidade do mercado se preparar para tal variação. A Taxa de Juros (SELIC) e o Produto Interno Bruto (PIB) podem possuir relações lineares e não lineares com a inflação, sendo assim, suas variações podem impactar uns aos outros e, por essa razão, também devem ser observadas.

A taxa de juros é o principal método utilizado pelo governo brasileiro para controlar a inflação. Esse mecanismo tornou-se ainda mais importante quando o Brasil

adotou o sistema de metas inflacionárias em junho de 1999. Este sistema já vinha sendo utilizado na maioria dos países da América Latina. A taxa básica de juros brasileira, conhecida como SELIC, fornece um limite para os juros que os bancos pagam sobre depósitos a prazo. A partir disso, os bancos também determinam a taxa que cobrarão de pessoas que solicitam empréstimos (Gonçalves *et al.* 2017).

Acompanhar a variação da taxa básica de juros pode trazer informações sobre a atividade econômica do país. Em um ambiente cujos juros básicos estão mais baixos, a tendência é que as pessoas consumam mais, e os empresários invistam mais na expansão da produção via elevação de investimento em bens de capital. O inverso também pode acontecer: quando a taxa sofre aumento, os juros são impulsionados e as pessoas tendem a consumir menos. A redução do consumo é um indicador de desaceleração, refletindo no PIB.

O desenvolvimento de redes neurais remonta ao início dos anos 1940 e experimentou um aumento de popularidade no final dos anos 1980. Este é o resultado da descoberta de novas tecnologias, do desenvolvimento da tecnologia de hardware de computador e do progresso geral. As redes neurais são frequentemente usadas para análise estatística e modelagem de dados, em que seu papel é percebido como uma alternativa à regressão não linear ou técnicas de análise de cluster. eles são normalmente usados em problemas que podem ser expressos em termos de classificação ou previsão. Alguns exemplos incluem reconhecimento de imagem e fala, reconhecimento de caracteres textuais e domínios de conhecimento humano, como diagnóstico médico, pesquisa geológica de petróleo e previsão de indicadores do mercado financeiro (GURNEY, 2018).

O presente trabalho propõe analisar as capacidades dos modelos de redes neurais e ressaltar os benefícios que eles proporcionam ao prever as variáveis macroeconômicas mais relevantes de um país. Isto, pois, ao aplicar modelos de redes neurais para previsões de variáveis macroeconômicas, neste caso, a inflação (IPCA), o PIB (a partir do PIM) e a taxa de juros (SELIC), é possível compreender o cenário econômico, a partir de informações matemáticas. Tais informações podem orientar políticas e estratégias de desenvolvimento.

A previsão de séries macroeconômicas é muito importante para as decisões tomadas em várias áreas da economia. Por isso, a implementação de técnicas avançadas, como redes neurais, tem se mostrado um método promissor para aumentar a precisão das previsões nesses casos. Isto é, devido à habilidade que as redes neurais têm de lidar com volumes extensos de dados e o processamento de informações não lineares, ela poderá capturar relações sutis e interdependências entre as variações dos diferentes indicadores econômicos. Na macroeconomia moderna, a importância das previsões de inflação não pode ser exagerada, dado o seu papel proeminente em muitas situações práticas (THEOHARIDIS *et al.* 2023).

A relevância do presente tema se dá em razão da importância em aprimorar os modelos de previsão de índices macroeconômicos e demonstrar a sua utilidade e acurácia. A partir disso, seria possível monitorar os comportamentos desses índices e aplicar os métodos de previsão com os modelos *LSTM*. A inflação, por exemplo, é um índice muito importante, pois influencia nas decisões do poder público, que irá responder com políticas monetárias. Assim como influência no setor privado, que precisa periodicamente atualizar seus preços em razão da variação desse índice. Além disso, a população em geral é influenciada pela inflação, pois ela irá impactar no rendimento de sua poupança ou investimentos e, ainda, irá determinar o poder de compra no de todos os setores econômicos.

A abordagem de previsão realizada a partir de redes neurais possui diversos benefícios. O primeiro deles é quanto à possibilidade de uma análise mais abrangente e em tempo real, o que irá oferecer prontamente informações valiosas para o tomador de decisões. Além disso, no modelo de redes neurais em que é utilizada uma ampla gama de variáveis macroeconômicas, é possível capturar nuances e correlações que outros métodos ou modelos de previsões não conseguiriam detectar. Uma outra vantagem importante é a sua capacidade de adaptação. Isto é, os modelos podem ajustar-se de forma automática aos novos dados incluídos em suas bases, utilizando tais informações para aprendizagem. Sendo assim, tal método é muito eficiente quando se observa a existência de um ambiente econômico tão complexo e em constante mudança.

Para estruturar o presente estudo, apresentaram-se os indicadores econômicos. Posteriormente, versou-se sobre os dados utilizados para realizar o treinamento dos modelos de redes neurais, no intuito de abordar a história de sua invenção, seu propósito e seus principais modelos. Na sequência, os resultados e as conclusões foram expostos.

