

UNIVERSIDADE FEDERAL DE ALAGOAS
CAMPUS SERTÃO
ENGENHARIA DE PRODUÇÃO

JAILSON DOS SANTOS VIEIRA

PREVISÃO DA UTILIZAÇÃO DO AMBULATÓRIO DE UMA FÁBRICA DE
CALÇADOS POR MEIO DE REDES NEURAIIS RECORRENTES LSTM.

Delmiro Gouveia - AL

2023

JAILSON DOS SANTOS VIEIRA

PREVISÃO DA UTILIZAÇÃO DO AMBULATÓRIO DE UMA FÁBRICA DE
CALÇADOS POR MEIO DE REDES NEURAIIS RECORRENTES LSTM.

Trabalho de Conclusão de Curso apresentado
ao Curso de Engenharia de Produção da
Universidade Federal de Alagoas, como
requisito parcial à obtenção do título de
Bacharelado em Engenharia de Produção.

Orientador: Prof. Dr. Jonhatan Magno Norte da Silva

Delmiro Gouveia - AL

2023

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca do Campus Sertão
Sede Delmiro Gouveia

Bibliotecária responsável: Renata Oliveira de Souza CRB-4/2209

V657p Vieira, Jailson dos Santos

Previsão da utilização do ambulatório de uma fábrica de calçados por meio de redes neurais recorrentes LSTM / Jailson dos Santos Vieira. – 2023.

61 f. : il.

Orientação: Jonhatan Magno Norte da Silva.

Monografia (Engenharia de Produção) – Universidade Federal de Alagoas. Curso de Engenharia de Produção. Delmiro Gouveia, 2023.

1. Engenharia de produção. 2. Previsão de demanda. 3. Ambulatório. 4. Long Short-Term Memory – LSTM. 5. Insumos. 6. Redes neurais. 7. Saúde ocupacional. I. Silva, Jonhatan Magno Norte da. II. Título.

CDU: 658.52.011.56

Folha de Aprovação

JAILSON DOS SANTOS VIEIRA

PREVISÃO DA UTILIZAÇÃO DO AMBULATÓRIO DE UMA FÁBRICA DE
CALÇADOS POR MEIO DE REDES NEURAIS RECORRENTES LSTM.

Trabalho de Conclusão de Curso submetido
à banca examinadora do curso de
Engenharia de Produção da Universidade
Federal de Alagoas e aprovada em 25 de
maio de 2023.

Banca examinadora:



Documento assinado digitalmente
JONHATAN MAGNO NORTE DA SILVA
Data: 25/05/2023 16:48:52-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Jonhatan Magno Norte da Silva – UFAL/Sertão



Documento assinado digitalmente
MANOEL GERONIMO LINO TORRES
Data: 29/05/2023 16:45:37-0300
Verifique em <https://validar.iti.gov.br>

Prof. M.Sc. Manoel Gerônimo Lino Torres – UFAL/Sertão



Documento assinado digitalmente
IGOR EDUARDO SANTOS DE MELO
Data: 25/05/2023 20:36:50-0300
Verifique em <https://validar.iti.gov.br>

M.Sc, Igor Eduardo Santos de Melo –
PPGEP/UFRGS

Dedico este trabalho aos meus pais, minha
família e amigos que sempre me incentivaram.

AGRADECIMENTOS

Agradeço aos meus professores e colegas por me ajudarem a desenvolver este trabalho.

Epígrafe

“Estude o passado, aprenda com o presente e preveja o futuro.” (Vieira, J. S.).

RESUMO

A previsão de demanda é um problema crítico para a gestão dos recursos em uma fábrica, especialmente no que diz respeito aos insumos e capacidade do ambulatório. O objetivo desse trabalho foi desenvolver um modelo de aprendizado de máquina capaz de estimar a demanda do atendimento médico dos funcionários da fábrica de grande porte localizado na região nordeste do Brasil, considerando a quantidade de atendimentos no ambulatório da fábrica. A metodologia utilizada consiste em coletar e tratar os dados históricos do ambulatório, aplicar a técnica de LSTM para treinar e testar o modelo, e avaliar o desempenho do mesmo por meio de métricas como erro quadrático médio. Os resultados obtidos indicam que o modelo proposto é capaz de prever com certa precisão a utilização do ambulatório, podendo auxiliar na gestão dos recursos humano. O trabalho contribui para o avanço do conhecimento na área de inteligência artificial aplicada à saúde ocupacional.

Palavras-chave: previsão; ambulatório; redes neurais; LSTM; saúde ocupacional.

ABSTRACT

Demand forecasting is a critical problem for resource management in a factory, especially regarding inputs and the capacity of the clinic. The objective of this study was to develop a machine learning model capable of estimating the demand for medical care for employees at a large-scale factory located in the northeast region of Brazil, considering the number of visits to the factory clinic. The methodology used consists of collecting and processing historical data from the clinic, applying the LSTM technique to train and test the model, and evaluating its performance using metrics such as mean squared error. The obtained results indicate that the proposed model is capable of predicting the clinic's utilization with a certain level of accuracy, thus assisting in human resource management. This study contributes to the advancement of knowledge in the field of artificial intelligence applied to occupational health.

Keywords: Forecasting; Clinic; Neural networks; LSTM; Occupational health.

Lista de Ilustrações

Figura 1 — Representação de um neurônio artificial.....	20
Figura 2 — Função Sigmoide.....	21
Figura 3 — Função Tangente Hiperbólica	21
Figura 4 — Rede Neural Recorrente Simples	22
Figura 5 — Célula de memória de uma rede LSTM	23
Figura 6 — Visitas ao ambulatório para o ano de 2018	29
Figura 7 — Visitas ao ambulatório para o ano de 2019	29
Figura 8 — Gráfico dos outliers	30
Figura 9 — Dados de treinamento normalizados	30
Figura 10 — Transformação de uma série temporal em janelas temporais	27
Figura 11 — Função de autocorrelação para determinar janela temporal.....	31
Figura 12 — Perdas durante o treinamento do modelo	34
Figura 13 — Previsão de 146 dias.....	32
Figura 14 — Primeira parte da previsão dos 146 dias.....	32
Figura 15 — Segunda parte da previsão de 146 dias	33
Tabela 1 — Visitas ao Ambulatório	38

LISTA DE ABREVIATURAS E SIGLAS

ARIMA	AutoRegressive Integrated Moving Average
COVID-19	Coronavírus SARS-CoV-2
EAM	Erro Absoluto Médio
EQM	Erro Quadrático Médio
LSTM	Long Short-Term Memory
RNA	Redes Neurais Artificiais
RNR	Redes Neurais Recorrentes
TangH	Tangente Hiperbólica
TCC	Trabalho de Conclusão de Curso

SUMÁRIO

1	INTRODUÇÃO	13
1.1	MOTIVAÇÃO	14
1.2	JUSTIFICATIVA.....	15
1.3	OBJETIVOS.....	16
1.3.1	OBJETIVO GERAL	16
1.3.2	OBJETIVOS ESPECÍFICOS.....	16
1.4	DELIMITAÇÃO DO ESTUDO	16
1.5	ESTRUTURA DO TRABALHO DE CONCLUSÃO DE CURSO	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	INTRODUÇÃO À SÉRIES TEMPORAIS.....	18
2.2	MÉTODOS TRADICIONAIS PARA PREVISÃO DE SÉRIES TEMPORAIS ...	18
2.3	REDES NEURAIS ARTIFICIAIS	19
2.3.1	FUNÇÕES DE ATIVAÇÃO	20
2.3.2	REDES NEURAIS RECORRENTES	22
2.3.3	Rede LSTM (Long Short-Term Memory)	22
2.4	MEDIDAS DE AVALIAÇÃO.....	24
2.4.1	ERRO QUADRÁTICO MÉDIO	24
3	METODOLOGIA	25
3.1	BASE DE DADOS.....	25
3.1.1	DADOS DE TREINAMENTO E TESTE	25
3.1.2	NORMALIZAÇÃO	26
3.1.3	JANELA TEMPORAL	26
3.1.4	AUTOCORRELAÇÃO	27
3.2	HIPERPARÂMETROS UTILIZADOS NO MODELO DE TREINAMENTO	27
4	RESULTADOS.....	29
4.1	ANÁLISE DOS DADOS.....	29
4.2	MODELO DE PREVISÃO CONSTRUÍDO	31
4.3	PRECISÃO DO MODELO DE PREVISÃO.....	33
5	CONCLUSÃO	35
	REFERÊNCIAS	36
	APÊNDICE A — Dados das visitas ao ambulatório de uma fábrica de calçados entre janeiro de 2018 e dezembro de 2019	38
	APÊNDICE B — Algoritmo em linguagem de programação python utilizado para a previsão de redes neurais recorrentes LSTM	56

1 INTRODUÇÃO

A previsão de demanda é um problema crítico para a gestão de recursos em uma fábrica, especialmente no que diz respeito às visitas ao ambulatório. A capacidade de prever o número de visitas ao ambulatório permite que as empresas possam garantir a disponibilidade de recursos, tais como profissionais de saúde (médicos, enfermeiros etc.), medicamentos diversos (para dor, enjoo, cicatrização etc.), matérias de primeiros socorros básicos (gaze, curativos etc.), além de reduzir os tempos de espera para os funcionários que demanda algum atendimento em situação de desconforto.

Com a necessidade de adicionar novas ferramentas para auxiliar na eficiência de tratamentos médicos, foi realizado um trabalho de previsão para estimar o número de casos de dengue no município de Cascavel-PR, utilizaram além do número de casos, informações climáticas para complementar as estimativas e que foi aplicado metodologias estatísticas e de aprendizagem de máquina (RNR). Obtendo resultados promissores para estimar os números de casos (Orssato et al.,2022).

É senso comum que atualmente os profissionais das mais diversas áreas estão demandando conhecer ferramentas e metodologias para o tratamento de dados. Neste sentido diversas ferramentas estatísticas e computacionais possibilitaram uma melhor análise e compreensão de diversos tipos de dados. Dentre diversas ferramentas destaca-se aqui as redes neurais.