O presente trabalho foi realizado usando a linguagem de programação Python e as bibliotecas Numpy (para o tratamento dos dados extraídos), Keras (para funções de ativação das redes neurais e das camadas de neurônios das *LSTM*) e o Mathplotlib (para os gráficos das análises dos dados). Conforme a metodologia proposta, presentes nos tópicos seguintes, foram desenvolvidos modelos de previsão para os indicadores econômicos PIB, IPCA e SELIC. Após o tratamento dos dados, foi criada uma arquitetura de rede neural *LSTM* em Python para realizar a aprendizagem e previsões com base nos dados fornecidos. Ao final, apresentou-se os resultados da pesquisa, demonstrando os dados, a criação dos modelos, assim como a demonstração dos parâmetros de métricas para avaliar a acurácia da previsão.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neurocientistas e psicólogos estão interessados em redes porque as redes são modelos computacionais de cérebros de animais desenvolvidos pela abstração de propriedades de tecido neural real, consideradas necessárias para o processamento de informações. Os neurônios artificiais costumam ser versões extremamente simplificadas de suas contrapartes biológicas GURNEY (2018).

Um neurônio pode resolver problemas bastante simples e uma coleção de neurônios conectados de uma certa maneira pode resolver problemas bastante complexos. O tipo mais simples de rede neural é uma rede perceptron de camada única, que consiste em uma única camada de nós de saída. A entrada é alimentada diretamente na saída através de uma série de pesos. Entre os neurônios de entrada e saída podem ser definidas camadas adicionais, chamadas de camadas ocultas. Eles são responsáveis por construir abstrações complexas de alto nível (NAUMETC, 2017).

As Redes Neurais Artificiais (RNAs) — uma forma de modelos computacionais não algorítmicos caracterizados por sistemas que em algum grau se assemelham à



estrutura do cérebro humano — ganham relevância devido ao seu potencial para gerar previsões mais assertivas. As RNA são sistemas distribuídos paralelos compostos por unidades de processamento simples – geralmente chamadas de neurônios artificiais em analogia aos neurônios biológicos – dispostas em uma ou mais camadas e interligadas por múltiplas conexões (Siqueira *et. al.* 2014).

Inicialmente, inovando nessa área de pesquisa, Tkacz e Hu (1999) utilizaram Redes Neurais (RNAs) para prever o crescimento da produção e compararam o resultado obtido com modelos tradicionais. Os autores investigaram a capacidade das RNA's de detectar padrões complexos e não lineares encontrados em dados econômicos. A partir deles, percebe-se que, depois do quarto horizonte, as RNAs fazem melhores previsões do PIB. Isso fornece uma alternativa promissora para a precisão das previsões macroeconômicas. Ainda em relação aos seus estudos, as RNAs utilizadas demonstraram ser eficazes na modelagem das relações de curto e longo prazo. Nesse sentido, a capacidade de capturar dinâmicas temporais é o principal para a previsão precisa dos índices macroeconômicos, uma vez que a economia é influenciada por fatores interconectados e evolui ao longo do tempo.

Para entender a evolução dos estudos na área de *Machine Learning*, é interessante apontar a pesquisa realizada no início do século XXI por Zhang *and* Berandi (2001). Tal pesquisa propõe a combinação de diferentes modelos de redes neurais para aprimorar o desempenho de um “modelo único”, com a finalidade de realizar uma predição da taxa de câmbio. Os resultados demonstram que, ao combinar os modelos de redes neurais baseadas em diferentes partições do banco de dados, tem-se uma maior acurácia do que utilizando os modelos individuais. Porém, Zhang *and* Berandi (2001) explicitam que, apesar do modelo combinado deles ter vantagens se comparados com os métodos padrões de previsão, ele não tem uma melhora significativa, quanto a predição da taxa de câmbio, nas comparações realizadas a partir de testes com o modelo de *baseline Random Walk*, isto é, um grupo controle que tem como variável a aplicação de um processo discreto, mas puramente aleatório. Sendo assim, foi constatado que os valores apontados na previsão do modelo combinado não eram suficientemente distintos do modelo de *baseline Random Walk*, então, foi possível inferir que a predição não ocorreu de forma

suficientemente satisfatória. É importante ressaltar que a previsão de variáveis macroeconômicas é fundamental para desenvolver uma visão sobre as perspectivas econômicas de um país, por isso, o estudo de modelos utilizando o aprendizado de máquinas, para a aperfeiçoar, é tão importante.

A maioria dos modelos de previsão tradicionais depende do ajuste das informações para obter uma relação pré especificada entre dados de entrada e variáveis de saída, assumindo, assim, um processo funcional e estocástico específico subjacente a esse processo (JUNG, *et al.* 2018). Em uma nova abordagem para previsão de variáveis macroeconômicas, foram empregados uma série de aprendizado de máquina com algoritmos, isto é, um método baseado em dados, que impôs restrições limitadas sobre a natureza da verdadeira relação entre as variáveis de entrada e saída. Jung *et. al.* aplica os conceitos de *Elastic Net*, *SuperLearner* e Algoritmos recorrentes de rede neural em dados macroeconômicos. Os resultados obtidos são que o aprendizado de máquina pode fazer previsões mais precisas para variáveis econômicas como inflação, desemprego e crescimento do PIB. Esses modelos também podem ser úteis na captura de mudanças estruturais e no processamento de dados de séries temporais complexas.

As aplicações de redes neurais recorrentes (RNN), presentes nos modelos LSTM, são importantes para a previsão da inflação e são discutidas no artigo de Almosova e Andresen (2023), publicado no *Journal of Forecasting*. Eles examinam a capacidade das RNNs de detectar a não linearidade encontrada nos dados econômicos. Além disso, os autores também examinaram seu desempenho em comparação com outros modelos de previsão. Em suas pesquisas, eles apontam que, em comparação com outros modelos convencionais, como modelos ARIMA e redes neurais *feedforward*, as RNNs mostraram-se melhores na previsão da inflação. As RNNs podem capturar a dependência temporal das séries, lidar com relações não lineares e modelar a dinâmica da inflação ao longo do tempo de forma eficaz. Por essa razão, o presente estudo opta por utilizá-las nos testes dos modelos de LSTM.

Em continuidade ao tema proposto, Araujo e Gaglianone (2023) utilizaram em seu artigo métodos de aprendizado de máquina na previsão da inflação no Brasil, comparando-os com modelos clássicos. O estudo examina diversos algoritmos de

aprendizado de máquina e como eles podem capturar a dinâmica complexa dos dados econômicos do Brasil, demonstrando que a *Machine Learning* consistentemente possui métricas melhores do que o modelo autorregressivo de médias móveis e, em muitos casos, conseguem alcançar ganho de dois dígitos de acurácia.

### 3 ÍNDICES ECONÔMICOS

Inicialmente, trataremos nesse texto o conceito de inflação, isto é, o aumento geral e sustentado dos preços dos bens e serviços em uma economia ao longo de um ano. Esse indicador é exposto pelo IBGE, com frequência mensal, em seus relatórios, mostrando se a moeda brasileira está desvalorizada dentro do seu próprio país. A principal métrica do cálculo da inflação é o Índice Nacional de Preços ao Consumidor Amplo (IPCA). Tal índice tem como função medir a variação de preços de uma cesta de produtos e serviços consumidos pela população do país. Essa cesta é definida pela Pesquisa de Orçamento Familiares do IBGE, encontrada no site do próprio instituto. A pesquisa trata-se da análise do que a população consome e quanto é gasto, proporcionalmente à renda mensal da família, em cada item da cesta. Sendo assim, é possível determinar quais produtos impactam mais as despesas do consumo cotidiano do brasileiro. Portanto, é de suma importância monitorar este indicador para diagnosticar a situação do país. Ademais, o mercado observa esse indicador para as suas decisões estratégicas, ao analisar os riscos ou oportunidades que surgem ao longo dos ciclos econômicos.

Ainda, definido pelo Banco Central, a taxa de juros (SELIC) é a remuneração do retorno do capital realizado em investimentos. Também podemos explicá-lo como um incentivo para àquele que posterga seu consumo e usa o valor para emprestar. Ademais, a determinação da taxa de juros ocorre no mercado monetário, resultando de um equilíbrio entre a oferta e a demanda de dinheiro, ou, então, de um equilíbrio entre a quantidade de moeda e da preferência pela liquidez (KEYNES, 1936).

A taxa de juros pode ser dividida em duas categorias: nominal e real. O primeiro conceito é o valor expresso em moeda corrente do país, sem contar as expectativas da inflação. Já o segundo, é o cálculo contabilizando uma cesta de consumo de bens, no qual realizamos a taxa de juros dividida pela expectativa de inflação do período seguinte. A



taxa de juros real é calculada subtraindo a taxa de inflação da taxa de juros nominal. Já a taxa de juros nominal, é a taxa de juros observadas em determinados investimentos ou empréstimos.

Por fim, o Produto Interno Bruto (PIB) representa a soma de todos os bens e serviços finais produzidos numa determinada região durante um período determinado (PIRES *et. al* 2023). Para evitar dupla contagem, é preciso excluir do cálculo os bens de consumo intermediários gerados através da cadeia de produção. Isto é necessário, por exemplo, para evitar a repetição na contagem do valor do trigo, que é contabilizado como bem intermediário na produção de farinha e como bem final na produção de pães, o que resultaria em uma superestimação do valor da atividade econômica (MANKIW, 2023).

Destarte, o PIB funciona como um monitoramento dos ciclos econômicos de uma sociedade (SICSÚ *et al.*, 2009). Os Bancos Centrais, buscando a estabilidade de preços, analisa esse indicador para a análise das pressões inflacionárias dentro da economia, assim realizando as devidas políticas monetárias e fiscais para equilibrar os indicadores econômicos.

#### 4 METODOLOGIA

Os dados utilizados nesta análise são datados do período de janeiro de 2002 até março de 2023. A definição do período de tempo disponível no banco de dados da máquina se deu em razão das informações disponibilizadas pelos órgãos públicos. Como resultado, foram utilizadas 254 entradas de informações. Tais entradas foram divididas para compor 70% o treino do modelo e 30% em seu teste, por fim, o modelo foi programado para realizar 350 épocas de treinamento.