É senso comum que atualmente os profissionais das mais diversas áreas estão demandando conhecer ferramentas e metodologias para o tratamento de dados. Neste sentido diversas ferramentas estatísticas e computacionais possibilitaram uma melhor análise e compreensão de diversos tipos de dados. Dentre diversas ferramentas destaca-se aqui as redes neurais.

Neste trabalho, propõe-se utilizar redes neurais recorrentes (RNR) para prever o número de visitas ao ambulatório de uma fábrica. As RNR são uma classe de modelos de redes neurais que possuem conexões de retroalimentação e são capazes de lidar com sequências de dados, o que as torna ideais para prever eventos futuros com base em informações históricas. A base de dados utilizada será composta por informações sobre o número de visitas ao ambulatório.

Com o grande volume de dados obtidos com a pandemia de COVID-19 no Brasil, um trabalho de Araújo (et al.,2021) investigou a efetividade das redes neurais em prever a perda de vidas durante a pandemia. Resultando em um modelo com uma boa eficácia geral

que pode estimar os números da pandemia do COVID-19 no Brasil (Araújo et al., 2021).

Utilizando redes neurais LSTM com objetivo de auxiliar os cálculos previdenciários, Nascimento et al (2022) desenvolveu um modelo LSTM que estima valores para tábuas de mortalidade (distribuição de óbitos e sobreviventes de uma determinada população). Através de sua pesquisa o autor, observou que o modelo pode ser utilizado como ferramenta para auxiliar a construção de tabuas de mortalidade (Nascimento et al., 2022).

Com a necessidade crescente de adotar novas tecnologias para melhorar a eficiência dos serviços de saúde, as redes neurais têm sido amplamente estudadas e aplicadas nesse contexto. Um exemplo disso é o uso dessas técnicas para prever e estimar o número de casos de doenças, como a dengue em Cascavel-PR, e até mesmo para prever a perda de vidas durante a pandemia de COVID-19 no Brasil. Além disso, as redes neurais também podem ser aplicadas em áreas como cálculos previdenciários e construção de tábuas de mortalidade. As vantagens e benefícios encontrados no uso dessas técnicas incluem uma maior precisão na previsão de eventos, como a ocorrência de doenças, bem como uma maior rapidez e eficiência no processamento e análise de grandes volumes de dados. Nesse sentido, é importante continuar explorando e aprimorando o uso de redes neurais em serviços de saúde, a fim de melhorar a qualidade do atendimento e prevenir doenças de maneira mais eficaz.

Neste contexto, o presente trabalho tem a finalidade de testar a capacidade de previsão de redes neurais recorrentes no contexto de previsão de visitas ao ambulatório de uma fábrica. Este trabalho busca contribuir para a área de gestão de recursos em uma fábrica, oferecendo uma metodologia baseada em inteligência artificial para previsão de visitas ao ambulatório.

1.1 MOTIVAÇÃO

Prever a demanda de visitas ao ambulatório de uma fábrica de calçados é importante porque permite que a empresa gerencie seus recursos de forma mais eficiente e tome decisões estratégicas. Por exemplo, a previsão da demanda pode ajudar a empresa a:

- Planejar o número de profissionais de saúde necessários para atender à demanda, evitando problemas de superlotação ou falta de pessoal;

- Identificar tendências e picos na utilização do ambulatório, o que pode indicar problemas de saúde ocupacional ou necessidade de mudanças na organização do trabalho;
- Alocar recursos financeiros e materiais de forma mais eficiente, garantindo que haja sempre suficiente para atender à demanda;
- Melhorar a qualidade do atendimento aos funcionários, evitando longas esperas e garantindo que eles tenham acesso aos tratamentos de saúde de que precisam.

1.2 JUSTIFICATIVA

A escolha desse tema se baseia na convicção de que a saúde dos trabalhadores desempenha um papel essencial na produtividade e na qualidade de vida nas fábricas. De acordo com a Norma Regulamentadora 4 (NR 4), é exigido que empresas com mais de 50 funcionários possuam um Serviço Especializado em Engenharia de Segurança e Medicina do Trabalho (SESMT), o que pode incluir um ambulatório médico interno. Entretanto, os ambulatórios das fábricas podem enfrentar uma escassez de recursos e insumos para atender à uma demanda dos trabalhadores em situações de acidentes ou doenças ocupacionais.

Um dos recursos mais importantes para o funcionamento de um ambulatório é o planejamento da demanda, ou seja, a previsão da quantidade e do tipo de atendimentos que serão necessários em um determinado período. Isso permite que o ambulatório se organize e evite desperdícios ou escassez de materiais e medicamentos.

Para realizar a previsão da demanda do ambulatório de uma fábrica de calçados, propus utilizar redes neurais recorrentes LSTM, que são um tipo de inteligência artificial capaz de aprender com dados históricos e temporais. As redes neurais recorrentes LSTM têm a vantagem de aprender com os dados ao longo do tempo, considerando os efeitos sazonais, tendenciais e aleatórios que podem influenciar na demanda do ambulatório. Além disso, elas podem se adaptar a mudanças nos dados e incorporar novas informações à medida que elas se tornam disponíveis.

Existe a possibilidade que esse método seja melhor do que outros métodos tradicionais de previsão, como regressão linear, pois ele é mais flexível, robusto e preciso. Além disso, ele pode gerar *insights* sobre os fatores que afetam a demanda do ambulatório, como o perfil dos trabalhadores, as condições de trabalho, as medidas preventivas, entre outros.

Não foi encontrado na literatura estudos que utilizassem redes neurais recorrentes

LSTM para prever a demanda de ambulatórios de fábricas. Portanto, considero que esse estudo tem indícios fortes de ineditismo e relevância científica e social.

1.3 OBJETIVOS

Esse trabalho de conclusão de curso (TCC) pretende atingir os objetivos especificados detalhados a seguir.

1.3.1 OBJETIVO GERAL

O objetivo geral desse trabalho é desenvolver uma ferramenta de previsão da demanda das visitas ao ambulatório de uma fábrica de calçados utilizando redes neurais recorrentes.

1.3.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos para prever as visitas ao ambulatório de uma fábrica de calçados seria:

- Estudo e aplicação de técnicas de redes neurais recorrentes para previsão da demanda futura do ambulatório.
- Avaliação da precisão das previsões obtidas pelo modelo de redes neurais recorrentes.
- Verificação da capacidade do modelo em identificar aumentos ou reduções inesperados na demanda do ambulatório.
- Utilizar o modelo para prever a demanda de visitas ao ambulatório para um período futuro e fornecer recomendações para o gerenciamento dos recursos do ambulatório, com o objetivo de melhorar a eficiência e qualidade dos serviços prestados.
- Utilizar códigos na linguagem *Python* no desenvolvimento de uma ferramenta de previsão, no intuito de melhorar a eficiência e qualidade dos serviços prestados pelo ambulatório da fábrica

1.4 DELIMITAÇÃO DO ESTUDO

Esse Trabalho de Conclusão de Curso se delimitou a analisar o número de visitas que ocorreram no ambulatório de uma única fábrica de calçados de grande porte localizada no

Estado do Nordeste. Portanto, não se pode generalizar os resultados desse Trabalho a outras fábricas semelhantes. Também não são detalhadas informações sobre o tipo de atendimentos realizados, nem sobre os indivíduos que visitaram o ambulatório ou sobre os tipos de tratamentos clínicos administrados pelos profissionais de saúde.

1.5 ESTRUTURA DO TRABALHO DE CONCLUSÃO DE CURSO

Esse trabalho de conclusão está estruturado na forma de cinco capítulos, a saber:

- Capítulo 1 – Introdução: esse mesmo capítulo no qual são apresentados elementos sobre o tema pesquisado, justificativa, objetivos, delimitação e estrutura do TCC;
- Capítulo 2 – Fundamentação Teórica: é apresentado um breve histórico das redes neurais recorrentes e sua evolução ao longo dos anos. Em seguida, será discutido a previsão de visitas ao ambulatório e a importância da gestão de recursos. Depois, será descrito o processo de treinamento, validação e avaliação do modelo de RNR proposto;
- Capítulo 3 – Metodologia: é apresentada a classificação da pesquisa, levantamento dos dados, procedimentos computacionais e como se deu a avaliação do processo de previsão da rede neural;
- Capítulo 4 – Resultados: são apresentados os principais resultados encontrados com a utilização das redes neurais recorrentes;
- Capítulo 5 – Conclusão: são apresentadas as considerações finais do trabalho de conclusão, limitações e sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 AMBULATÓRIOS DENTRO DAS FÁBRICAS

Os ambulatórios das fábricas são espaços destinados para o atendimento médico e de enfermagem dos trabalhadores industriais. Eles fazem parte da medicina corporativa, que é o conjunto de ações de manutenção da saúde física e psicológica dos colaboradores. A medicina corporativa tem como objetivo reduzir riscos, prevenir e controlar doenças, além de socorrer os funcionários em caso de acidentes.

De acordo com as informações encontradas no Instituto INE (2021), os ambulatórios de fábricas podem realizar exames ocupacionais, consultas preventivas, campanhas de conscientização, atendimentos de urgência e emergência, entre outras atividades. Eles contribuem para a saúde e a qualidade de vida dos trabalhadores, bem como para a produtividade e a imagem da empresa no mercado.

2.2 INTRODUÇÃO À SÉRIES TEMPORAIS

De acordo com Morettin e Toloi (2006), uma série temporal é qualquer conjunto de observações ordenadas no tempo.

A série temporal é contínua quando os dados observados são coletados continuamente ao longo do tempo. Quando as observações são registradas em intervalos de tempo iguais, a série é discreta. Assim, a classificação da série temporal não está relacionada ao tipo de variável que está sendo registrada, mas ao intervalo de tempo das observações.