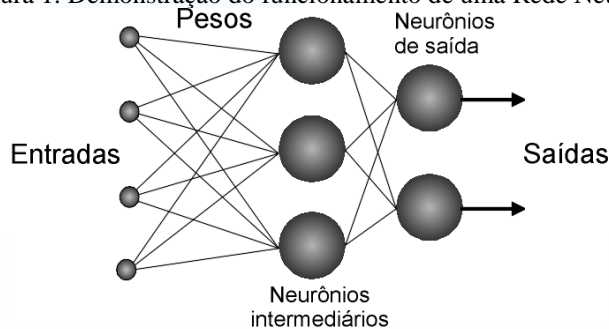
Para a taxa Selic, é possível encontrar através do BACEN (2023) informações mensais desde 26/06/1996 até 01/03/2023, em forma decimal. Já quanto ao IPCA, possuímos as informações mensais desde 01/01/1994 até 01/03/2023, em formato percentual, através do banco de dados do IBGE (2023). Para o PIB, com o intuito de realizar o estudo com dados mensais, iremos utilizar os valores do PIM (produto industrial mensal) e extrair as informações que possuem ajustes sazonais do IBGE, desde

01/01/2002 até 01/03/2023. Por fim, o histórico da taxa do câmbio foi retirado do site do Banco Central com um histórico desde 01/01/1956 até 01/03/2023.

Cabe destacar que existem diversos tipos de aprendizado distintos, dentre os mais utilizados estão o aprendizado Supervisionado: quando os dados de entrada já estão rotulados e mostramos para o modelo o resultado que desejamos. Sendo assim, nesse tipo de aprendizado, os dados de saída e entrada já são previamente especificados, com o objetivo inicial de realizar o treinamento do modelo LSTM. O teste nesse tipo de aprendizado pode ser realizado a partir da comparação entre as previsões e a informação presente nos dados. Sendo assim, é preciso separar uma porcentagem da amostragem para treinamento do modelo e outra para testes, e o Aprendizado Não-Supervisionado: quando estão disponíveis para o modelo apenas dados não classificados, dessa forma, o próprio algoritmo irá realizar a análise e classificação das informações disponíveis.

No presente trabalho, o método de aprendizado supervisionado fora o escolhido. Sendo assim, será fornecido ao modelo os dados já com as suas classificações. Isto é, a partir de um dado de entrada, será especificado para o modelo quais são os dados que devem estar na saída. É importante fazer a ressalva de que aqui estamos tratando apenas do método de aprendizado, sendo assim, a partir da apresentação reiterada dessas informações, o modelo LSTM poderá realizar as suas próprias previsões, tendo em vista a sua função de verificar e aprender os padrões. No final da pesquisa, a acurácia de suas previsões é testada a partir da separação de uma quantidade de dados utilizados apenas para comparações.

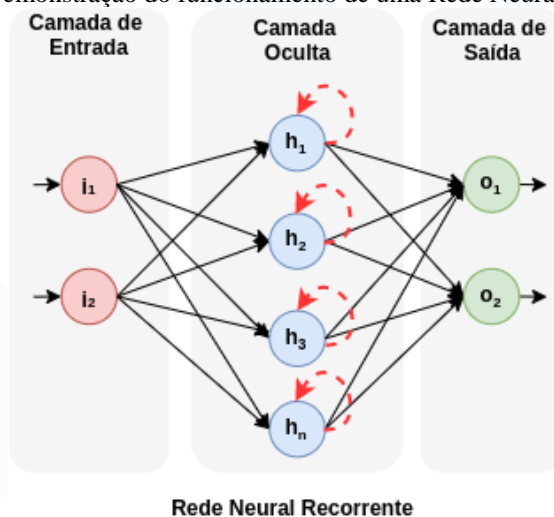
Figura 1. Demonstração do funcionamento de uma Rede Neural.



Fonte: Elaboração própria.

As redes neurais recorrentes (RNN) são um tipo de rede neural que foi desenvolvida para identificar padrões em um banco de dados. Através dos algoritmos da rede neural, ele é capaz de identificar, dentro de uma sequência de informações, similaridades de palavras, de texto, imagens e números, utilizando esses padrões para aprender e futuramente entender o que o está sendo “apresentado”. O aprendizado dessa rede é chamado de recorrente pois ela não toma como entrada apenas os dados inicialmente disponíveis. Desse modo, em seu aprendizado, a *RNN* também é influenciada pelas decisões que elas mesmas tomaram anteriormente. Logo, o algoritmo irá combinar o que foi aprendido e o que foi inserido no passado, concatenando essas duas informações decidir. Essas informações são guardadas no que chamamos de camada oculta da rede. Essa camada funciona como uma memória do modelo, que serve para retroalimentação do neurônio.

Figura 2. Demonstração do funcionamento de uma Rede Neural Recorrente.



Fonte: Barbosa et al. (2021)

Com a utilização dessa rede para analisar dados de longo período, é necessário a implementação de cada vez mais neurônios. É nesse momento que pode ocorrer o problema que chamamos de “Desaparecimento de gradiente”, isto ocorre quando os pesos são multiplicados por números cada vez menores até chegarem em um número bem próximo de zero, resultando em uma dificuldade maior para o treinamento dos neurônios

mais distantes da rede, chegando a um fim no aprendizado do modelo. Para mitigar tal problema, é preciso implementar a Memória LST, presente no tópico seguinte. Memória LST (Long-Short Term) é uma versão sofisticada das RNN (*recurrent neural networks*) desenvolvido por Hochreiter and Schmidhuber (1997). Tal versão foi criada para representar sequências cronológicas mais precisamente do que os outros modelos de redes neurais.

A arquitetura dessa rede é criada de uma forma em que é possível armazenar memórias de curto prazo e de longo prazo. Assim, as memórias de curto prazo atuam quando é realizada a entrada de dados no modelo. Já a memória de longo prazo será frequentemente, frequentemente analisada pelo modelo, apontando se é necessário descartá-la ou manter para utilização futura.

Para comparar o desempenho dos modelos utilizados neste estudo, foram calculados o erro quadrático médio (*MSE*), o erro percentual médio absoluto (*MAPE*) e o erro absoluto médio (*MAE*), disponíveis na tabela abaixo. No erro quadrático médio (*MSE*), pega-se a diferença entre o valor predito pelo modelo e o valor da realidade, eleva-se essa diferença ao quadrado, esse cálculo deve ser feito em todos os pontos, somando-os e dividindo pelo número de pontos utilizados. Já no erro absoluto médio (*MAE*), será medido a diferença absoluta entre valores previstos no modelo e os valores que de fato foram observados na realidade. Por fim, no erro percentual médio absoluto (*MAPE*), também conhecido como desvio percentual médio absoluto, é a expressão em porcentagem da diferença anterior, isto é, a diferença absoluta entre valores previstos no modelo e os valores que de fato foram observados na realidade.

## 5 IMPLEMENTAÇÃO DOS MODELOS

O presente estudo desenvolve modelos de redes neurais recorrentes *LSTM* para prever indicadores econômicos importantes: PIM, IPCA e SELIC. Analisamos três variantes desses modelos: Multivariado, Univariado *Stacked* e Univariado Bidirecional. Através de uma análise comparativa de desempenho, foi possível analisar o modelo com o menor percentual de erros, fornecendo *insights* para futuras previsões e análises econômicas mais precisas.

As redes neurais recorrentes (*RNN*) são um tipo de rede neural que foi desenvolvida para identificar padrões em um banco de dados. Através dos algoritmos da rede neural, ele é capaz de identificar, dentro de uma sequência de informações, similaridades de palavras, de texto, imagens e números, utilizando esses padrões para aprender e futuramente entender o que o está sendo “apresentado”.

Memória *LST* (*Long-Short Term*) é uma versão sofisticada das *RNN* (*recurrent neural networks*) desenvolvido por Hochreiter and Schmidhuber (1997). Tal versão foi criada para representar sequências cronológicas mais precisamente do que os outros modelos de redes neurais. A arquitetura dessa rede é criada de uma forma em que é possível armazenar memórias de curto prazo e de longo prazo. Assim, as memórias de curto prazo atuam no momento em que é realizada a entrada de dados no modelo. Já a memória de longo prazo será frequentemente analisada pelo modelo, apontando se é necessário descartá-la ou manter para utilização futura.

Como supramencionado, a Rede Neural *LSTM* Univariado é um modelo que tem apenas uma série de observação, e a partir disso, o modelo aprenderá o padrão para prever o próximo valor da sequência. Dentro desta classificação, podemos destacar dois outros modelos, o Univariado *Stacked* e o Univariado Bidirecional. O primeiro possui uma arquitetura de mais de uma camada de neurônio escondida, no qual cada camada contém múltiplas células de memória. Já no segundo, são introduzidos dois meios de aprendizagem, um que aprende a sequência da entrada fornecida e o outro que aprende o inverso dessa sequência, propiciando uma maior quantidade de informações para o treinamento.

A segunda classificação de rede neural presente neste artigo é a Multivariado com Séries de Múltiplas Entradas. Esse tipo de modelo utiliza mais de uma variável como entrada para realizar previsões. Ainda, esse modelo será utilizado quando a variável alvo é dependente de mais de uma série temporal.

Em suma, os modelos foram implementados utilizando a linguagem de programação Python e as bibliotecas Numpy (para o tratamento dos dados extraídos) e Keras (para funções de ativação das redes neurais e das camadas de neurônios das *LSTM*). O código foi desenvolvido como uma ferramenta para a criação dos modelos, possuindo



mais de 50 linhas de comando e diversas funções. Nelas, se encontram todos os procedimentos necessários, isto é, a definição de matriz alvo, os parâmetros disponíveis da rede neural, os treinamentos realizados e, ainda, as simulações feitas.

Por fim, deve-se salientar que o processo de desenvolvimento realizado para criar uma rede neural é demorado, pois as informações que serão utilizadas como parâmetros dos modelos deverão ser pensadas e testadas exaustivamente. Além disso, em diversos momentos, é necessário que o projetista do código, a partir de uma análise própria dos dados, defina quais serão as métricas mais importantes, tendo em vista o objetivo do modelo e os resultados dos testes. Dessa forma, aqui será elencado as características que tiveram os melhores resultados. No geral, a partir dos testes que comparam o percentual de erro e acerto, foi observado que os modelos que utilizam quatro camadas de cinquenta neurônios e uma camada densa com um neurônio como *output* obtiveram os melhores resultados, ou seja, menos erros. Além disso, foi observado que o melhor modelo advém do uso das quatro variáveis macroeconômicas aqui elencadas, ao mesmo tempo.