2.3 MÉTODOS TRADICIONAIS PARA PREVISÃO DE SÉRIES TEMPORAIS

Alguns dos métodos tradicionais de previsão de séries temporais segundo Lustosa (2008), incluem:

I. **Média móvel:** A média móvel é uma técnica de processamento de dados usada para suavizar séries temporais, removendo flutuações aleatórias e destacando tendências e padrões subjacentes nos dados. É calculada tomando a média de um número de valores consecutivos de uma série temporal, criando assim uma nova série de valores suavizados. O comprimento da janela de média móvel, ou seja, o número de valores usados para cada cálculo de média, é um parâmetro que pode ser ajustado para obter diferentes níveis de suavização.

II. Suavização exponencial: A suavização exponencial é uma técnica de suavização de séries temporais que busca minimizar a influência de valores pontuais (outliers) ou de flutuações transitórias nos dados. Ela combina uma média móvel com ponderação exponencial, dando mais peso aos valores mais recentes. Cada valor da série é combinado com o valor anterior suavizado, usando uma taxa de suavização (ou fator de suavização) para controlar o peso dado a cada valor. Quanto maior a taxa de suavização, mais influência terão os valores mais recentes na suavização, e menos influência terão os valores mais antigos.

III. Box-jenkins (ARIMA): A metodologia Box-Jenkins é uma abordagem de análise de séries temporais que busca modelar e prever comportamentos futuros em séries temporais. É composta por três etapas principais: identificação, estimação e verificação. A etapa de identificação envolve a seleção de um modelo adequado para a série temporal, baseado na análise de gráficos e estatísticas. A etapa de estimação envolve a estimativa dos parâmetros do modelo selecionado, geralmente por métodos de otimização. Finalmente, a etapa de verificação envolve a avaliação da qualidade do modelo estimado, usando testes estatísticos e análise de resíduos.

2.4 REDES NEURAIS ARTIFICIAIS

Rede Neural Artificial (RNA) é um modelo matemático inspirado na estrutura e no funcionamento do sistema nervoso humano. Elas são compostas por unidades básicas, chamadas de neurônios, que se conectam entre si e formam camadas. Para definir mais precisamente uma rede neural artificial, recorremos a Haykin (2001):

Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro em dois aspectos:

O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem. Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

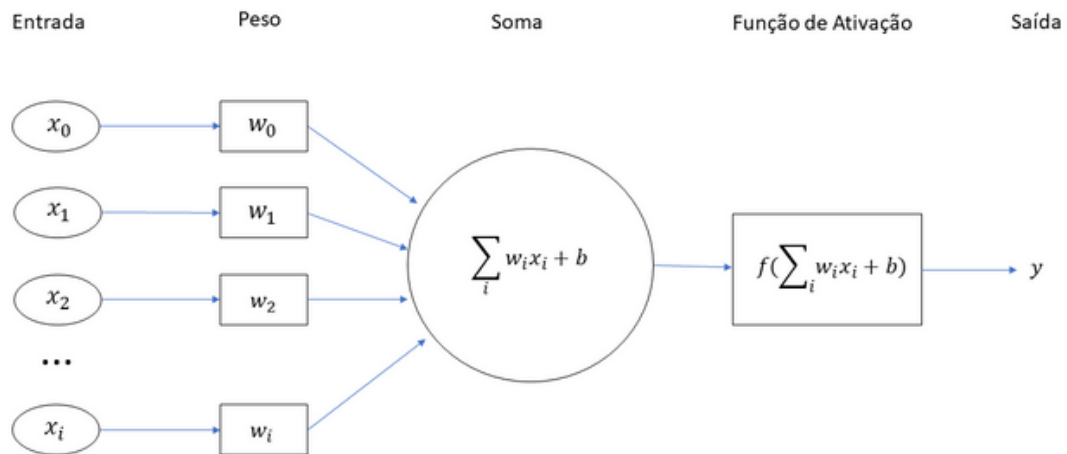
Uma RNA é concebida para detectar uma tendência fundamental em um conjunto de dados e formular uma generalização a partir dela. A maior vantagem de uma RNA, quando comparada a outras técnicas de previsão, é a sua capacidade de aprender os padrões

subjacentes dos dados, onde a maioria dos métodos convencionais tende a falhar (Zhang,1998).

As conexões entre esses neurônios são referidas como sinapses, e cada neurônio da rede recebe informação de outro como entrada através dessas sinapses. Essa informação de entrada é multiplicada por um peso, definido pelo criador da rede, que é utilizado para ajustar a importância de diferentes resultados computacionais obtidos pelo neurônio.

O neurônio artificial, representado na Figura 1, possui 'i' entradas (x_i) e para cada entrada é conectada ao neurônio através de sinapses com pesos (w_i).

Figura 1 — Representação de um neurônio artificial



Fonte: Adaptado de Hiransha.

O neurônio soma as entradas multiplicadas pelos pesos através da equação:

$$A = \sum_i x_i w_i + b$$

Onde A é o resultado da soma e b é o valor do viés. Para calcular a saída, à soma é aplicada uma função chamada de função de ativação (HIRANSHA, 2018).

$$y = f(A)$$

2.4.1 FUNÇÕES DE ATIVAÇÃO

A finalidade da função de ativação é limitar a saída de acordo com o problema proposto. Devido à natureza não-linear dos dados em estudo, sugere-se o uso de funções não-

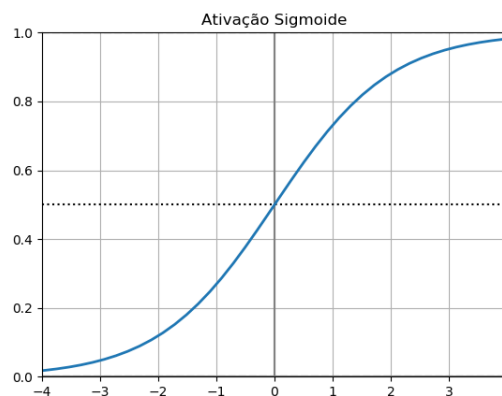
lineares, como a Sigmoide e a Tangente Hiperbólica, que são frequentemente utilizadas para a previsão com séries temporais. Essas funções são continuamente diferenciáveis, o que é uma propriedade desejável durante o treinamento da rede.

- Função Sigmoide

$$f(A) = \frac{1}{1 + e^{-\alpha A}}$$

A função sigmoide restringe os valores de retorno entre zero e um, o α é o parâmetro de inclinação.

Figura 2 — Função Sigmoide



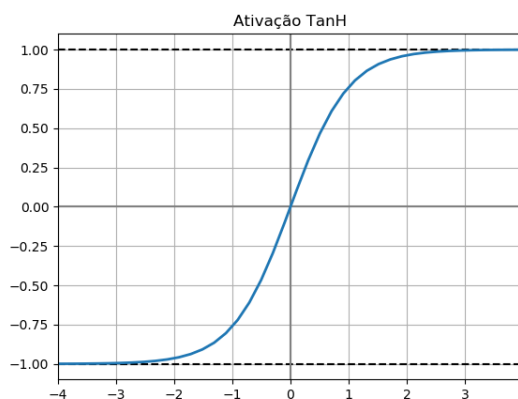
Fonte: elaborado pelo autor (2023).

- Função Tangente Hiperbólica

$$f(A) = \tanh(A)$$

Na função Tangente Hiperbólica, o retorno é expandido para o intervalo de um negativo até o um positivo.

Figura 3 — Função Tangente Hiperbólica

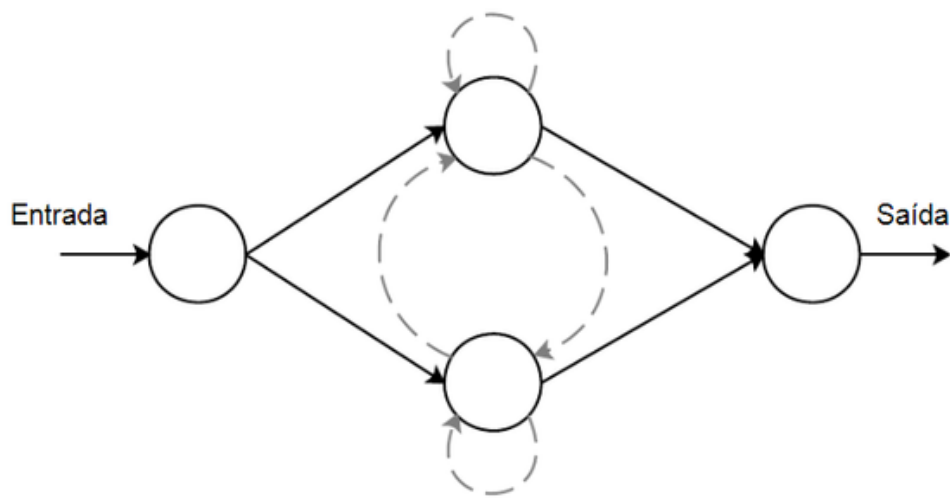


Fonte: elaborado pelo autor (2023).

2.4.2 REDES NEURAIS RECORRENTES

As redes neurais recorrentes (RNRs) são tipos de redes neurais que se diferenciam das redes neurais convencionais por terem uma estrutura de ciclo fechado, permitindo que elas guardem informações do passado. Essa característica as torna adequadas para problemas de sequência, como na previsão de séries temporais, tradução automática e processamento de linguagem natural.

Figura 4 — Rede Neural Recorrente Simples



Fonte: Bitaraes (2019).