A seguir apresenta-se um passo a passo simplificado da modelagem aplicada no presente artigo. Primeiramente, para prever os índices macroeconômicos alvos, isto é, SELIC, PIM ou IPCA, com base no histórico de dados, foi utilizado o modelo LSTM multivariante, usando todas nossas variáveis macroeconômicas como inputs (Câmbio, Inflação, Selic, PIM e a variável que queremos prever acumulada no ano):

1. Importou-se as bibliotecas necessárias, NumPy, Pandas, Sequential e LSTM Layers e Dense do TensorFlow.
2. Realizou-se o Input do banco de dados.
3. Dividiu-se os dados normalizados em conjuntos de teste e treinamento.
4. Criou-se uma sequência de dados para alimentar a LSTM multivariante.
5. Foram separados os dados para treinamento e teste.
6. Foi construído um modelo LSTM multivariante usando quatro camadas LSTM com ativação ReLU e uma camada densa de saída.
7. Usando o otimizador "Adam" e a métrica de perda "mean\_squared\_error", compilou-se o modelo.

8. Utilizou-se o modelo treinado para os dados de treinamento e teste para fazer previsões.
9. Para obter os valores reais em escala original, foi desnormalizada as previsões.
10. Calculou-se as medidas de erro.

Quadro I: Demonstração do Pseudocódigo Multivariante *LSTM*.

```
# Importar as bibliotecas necessárias
import pandas as pd
import numpy as np
from keras.models import Sequential
from keras.layers import LSTM, Dense
# Criar sequências de dados para alimentar o modelo LSTM
def create_sequences(data, n_steps):
    X, y = [], []
    for i in range(len(data) - n_steps):
        seq_x = data[i : i + n_steps]
        seq_y = data[i + n_steps][-1] # Considerar apenas o último valor do IPCA
        X.append(seq_x)
        y.append(seq_y)
    return np.array(X), np.array(y)
# Carregar e pré-processar os dados do IPCA
data = pd.read_excel("dados.xlsx")
# Pré-processamento dos dados...
# Definir o número de passos (lookback) e features
n_steps = 12 # Número de meses anteriores para a previsão
n_features = 5 # Número de variáveis de entrada (features)
# Dividir os dados em conjuntos de treinamento e teste
X, y = split_sequences(df[['IPCA', 'SELIC', 'PIM', 'CÂMBIO', 'IPCA_ACUMULADO_ANO']],
                        df['IPCA_12_Meses_Shift'], n_steps)
# Normalizar os dados
X_train, X_test = X_train.astype(float), X_test.astype(float)
y_train, y_test = y_train.astype(float), y_test.astype(float)
X_max_train = np.max(X_train)
X_max_test = np.max(X_test)
X_train = X_train / X_max_train
X_test = X_test / X_max_test
# Criar o modelo LSTM multivariante
model = Sequential()
model.add(LSTM(50, activation='relu', return_sequences=True, input_shape=(n_steps, n_features)))
model.add(LSTM(50, activation='relu', return_sequences=True))
model.add(LSTM(50, activation='relu', return_sequences=True))
model.add(LSTM(50, activation='relu'))
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse')
# Treinar o modelo
model.fit(X_train, y_train, epochs=350, batch_size=32, verbose=1)
# Fazer previsões
predictions = model.predict(X_test)
```

```
# Desfazer a normalização das previsões
predictions_denormalized = scaler.inverse_transform(predictions)
# Avaliar o modelo
# MAE, MSE, MAPE
# Fim do pseudocódigo
```

Fonte: Elaboração própria com base nos dados.

## 6 RESULTADO E DISCUSSÃO

Para a compreensão das tabelas 1 e 2, é preciso se atentar para a legenda criada em relação à letra correspondente a cada variável macroeconômica, que são: I, S, P e C, respectivamente, IPCA, SELIC, PIM e Câmbio. Ainda, é preciso relembrar que os índices macroeconômicos utilizados para fins de testes de previsão foram apenas os índices do IPCA, SELIC e PIM. Por fim, cada modelo testado, presentes nas tabelas abaixo, foi realizado de maneira separada, com o objetivo de comparar quais dos três tipos seria mais interessante para o projetista implementar em sua rede neural em uma primeira análise. A distinção dos três modelos é interessante para compreender as possíveis implementações no aprendizado da máquina e seus impactos.

Como pode ser visto na tabela 1, os modelos mais performáticos foram o de tipo Multivariado (com Séries de Múltiplas Entradas), isto é, que possuem os quatro indicadores macroeconômicos, pois apresentaram o menor percentual de erro (*MSE*, *MAE* e *MAPE*) em todas as previsões realizadas nos testes. Nos testes realizados pelo próprio modelo, constata-se que o desvio percentual médio absoluto foi de 16,5% para o IPCA, 11,9% para a SELIC e 6,8% para o PIM. Tais percentuais são suficientemente satisfatórios quando diante de índices tão variáveis ao longo do tempo, pois esses valores são diretamente influenciados por acontecimentos externos e imprevisíveis. Ainda assim, por se tratar de uns métodos precisos, baseados em cálculos matemáticos, os números encontrados poderiam refletir adequadamente o índice esperado em um determinado período de tempo, desde que possuíssem dados anteriores.

Tabela 1 – Métricas de erro do modelo Multivariado.