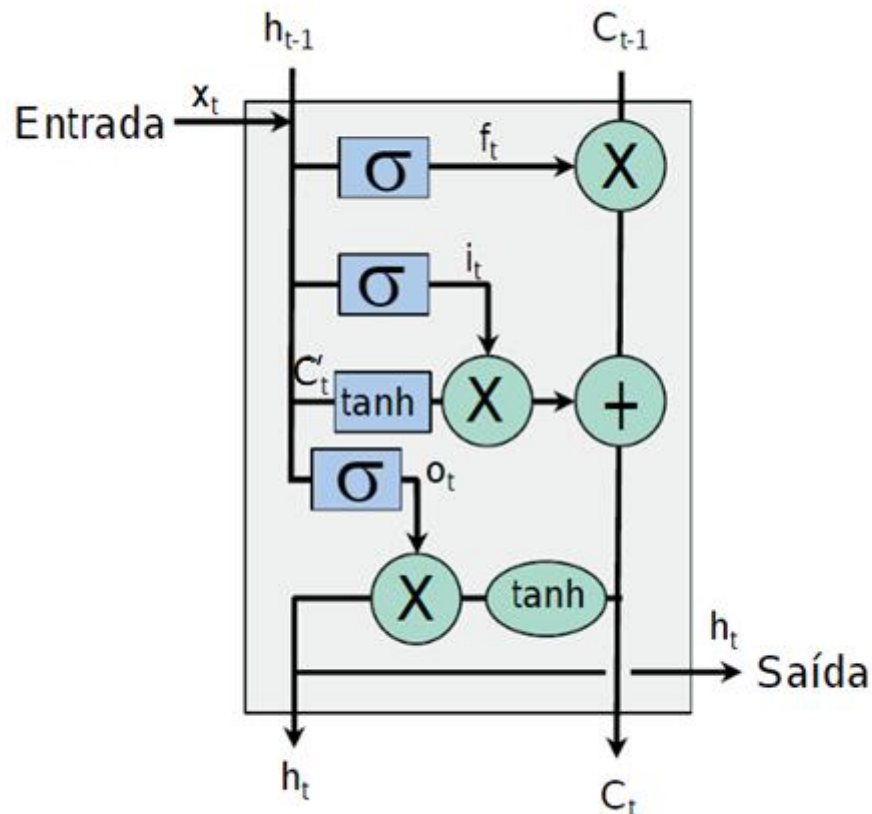
Na rede recorrente ilustrada na Figura 4, a cada momento os sinais são transmitidos ao longo dos caminhos sólidos. Os caminhos traçados ligam o neurônio de origem em cada momento com o neurônio de destino no momento seguinte. Os neurônios com conexões recorrentes recebem a entrada de dados atual, bem como valores dos neurônios do estado anterior da rede. A entrada anterior pode influenciar a saída, e posteriormente, por meio das conexões recorrentes.

2.4.3 Rede LSTM (Long Short-Term Memory)

O modelo LSTM é uma rede neural recorrente e profunda, introduzido por Schmidhuber e Hochreiter em 1997. Ele se assemelha a uma rede neural recorrente convencional com uma camada oculta, mas a diferença é que cada neurônio na camada oculta

é substituído por uma célula de memória. Cada célula de memória possui um neurônio interconectado com peso fixo de um e mecanismos para liberar, adicionar e ler informações na memória.

Figura 5 — Célula de memória de uma rede LSTM



Fonte: Adaptado de Bitaraes (2019).

O primeiro passo decide qual dado na entrada vai ser apagado, os sinais de x_t e h_{t-1} , se o resultado da função sigmoide for zero, o dado será excluído, mas se o resultado for um, o dado não será apagado da memória. O próximo passo é determinar quais novos dados serão armazenados na memória. Esse processo consiste em duas etapas: Em primeiro lugar, uma camada de função sigmoide determina quais valores serão atualizados. Em seguida, uma camada $\tanh$ cria um vetor de novos valores candidatos, C'_t que poderia ser adicionado ao estado. Em seguida, esses dois resultados são combinados para gerar uma atualização para o estado atual.

Nesse passo, o estado da célula C_{t-1} é atualizado para C_t . Os passos anteriores já

determinaram o que deve ser excluído e o que deve ser armazenado, e este passo apenas implementa essas ações. Por fim, uma função sigmoide e a tangente hiperbólica identificam o que é realmente importante para transmitir aos próximos neurônios, agindo como uma espécie de filtro.

2.5 MEDIDAS DE AVALIAÇÃO

A maneira mais comum de avaliar a precisão de um modelo de rede neural que prevê o desempenho futuro de ações é calculando o erro do modelo. Quanto menor o erro, melhor o modelo se adapta aos dados e mais próximas suas previsões estão dos valores reais. Neste trabalho, será apresentada a métrica de Erro Quadrático Médio (EQM).

2.5.1 ERRO QUADRÁTICO MÉDIO

O Erro Quadrático Médio (EQM) é uma medida adicional usada para avaliar a precisão e o erro dos modelos preditivos utilizados no trabalho. A fórmula para o cálculo do EQM pode ser definida como:

$$EQM = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2$$

no qual x_i é o i -ésimo valor real, y_i é o i -ésimo valor predito e n é o número de dados utilizados.

3 METODOLOGIA

Este trabalho tem como objetivo prever a utilização do ambulatório de uma fábrica de calçados por meio de redes neurais recorrentes LSTM. Para isso, foi utilizado dados que tem uma natureza quantitativa e uma abordagem indutiva.

Primeiramente, definimos a base de dados que utilizaremos para treinar a rede neural. Iremos utilizar uma variável numérica dos dados que representa as visitas ao ambulatório para compor a nossa série temporal. Separamos a série temporal em duas, treinamento e teste. Essa divisão é fundamental para permitir que se pudesse avaliar, de forma efetiva, a eficácia da metodologia proposta, bem como identificar erros ou discrepâncias nos resultados encontrados. O código-fonte completo descrito durante as etapas seguintes pode ser encontrado no apêndice B.

3.1 BASE DE DADOS

Os dados selecionados para o desenvolvimento deste trabalho foram cedidos pelo engenheiro responsável por monitorar a utilização do ambulatório da fábrica. Assim, trata-se de um banco de dados fornecidos por uma grande fábrica de calçados localizada na região Nordeste.

O conjunto continha 730 amostras. Os dados foram distribuídos entre as datas de 01 de janeiro de 2018 até 31 de dezembro de 2019. As informações principais contidas no banco de dados se referem ao dia, mês, ano e quantidade de visitas ocorridas nestas datas.

Vale salientar que, em nenhum momento, os indivíduos que fizeram visitas ao ambulatório tiveram suas informações pessoais fornecidas. Assim, não há necessidade de que os dados sejam apreciados pelo Comitê de Ética em Pesquisa.

3.1.1 DADOS DE TREINAMENTO E TESTE

As 730 visitas foram separadas em dois bancos de dados, o primeiro utilizado para treinamento da rede neural; e o segundo para testar os resultados das previsões. Os dados de treinamento foram utilizados para alimentar a rede neural, seguindo uma proporção de 80% de treinamento e 20% de teste (CHOLLET, 2017). No processo de treinamento, são atribuídos pesos as sinapses que conectam os neurônios da rede neural. Tais pesos são corrigidos e modificados pelo processo *backpropagation* (retropropagação), que é responsável por ajustar

o valor dos pesos com a finalidade de reduzir o erro final da rede neural (LILICRAP et al., 2020). Para o caso em questão, reduzir o erro relacionado ao número de visitas ocorridas ao ambulatório.

Já os dados de teste são uma parte do banco de dados separados para verificar o quanto que a previsão da rede neural consegue prever com precisão os valores futuros. Assim, os valores de treinamento são comparados com os valores de teste fornecendo indícios de que a rede neural consegue fazer boas previsões (Bortolot, 2005).

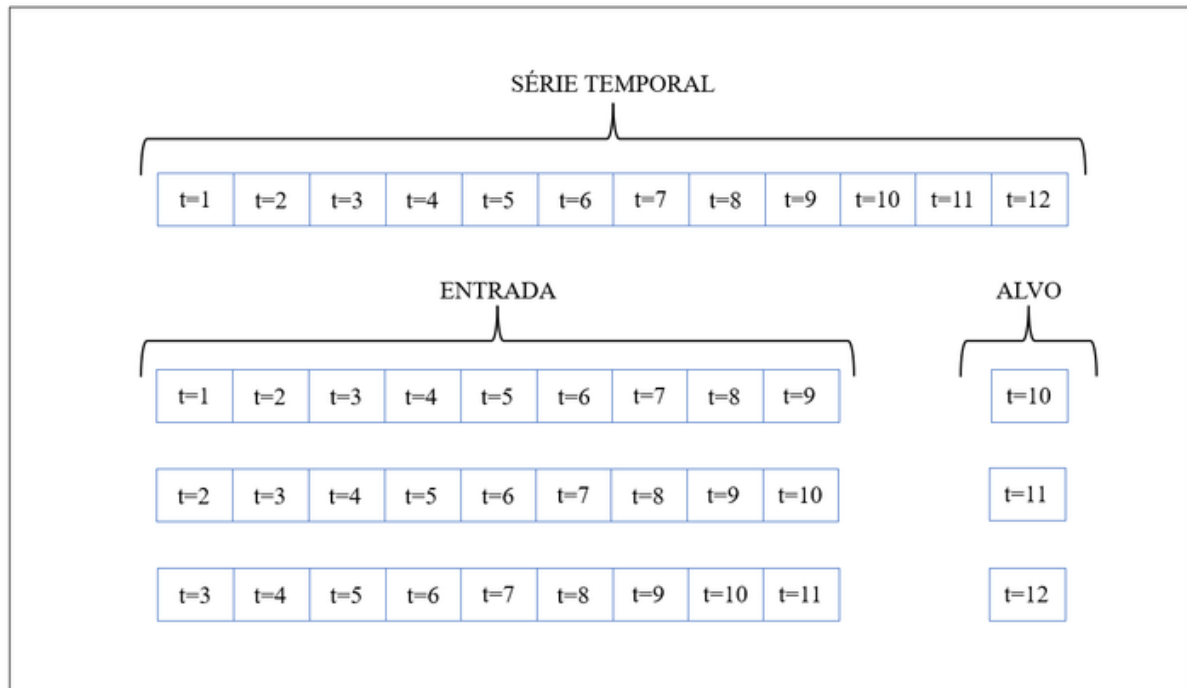
3.1.2 NORMALIZAÇÃO

Como os dados de treinamento que alimentam a entrada da rede neural são multiplicados por pesos sinápticos e somados com a função de ativação, entradas com ordens de grandeza muito maior que outras podem influenciar no sinal de ativação, essa é uma das razões de remover os outliers dos dados de entrada. Como sugeridos por Haykin (2001), para reduzir essa influência no sinal de ativação da rede neural, limitamos a amplitude com a normalização, utilizando intervalos normalizados entre $[0,1]$.

3.1.3 JANELA TEMPORAL

O modelo de redes neurais recorrentes para ser treinado necessita que o conjunto de treinamento seja fragmentado em listas de amostras com valores alvo associados. Para esse processo é utilizado o modelo de janelas temporais (Figura 10).

Figura 10 — Transformação de uma série temporal em janelas temporais



Fonte: Autor (2023).

Desse modo, um dado conjunto de valores alimenta a rede neural recorrente. Em seguida, uma nova observação temporal é acrescentada a série de valores, de tal modo que é verificado se existe um ganho de informação (redução do erro de previsão) com a entrada dessa nova observação. Tal nova observação é incorporada a série temporal, e a informação mais antiga é descartada, para que a próxima observação seja prevista. Esse processo se repete até serem processados todos os dados do banco de treinamento.

3.1.4 AUTOCORRELAÇÃO

Para determinar o tamanho da janela temporal foi utilizado a função de autocorrelação para indicar o valor para fragmentar o conjunto de treinamento. A autocorrelação expressa a dependência dos dados, ou seja, se um determina observação temporal se relaciona com as demais observações temporais presentes na série de dados. De modo geral, o gráfico de autocorrelação e de autocorrelação parcial são utilizados para apresentar a dependência que os dados têm entre si (Casanova et al., 2012).

3.2 HIPERPARÂMETROS UTILIZADOS NO MODELO DE TREINAMENTO

O modelo neural em questão foi construído com cinco camadas distintas de neurônios. A primeira camada, também conhecida como camada de entrada, recebe os dados de treinamento e é composta por 100 neurônios recorrentes. Esses neurônios são projetados para lidar com sequências de dados e têm a capacidade de reter informações de estados anteriores.

A segunda e terceira camadas, que seguem a camada de entrada, são constituídas por 50 neurônios cada, também no modelo recorrente. Essas camadas são responsáveis por processar as informações recebidas da camada de entrada e extrair características relevantes que possam ser usadas para fazer previsões precisas.

A quarta camada, com seus 50 neurônios, atua como uma camada densa que é projetada para agregar as informações coletadas nas camadas anteriores e sintetizar os recursos necessários para a tarefa de aprendizado.

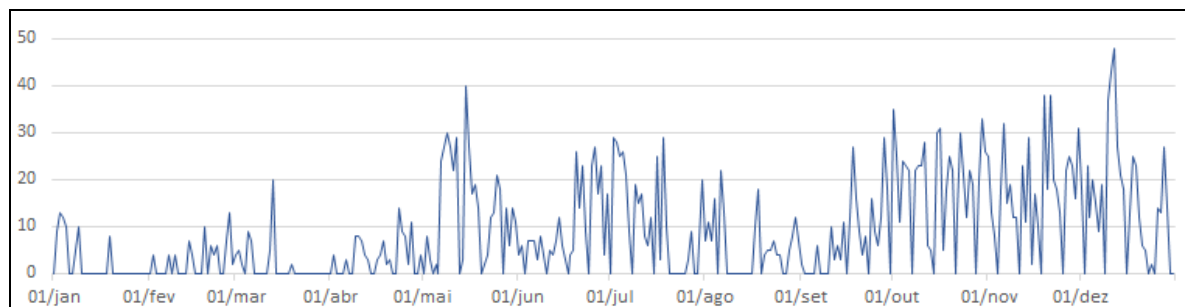
Por fim, a quinta e última camada do modelo é composta por apenas um neurônio e utiliza a função de ativação sigmoide. Essa camada é responsável por fazer a previsão final com base nos recursos agregados nas camadas anteriores.

4 RESULTADOS

4.1 ANÁLISE DOS DADOS

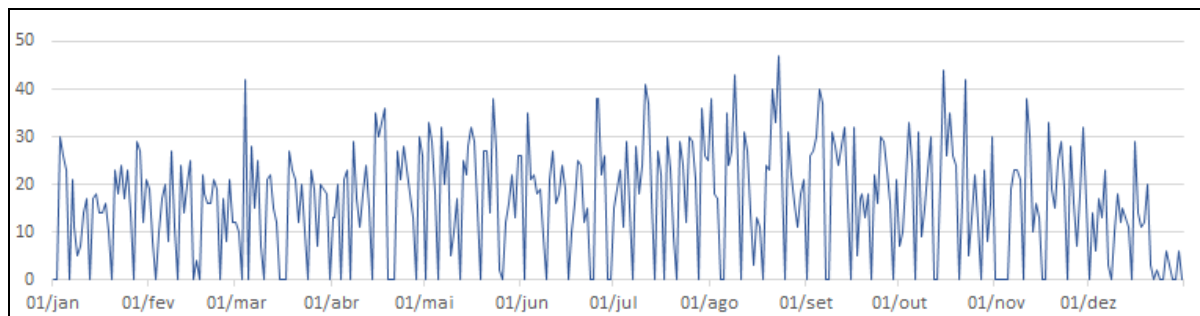
Foram observadas 730 visitas ao ambulatório da fábrica de calçados. Os dados foram plotados em dois gráficos, sendo o primeiro relacionado as visitas ocorridas no ano de 2018 (Figura 6) e no ano de 2019 (Figura 7). Até o início do mês de maio de 2018 poucos são os dias com mais de 10 visitas ao ambulatório. Após o mês de maio de 2018 existe um aumento de dias com mais de 20 visitas ao ambulatório. Após setembro de 2018 existe diversos dias com mais de 30 visitas ao ambulatório. Tal comportamento indica que os trabalhadores passaram a fazer mais visitas com o passar dos dias, provavelmente devido ao desgaste das atividades laborais. Dos Santos Leite et al. (2021) destaca que a indústria de calçados brasileira apresenta um grande de número de riscos ocupacionais. A Figura 7 evidencia que o número de visitas se manteve elevada com a entrada do ano de 2019. De modo geral, o número de visitas ao ambulatório cresceu de forma significativa entre os anos de 2018 e 2019.

Figura 6 — Visitas ao ambulatório para o ano de 2018



Fonte: Autor (2023).

Figura 7 — Visitas ao ambulatório para o ano de 2019

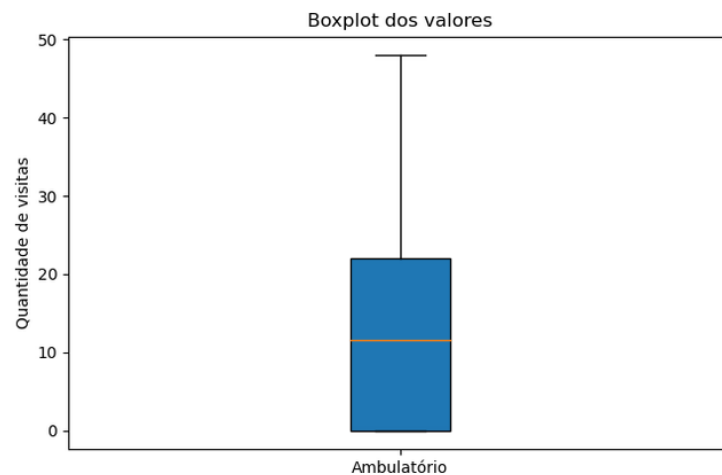


Fonte: Autor (2023).

Devido a essa diferença dos valores das visitas ao ambulatório entre os anos,

buscou-se a presença de outliers nos dados. Verificou-se que não existem outliers nos dados (Figura 8). Assim, nenhuma das observações ficaram fora dos limites delimitados pelo Boxplot. A mediana dos valores ficou um pouco acima de 10 visitas ao ambulatório, e o primeiro e terceiro quartil ficou próximo a zero e um pouco acima de 20 visitas.

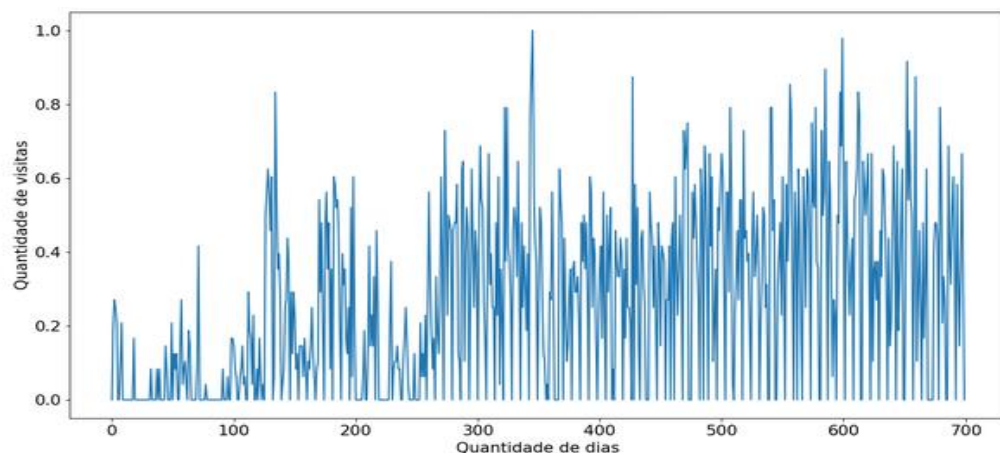
Figura 8 — Gráfico dos outliers



Fonte: elaborado pelo autor (2023).

Como dito, anteriormente, uma proporção de 80% (700 observações da amostra dos dados) foi utilizada como treinamento. Os valores foram normalizados (Figura 9), tal procedimento faz com que os parâmetros da rede neural sejam estimados com menor esforço computacional.