INPUT	MSE	MAPE	MAE	OUTPUT
I, S, P, C	1,	0,165	0,752	IPCA
I, S, P, C	3,	0,119	1,	SELIC
I, S, P, C	2,	0,068	0,861	PIM

Fonte: Elaborado pelos autores.

Na tabela 2, é possível verificar os resultados dos modelos Univariados *Stacked*. Esses modelos são caracterizados por possuírem apenas uma variável como input e outra como output. Sendo assim, os modelos presentes nos testes englobam todas as 12 combinações entre os índices macroeconômicos. Esse modelo é interessante quando há uma estrutura hierárquica na sequência de dados, isto é, quando a importância de cada dado atribuída pela *LSTM* interfere na qualidade da previsão. A partir da análise da tabela, o melhor resultado desse modelo foi ao utilizar o SELIC de *input* para prever os três diferentes tipos de indicadores, apresentando um *MAPE* de 19,1%, 15,3% e 26,2%, quanto à previsão de IPCA, SELIC e PIM, respectivamente.

Tabela 2 – Métricas de erro do modelo Univariado *Stacked*.

UNIVARIANTE - STACKED				
INPUT	MSE	MAPE	MAE	OUTPUT
IPCA	2,74	0,25	1,27	IPCA
SELIC	1,58	0,19	0,95	IPCA
PIM	3,99	0,28	1,47	IPCA
CÂMBIO	3,43	0,29	1,43	IPCA
IPCA	8,44	0,24	2,43	SELIC
SELIC	4,16	0,15	1,41	SELIC
PIM	26,31	0,43	4,11	SELIC
CÂMBIO	14,19	0,40	3,33	SELIC
IPCA	203,90	0,72	9,64	PIM
SELIC	25,52	0,26	3,57	PIM
PIM	121,89	0,77	7,78	PIM
CÂMBIO	11,62	0,35	2,80	PIM

Fonte: Elaborado pelos autores.

Já na tabela 3, estão expostas as métricas de desempenho dos modelos Univariados bidirecionais, que em vez de treinar um único modelo, introduzimos dois. O primeiro modelo aprende a sequência da entrada fornecida e o segundo modelo aprende o inverso dessa sequência, propiciando um aumento de informações que podem ser utilizadas pelo modelo ao realizar as suas previsões na fase de treinamento. Esse modelo não obteve um percentual de êxito interessante em seus resultados, apresentando um *MAPE* maior do que 19% em todos os testes. Mesmo assim, é importante ressaltar que ele foi utilizado no presente trabalho de forma singular. Além disso, o presente modelo tem como premissa ter um único índice em seu *input*, prevendo outro índice, em seu *output*, o que certamente impactou em seu percentual de acertos. Exemplo disso é que o

modelo de Multivariado com Séries Múltiplas de Entradas, presente na tabela 1, por si só, fora capaz de atingir resultados muito mais interessantes.

Tabela 3 – Métricas de erro do modelo Univariado Bidirecional

UNIVARIANTE - BIDIRECIONAL				
INPUT	MSE	MAPE	MAE	OUTPUT
IPCA	4,310	0,300	1,593	IPCA
SEUC	2,932	0,287	1,366	IPCA
PIM	3,571	0,242	1,430	IPCA
CÂMBIO	3,297	0,262	1,402	IPCA
IPCA	13,077	0,300	2,805	SEUC
SEUC	6,256	0,194	1,942	SEUC
PIM	13,129	0,360	2,965	SEUC
CÂMBIO	16,096	0,376	3,304	SEUC
IPCA	117,195	0,568	7,227	PIM
SEUC	26,296	0,356	4,065	PIM
PIM	18,093	0,257	2,709	PIM
CÂMBIO	32,525	0,201	3,023	PIM

Fonte: Elaborado pelos autores.

Mesmo com uma quantidade seleta de dados disponíveis para o estudo, o modelo elencado como o de menor percentual de erro, presente na tabela 1, demonstrou resultados satisfatórios para a previsão dos indicadores econômicos. É importante ressaltar que o objetivo das comparações acima é justamente responder a indagação de qual seria o modelo mais satisfatório em uma primeira análise. Sendo assim, no momento em que o projetista considerar a implementação de um dos modelos acima elencados, poderia ele preferencialmente considerar o Multivariado com Séries de Múltiplas Entradas. Ao mesmo tempo, se o projetista estiver interessado em mais de uma implementação, é importante que se atente aos outros modelos aqui disponíveis.

Por fim, a previsão de variáveis econômicas com aprendizado de máquina demonstrou poder ser suficientemente acurada para a sua utilização. Sendo assim, a sua implementação e análise tornam-se interessantes para viabilizar tomada de decisões futuras que dependem de tais índices. Entretanto, é preciso ressaltar que por mais matemático e preciso que o modelo seja, nunca será possível descartar a análise de indicadores fáticos e históricos. Então, para a redução dos riscos na análise e previsão de índices macroeconômicos, é preciso que as decisões sejam tomadas juntamente com a observância de demais indicadores técnicos e análise histórica.



## 7 CONCLUSÃO

O presente artigo apresentou modelos de redes neurais artificiais *LSTM* Multivariado, Univariado *Stacked* e Univariado Bidirecional para testar a acurácia da previsão dos indicadores econômicos IPCA, SELIC e PIB, ao longo do tempo. Nesse sentido, as métricas de erro *MAE*, *MAPE* e *MSE* demonstraram que, dentre os modelos acima, o modelo de rede neural artificial *LSTM* Multivariado é o que possui maior desempenho e acurácia.