Figura 9 — Dados de treinamento normalizados

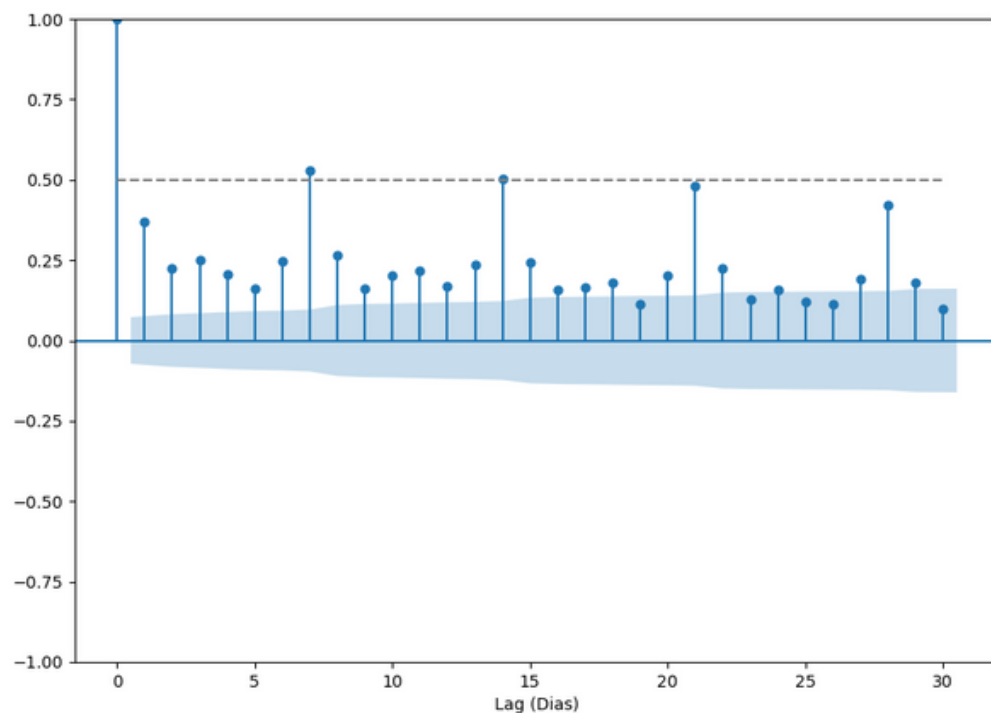


Fonte: Autor (2023).

A partir da Figura 11, foi observado que o valor calculado para compor o gráfico

de autocorrelação atinge ou fica bem próximo da linha tracejada que determina a existência de correlação. Portanto, o número sete (7 dias), que vai ser utilizado como o valor para fragmentar as janelas temporais. A função de autocorrelação calcula o grau de relacionamento linear entre os valores da série temporal (MAIA, 2017). Desse modo, 7 dias é um bom valor, dado que aos finais de semana existe uma pausa na jornada de trabalho e um certo repouso por parte dos trabalhadores que visitam o ambulatório.

Figura 11 — Função de autocorrelação para determinar janela temporal



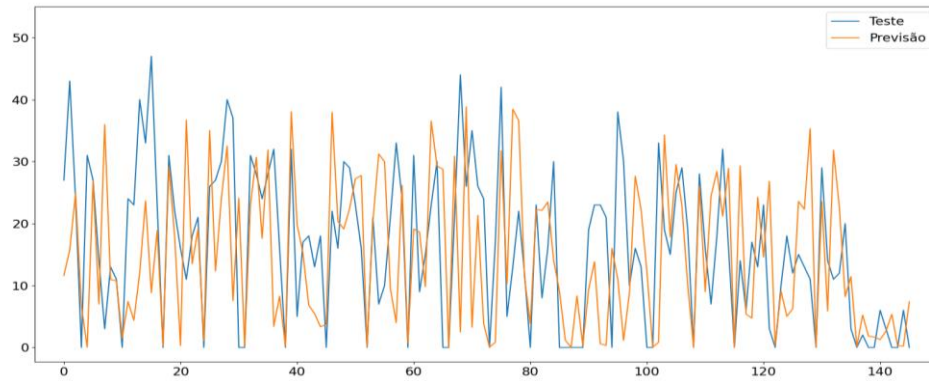
Fonte: elaborado pelo autor (2023).

Conforme ilustrado na figura 11, todos os pontos apresentam valores positivos e indicam uma correlação positiva entre os dados. Considerando o deslocamento temporal de 7 dias, podemos inferir que esses dados possuem uma periodicidade semanal.

4.2 MODELO DE PREVISÃO CONSTRUÍDO

O modelo previu uma sequência de 146 dias que foram plotados juntamente com os dados reais que equivalem ao mesmo período de análise (Figura 12). Para facilitar a visualização dos 146 dias da previsão (dados de teste), o gráfico foi dividido em duas partes com 73 dias cada (Figura 13 e 14).

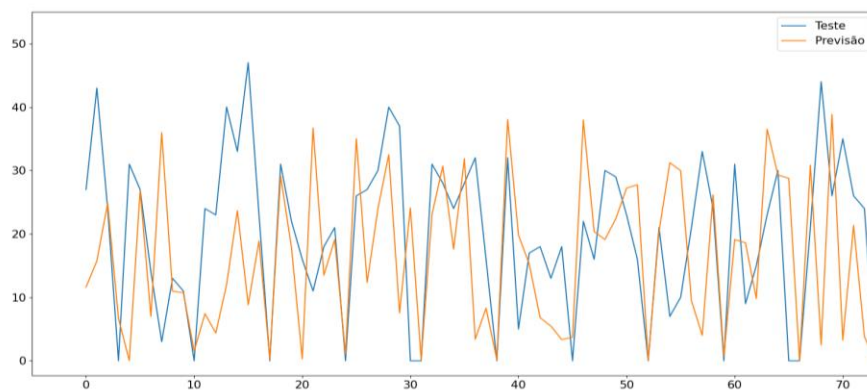
Figura 12 — Previsão de 146 dias



Fonte: Autor (2023).

Na figura 12, a linha em azul representa os dados de teste que representam os dados reais e a linha em laranja representa a previsão do modelo. Já é possível notar que o modelo consegue acompanhar os valores reais das visitas.

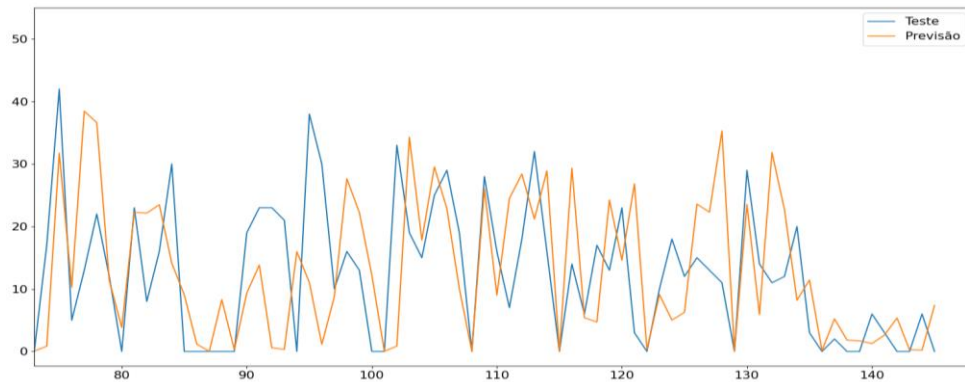
Figura 13 — Primeira parte da previsão dos 146 dias



Fonte: Autor (2023).

Nessa primeira parte, figura 13, é possível notar uma certa regularidade nos ciclos de altas de baixas nos valores dos dados.

Figura 14 — Segunda parte da previsão de 146 dias



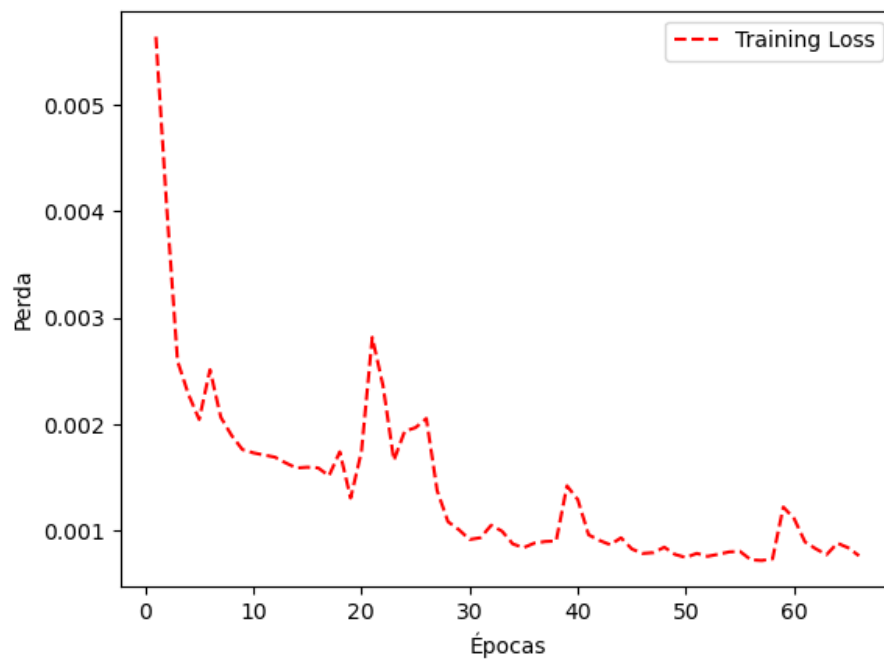
Fonte: Autor (2023).

É possível notar que o modelo conseguiu acompanhar as variações dos valores reais, mesmo quando houve uma variação abrupta entre os dias 130 e 140. Assim, a previsão consegue acompanhar a tendência (subida e descida) das visitas ao ambulatório da fábrica, fornecendo uma informação útil para que o ambulatório possa se preparar em termos de insumos e mão-de-obra.

4.3 PRECISÃO DO MODELO DE PREVISÃO

O modelo obteve um número baixo de perdas no treinamento do modelo de métricas (Figura 15) expresso pelo erro quadrático médio (EQM). Isso significa que as previsões são mais próximas dos valores reais. Desse modo, existem indícios de que o modelo está se ajustando bem aos dados de treinamento, ganhando propriedades que o tornam capaz de generalizar para novas observações, algo importante para se fazer previsões. Assim, a Figura 12 representa a perda no treinamento (*Training loss*) durante as épocas do treinamento do modelo.

Figura 15 — Perdas durante o treinamento do modelo



Fonte: Autor (2023).

Fica evidenciada a evolução da perda do modelo ao longo de 60 épocas de treinamento. A perda inicial do modelo era de pouco mais de 0.005 e, ao final do treinamento, a perda do modelo havia diminuído para abaixo de 0.001. A curva de perda é suave e consistente, sem grandes variações. Isso indica que o modelo está aprendendo bem e que o erro está diminuindo ao longo do tempo.

5 CONCLUSÃO

Com base nos resultados obtidos, podemos concluir que o modelo de redes neurais recorrentes treinado com 60 épocas foi capaz de prever com sucesso a utilização do ambulatório em uma fábrica de calçados para um período de 140 dias. O EQM obtido indica uma boa precisão na previsão da demanda futura.

Ao analisar os dados históricos de utilização do ambulatório, o modelo foi capaz de identificar padrões sazonais, além de prever aumentos ou reduções inesperados na demanda. Essa precisão na previsão é importante para a tomada de decisões estratégicas e para o planejamento dos recursos do ambulatório.

Esses resultados sugerem que as redes neurais recorrentes são uma técnica eficaz para prever a demanda futura de ambulatórios em fábricas, e podem ser aplicadas em outras empresas do setor calçadista ou em outros setores que possuam ambulatórios similares.

Portanto, concluímos que o uso de redes neurais recorrentes pode ser uma opção promissora para a previsão de demanda em ambulatórios, e pode contribuir para a melhoria da qualidade do atendimento aos funcionários e para a eficiência e produtividade da empresa como um todo.

Contudo, esse TCC apresenta algumas limitações. A primeira delas é o uso de um modelo com dados univariados (número de visitas ao ambulatório) para fazer a previsão. A segunda limitação foi o baixo horizonte de tempo, pois a pandemia impossibilitou a análise dos dados ambulatoriais devido a paralização por um tempo das atividades econômicas da Fábrica de calçados.

Assim, sugere-se que estudos futuros possam fazer uso de um modelo multivariados com redes neurais recorrentes de múltiplos previsores do tipo long-short term memory (LSTM) para fazer as previsões, dados que modelos multivariados tendem a ter um poder de previsão superior. Também recomenda-se avaliar as visitas após o período de pandemia, dado que um maior cuidado para as questões de saúde passou a ser demandado pelas empresas em geral.

REFERÊNCIAS

ARAUJO, R. de A.; SOARES, S. C. B.; MEIRA, S. R. de L. Um Modelo de Rede Neural Profunda para Previsão da Pandemia do SARS-COV-2 no Brasil. **Revista Semiárido De Visu**, 2021. Disponível em: <https://semiaridodevisu.ifsertao-pe.edu.br/index.php/rsdv/article/view/21>. Acesso em: 14 maio. 2023.

BITARAES, Santino Martins. **Previsão de séries temporais em processos de mineração utilizando redes neurais recorrentes e séries temporais nebulosa**. ResearchGate. Ouro Preto, 2019. Disponível em: https://www.researchgate.net/publication/338075558_Previsao_de_series_temporais_em_processos_de_mineracao_utilizando_redes_neurais_recorrentes_e_series_temporais_nebulosas. Acesso em: 17 abr. 2023.

Bortolot, Z. J., & Wynne, R. H. (2005). **Estimating forest biomass using small footprint LiDAR data: An individual tree-based approach that incorporates training data**. ISPRS Journal of Photogrammetry and Remote Sensing, 59(6), 342-360.

Casanova, S., de Araujo, V. M. G., da Silva, W. V., & da Rocha, D. T. (2012). Previsão da demanda turística da cidade de Foz do Iguaçu: Uma aplicação com os modelos ARIMA. *Turismo-Visão e Ação*, 14(3), 366-385.

CHOLLET, Francois. **Deep Learning with Python**. Manning Publications, 2017.

DOS SANTOS LEITE, W. K., da Silva Araújo, A. J., da Silva, J. M. N., Gontijo, L. A., de Araújo Vieira, E. M., de Souza, E. L., ... & da Silva, L. B. Risk factors for work-related musculoskeletal disorders among workers in the footwear industry: a cross-sectional study. *International Journal of Occupational Safety and Ergonomics*, 27(2), 2021.

G, Zhanj. **Forecasting with artificial neural networks**: The state of the art. *International journal of forecasting*, 1998.

HAYKIN, Simon. **Redes Neurais: Princípios e prática**. Tradução Paulo Martins Engel. 2ª ed. 2001. Tradução de: *Neural networks: a comprehensive foundation*, 2/E.

HIRANSHA, M.. **Nse stock market prediction using deep-learning models**. sciencedirect. International Conference on Computational Intelligence and Data Science. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1877050918307828>. Acesso em: 17 abr. 2023.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. **Long Short Term Memory**. 1997.

INE, Instituto. **AMBULATÓRIO DE SAÚDE OCUPACIONAL**. Disponível em: https://institutoine.com.br/arquivos/_5ee90d41402d9.pdf. Acesso em: 17 abr. 2023.

LILLICRAP, T. P.; SANTORO, A.; MARRIS, L.; AKERMAN, C. J.; HINTON, G. **Backpropagation and the brain**. *Nature Reviews Neuroscience*, v.21, n.6, p.335-346, 2020.

LUSTOSA, Leonardo. **Planejamento E Controle Da Produção**. 4ª reimpressão ed. Rio de Janeiro: Elsevier, 2008.

MAIA, Alexandre Gori. **Econometria: Conceitos e Aplicações**. Saint Paul Editora, v. 3, 2017.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Análise de séries temporais**. 2 ed. São Paulo: Blucher, 2006.

NASCIMENTO, J. D., Escovedo, T., & Kalinowski, M. (2022). **Construção de Tábuas de Mortalidade utilizando Redes Neurais LSTM Bidirecionais para Predição de Probabilidades de Morte**. ISys - Revista Brasileira de Sistemas de Informação, 2022. Disponível em: <https://sol.sbc.org.br/journals/index.php/isy/article/view/2230/1933>. Acesso em: 17 abr. 2023.

ORSSATTO, A., Rizzi, C., Rizzi, R. e Brun, A. 2022. **Aplicação de séries temporais na estimação do número de casos de dengue em Cascavel-PR**. Revista Brasileira de Computação Aplicada, 2022. Disponível em: <http://seer.upf.br/index.php/rbca/article/view/13483>. Acesso em: 17 abr. 2023.

**APÊNDICE A — Dados das visitas ao ambulatório de uma fábrica de calçados entre
janeiro de 2018 e dezembro de 2019**

Tabela 1 — Visitas ao Ambulatório

Índice	Dia	Mês	Ano	Quantidade
1	1	jan	2018	0
2	2	jan	2018	9
3	3	jan	2018	13
4	4	jan	2018	12
5	5	jan	2018	10
6	6	jan	2018	0
7	7	jan	2018	0
8	8	jan	2018	5
9	9	jan	2018	10
10	10	jan	2018	0
11	11	jan	2018	0
12	12	jan	2018	0
13	13	jan	2018	0
14	14	jan	2018	0
15	15	jan	2018	0
16	16	jan	2018	0
17	17	jan	2018	0
18	18	jan	2018	0
19	19	jan	2018	8
20	20	jan	2018	0
21	21	jan	2018	0
22	22	jan	2018	0
23	23	jan	2018	0
24	24	jan	2018	0
25	25	jan	2018	0
26	26	jan	2018	0
27	27	jan	2018	0
28	28	jan	2018	0
29	29	jan	2018	0
30	30	jan	2018	0
31	31	jan	2018	0
32	1	fev	2018	0
33	2	fev	2018	4
34	3	fev	2018	0
35	4	fev	2018	0
36	5	fev	2018	0

37	6	fev	2018	0
38	7	fev	2018	4
39	8	fev	2018	0
40	9	fev	2018	4
41	10	fev	2018	0
42	11	fev	2018	0
43	12	fev	2018	0
44	13	fev	2018	0
45	14	fev	2018	7
46	15	fev	2018	4
47	16	fev	2018	0
48	17	fev	2018	0
49	18	fev	2018	0
50	19	fev	2018	10
51	20	fev	2018	0
52	21	fev	2018	6
53	22	fev	2018	4
54	23	fev	2018	6
55	24	fev	2018	0
56	25	fev	2018	0
57	26	fev	2018	7
58	27	fev	2018	13
59	28	fev	2018	2
60	1	mar	2018	4
61	2	mar	2018	5
62	3	mar	2018	2
63	4	mar	2018	0
64	5	mar	2018	9
65	6	mar	2018	7
66	7	mar	2018	0
67	8	mar	2018	0
68	9	mar	2018	0
69	10	mar	2018	0
70	11	mar	2018	0
71	12	mar	2018	5
72	13	mar	2018	20
73	14	mar	2018	0
74	15	mar	2018	0
75	16	mar	2018	0
76	17	mar	2018	0
77	18	mar	2018	0
78	19	mar	2018	2

79	20	mar	2018	0
80	21	mar	2018	0
81	22	mar	2018	0
82	23	mar	2018	0
83	24	mar	2018	0
84	25	mar	2018	0
85	26	mar	2018	0
86	27	mar	2018	0
87	28	mar	2018	0
88	29	mar	2018	0
89	30	mar	2018	0
90	31	mar	2018	0
91	1	abr	2018	0
92	2	abr	2018	4
93	3	abr	2018	0
94	4	abr	2018	0
95	5	abr	2018	0
96	6	abr	2018	3
97	7	abr	2018	0
98	8	abr	2018	0
99	9	abr	2018	8
100	10	abr	2018	8
101	11	abr	2018	7
102	12	abr	2018	4
103	13	abr	2018	3
104	14	abr	2018	0
105	15	abr	2018	0
106	16	abr	2018	3
107	17	abr	2018	4
108	18	abr	2018	7
109	19	abr	2018	2
110	20	abr	2018	3
111	21	abr	2018	0
112	22	abr	2018	0
113	23	abr	2018	14
114	24	abr	2018	9
115	25	abr	2018	8
116	26	abr	2018	2
117	27	abr	2018	11
118	28	abr	2018	0
119	29	abr	2018	0
120	30	abr	2018	4