Assim, ao optar pela implementação dos modelos mencionados, é recomendado ao projetista dar preferência ao Multivariado, inicialmente. No entanto, como foi explicado no tópico acima, é válido considerar alternativas disponíveis caso haja interesse em explorar diferentes abordagens. Além disso, é importante compreender que a análise aqui fora feita de forma singular, mas que seria também interessante a implementação de diversos modelos para realizar uma previsão, a critério do projetista.

Por fim, o artigo avaliou também a possibilidade de implementar modelos *LSTM* na previsão dos índices elencados acima, como forma de auxiliar a tomada de decisão dos setores da economia de um país. Sendo assim, a partir dos testes realizados, de fato há uma acurácia satisfatória dos modelos de *LSTM*. No entanto, é fundamental destacar que, mesmo com a precisão matemática do modelo, a análise de outros indicadores jamais pode ser descartada. Portanto, a fim de reduzir os riscos associados à análise e previsão de índices macroeconômicos, é imprescindível que as decisões sejam baseadas não apenas no modelo de aprendizado de máquina, mas também em conhecimentos técnicos do profissional da área, contexto histórico e a partir da observação de outros índices macroeconômicos.

O presente estudo é uma pesquisa inicial com a finalidade de versar sobre as relações entre as variáveis supracitadas e as obtenções de previsões acertavam das mesmas. A continuação desta pesquisa se dará com a aplicação de modelos estatísticos já consagrados na literatura para que se tenha um *benchmark* e um parâmetro de comparação da robustez e grau de ajuste dos modelos aqui propostos.

## REFERÊNCIAS

ALMOSOVA, A., & ANDRESEN, N. **Nonlinear inflation forecasting with recurrent neural networks**. Journal of Forecasting, n.42(2), p. 240-259, 2022.

ARAÚJO, G. S., & GAGLIANONE, W. P. **Machine learning methods for inflation forecasting in Brazil: new contenders versus classical models**. Latin American Journal of Central Banking, 4(2), 100087, 2022.

Barbosa G., Bezerra Govinda, Medeiros D., Mattos. D, Andreoni M. Segurança em Redes 5G: Oportunidades e Desafios em Detecção de Anomalias e Predição de Tráfego Baseadas em Aprendizado de Máquina - **XXI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais**, 2021.

BRASIL. **Banco Central**. Disponível em: <https://dadosabertos.bcb.gov.br/>. Acesso em: 25 de maio de 2023.

BRASIL. **IBGE**. Disponível em: <https://www.ibge.gov.br/estatisticas/todos-os-produtos-estatisticas.html>. Acesso em: 25 de mai. de 2023.

GONÇALVES, N., MEIRA, L., GUSMÃO, G., CORDEIRO, T., & DIAS, C.. Análise da taxa SELIC, PIB E Inflação no Brasil, no período de 2007 a 2016. e3-**Revista de Economia, Empresas e Empreendedores na CPLP**, 3(2), 75-91. 2017.

GURNEY, Kevin. **An introduction to neural networks**. CRC press, 2018.

HOCHREITER S. AND J. SCHMIDHUBER. **Long Short-Term Memory**. *Neural Computation*, 9(8):1735–1780, 1997

JUNG, J. K., PATNAM, M., & TER-MARTIROSYAN, A. (2018). **An algorithmic crystal ball: Forecasts-based on machine learning**. International Monetary Fund, 2018

PIRES, JANDRESSON DIAS; DIAS, CAMILA RAFAELA GOMES. Estudo da relação entre o PIB do seguimento primário no ramo agrícola com as principais culturas da agricultura brasileira. **OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA**, v. 21, n. 6, p. 4120-4133, 2023.

KEYNES, J. M. **The General Theory of Employment, Interest and Money**. Nova York: Harcourt Brace, 1953.

MANKIW, N. Gregory. **Princípios de macroeconomia**. [s.l.] Cengage Learning, 2023.

NAUMETC, Daniil. **Building the artificial neural network environment: artificial neural networks in plane control**. 2017.

SICSÚ JOÃO; ARMANDO CASTELAR PINHEIRO. **Sociedade e economia: estratégias de crescimento e desenvolvimento.** Editora: Brasília: Ipea, 2009.

SIQUEIRA-BATISTA, R., VITORINO, R. R., GOMES, A. P., OLIVEIRA, A. D. P., FERREIRA, R. D. S., ESPERIDIÃO-ANTONIO, V., ... & CERQUEIRA, F. R.. Artificial neural networks and medical education. **Revista Brasileira de Educação Médica**, 38, 548-556. 2014..

THEOHARIDIS, Alexandre Fernandes; GUILLÉN, Diogo Abry; LOPES, Hedibert. Deep learning models for inflation forecasting. **Applied Stochastic Models in Business and Industry**, 2023.

TKACZ, G., & Hu, S. **Forecasting GDP growth using artificial neural networks.** Bank of Canada. 1993,

ZHANG G. P. AND V. L. BERARDI. Time series forecasting with neural network ensembles: an application for exchange rate prediction. **Journal of the Operational Research Society**, 2001.