121	1	mai	2018	0
122	2	mai	2018	8
123	3	mai	2018	3
124	4	mai	2018	0
125	5	mai	2018	2
126	6	mai	2018	0
127	7	mai	2018	24
128	8	mai	2018	27
129	9	mai	2018	30
130	10	mai	2018	27
131	11	mai	2018	22
132	12	mai	2018	29
133	13	mai	2018	0
134	14	mai	2018	3
135	15	mai	2018	40
136	16	mai	2018	27
137	17	mai	2018	17
138	18	mai	2018	19
139	19	mai	2018	14
140	20	mai	2018	0
141	21	mai	2018	2
142	22	mai	2018	4
143	23	mai	2018	12
144	24	mai	2018	13
145	25	mai	2018	21
146	26	mai	2018	18
147	27	mai	2018	0
148	28	mai	2018	14
149	29	mai	2018	6
150	30	mai	2018	14
151	31	mai	2018	11
152	1	jun	2018	4
153	2	jun	2018	6
154	3	jun	2018	0
155	4	jun	2018	7
156	5	jun	2018	7
157	6	jun	2018	7
158	7	jun	2018	3
159	8	jun	2018	8
160	9	jun	2018	4
161	10	jun	2018	0
162	11	jun	2018	5

163	12	jun	2018	4
164	13	jun	2018	7
165	14	jun	2018	12
166	15	jun	2018	6
167	16	jun	2018	3
168	17	jun	2018	0
169	18	jun	2018	4
170	19	jun	2018	5
171	20	jun	2018	26
172	21	jun	2018	14
173	22	jun	2018	23
174	23	jun	2018	9
175	24	jun	2018	0
176	25	jun	2018	23
177	26	jun	2018	27
178	27	jun	2018	17
179	28	jun	2018	23
180	29	jun	2018	4
181	30	jun	2018	17
182	1	jul	2018	0
183	2	jul	2018	29
184	3	jul	2018	28
185	4	jul	2018	25
186	5	jul	2018	26
187	6	jul	2018	21
188	7	jul	2018	9
189	8	jul	2018	0
190	9	jul	2018	19
191	10	jul	2018	15
192	11	jul	2018	17
193	12	jul	2018	8
194	13	jul	2018	6
195	14	jul	2018	12
196	15	jul	2018	0
197	16	jul	2018	25
198	17	jul	2018	3
199	18	jul	2018	29
200	19	jul	2018	10
201	20	jul	2018	0
202	21	jul	2018	0
203	22	jul	2018	0
204	23	jul	2018	0

205	24	jul	2018	0
206	25	jul	2018	0
207	26	jul	2018	3
208	27	jul	2018	9
209	28	jul	2018	0
210	29	jul	2018	0
211	30	jul	2018	7
212	31	jul	2018	20
213	1	ago	2018	7
214	2	ago	2018	11
215	3	ago	2018	7
216	4	ago	2018	16
217	5	ago	2018	0
218	6	ago	2018	22
219	7	ago	2018	13
220	8	ago	2018	0
221	9	ago	2018	0
222	10	ago	2018	0
223	11	ago	2018	0
224	12	ago	2018	0
225	13	ago	2018	0
226	14	ago	2018	0
227	15	ago	2018	0
228	16	ago	2018	0
229	17	ago	2018	10
230	18	ago	2018	18
231	19	ago	2018	0
232	20	ago	2018	4
233	21	ago	2018	5
234	22	ago	2018	5
235	23	ago	2018	7
236	24	ago	2018	4
237	25	ago	2018	4
238	26	ago	2018	0
239	27	ago	2018	0
240	28	ago	2018	5
241	29	ago	2018	8
242	30	ago	2018	12
243	31	ago	2018	7
244	1	set	2018	2
245	2	set	2018	0
246	3	set	2018	0

247	4	set	2018	0
248	5	set	2018	0
249	6	set	2018	6
250	7	set	2018	0
251	8	set	2018	0
252	9	set	2018	0
253	10	set	2018	0
254	11	set	2018	10
255	12	set	2018	3
256	13	set	2018	6
257	14	set	2018	3
258	15	set	2018	11
259	16	set	2018	0
260	17	set	2018	12
261	18	set	2018	27
262	19	set	2018	16
263	20	set	2018	9
264	21	set	2018	4
265	22	set	2018	8
266	23	set	2018	0
267	24	set	2018	16
268	25	set	2018	9
269	26	set	2018	6
270	27	set	2018	12
271	28	set	2018	29
272	29	set	2018	18
273	30	set	2018	0
274	1	out	2018	35
275	2	out	2018	25
276	3	out	2018	11
277	4	out	2018	24
278	5	out	2018	23
279	6	out	2018	22
280	7	out	2018	0
281	8	out	2018	22
282	9	out	2018	23
283	10	out	2018	23
284	11	out	2018	28
285	12	out	2018	6
286	13	out	2018	5
287	14	out	2018	0
288	15	out	2018	30

289	16	out	2018	31
290	17	out	2018	5
291	18	out	2018	18
292	19	out	2018	25
293	20	out	2018	22
294	21	out	2018	0
295	22	out	2018	13
296	23	out	2018	30
297	24	out	2018	21
298	25	out	2018	12
299	26	out	2018	22
300	27	out	2018	19
301	28	out	2018	0
302	29	out	2018	20
303	30	out	2018	33
304	31	out	2018	26
305	1	nov	2018	25
306	2	nov	2018	13
307	3	nov	2018	8
308	4	nov	2018	0
309	5	nov	2018	18
310	6	nov	2018	32
311	7	nov	2018	15
312	8	nov	2018	19
313	9	nov	2018	12
314	10	nov	2018	12
315	11	nov	2018	0
316	12	nov	2018	23
317	13	nov	2018	11
318	14	nov	2018	29
319	15	nov	2018	2
320	16	nov	2018	17
321	17	nov	2018	10
322	18	nov	2018	0
323	19	nov	2018	38
324	20	nov	2018	18
325	21	nov	2018	38
326	22	nov	2018	20
327	23	nov	2018	18
328	24	nov	2018	13
329	25	nov	2018	0
330	26	nov	2018	22

331	27	nov	2018	25
332	28	nov	2018	23
333	29	nov	2018	16
334	30	nov	2018	31
335	1	dez	2018	18
336	2	dez	2018	0
337	3	dez	2018	23
338	4	dez	2018	12
339	5	dez	2018	20
340	6	dez	2018	15
341	7	dez	2018	9
342	8	dez	2018	19
343	9	dez	2018	0
344	10	dez	2018	37
345	11	dez	2018	43
346	12	dez	2018	48
347	13	dez	2018	27
348	14	dez	2018	21
349	15	dez	2018	18
350	16	dez	2018	0
351	17	dez	2018	13
352	18	dez	2018	25
353	19	dez	2018	23
354	20	dez	2018	12
355	21	dez	2018	6
356	22	dez	2018	5
357	23	dez	2018	0
358	24	dez	2018	2
359	25	dez	2018	0
360	26	dez	2018	14
361	27	dez	2018	13
362	28	dez	2018	27
363	29	dez	2018	14
364	30	dez	2018	0
365	31	dez	2018	0
366	1	jan	2019	0
367	2	jan	2019	0
368	3	jan	2019	30
369	4	jan	2019	26
370	5	jan	2019	23
371	6	jan	2019	0
372	7	jan	2019	21

373	8	jan	2019	11
374	9	jan	2019	5
375	10	jan	2019	7
376	11	jan	2019	14
377	12	jan	2019	17
378	13	jan	2019	0
379	14	jan	2019	17
380	15	jan	2019	18
381	16	jan	2019	14
382	17	jan	2019	14
383	18	jan	2019	16
384	19	jan	2019	10
385	20	jan	2019	0
386	21	jan	2019	23
387	22	jan	2019	18
388	23	jan	2019	24
389	24	jan	2019	17
390	25	jan	2019	23
391	26	jan	2019	14
392	27	jan	2019	0
393	28	jan	2019	29
394	29	jan	2019	27
395	30	jan	2019	12
396	31	jan	2019	21
397	1	fev	2019	19
398	2	fev	2019	8
399	3	fev	2019	0
400	4	fev	2019	10
401	5	fev	2019	17
402	6	fev	2019	20
403	7	fev	2019	8
404	8	fev	2019	27
405	9	fev	2019	11
406	10	fev	2019	0
407	11	fev	2019	24
408	12	fev	2019	14
409	13	fev	2019	20
410	14	fev	2019	25
411	15	fev	2019	0
412	16	fev	2019	4
413	17	fev	2019	0
414	18	fev	2019	22

415	19	fev	2019	18
416	20	fev	2019	16
417	21	fev	2019	16
418	22	fev	2019	21
419	23	fev	2019	19
420	24	fev	2019	0
421	25	fev	2019	17
422	26	fev	2019	8
423	27	fev	2019	21
424	28	fev	2019	12
425	1	mar	2019	12
426	2	mar	2019	10
427	3	mar	2019	0
428	4	mar	2019	42
429	5	mar	2019	0
430	6	mar	2019	28
431	7	mar	2019	15
432	8	mar	2019	25
433	9	mar	2019	7
434	10	mar	2019	0
435	11	mar	2019	21
436	12	mar	2019	22
437	13	mar	2019	15
438	14	mar	2019	12
439	15	mar	2019	0
440	16	mar	2019	0
441	17	mar	2019	0
442	18	mar	2019	27
443	19	mar	2019	23
444	20	mar	2019	21
445	21	mar	2019	12
446	22	mar	2019	20
447	23	mar	2019	10
448	24	mar	2019	0
449	25	mar	2019	23
450	26	mar	2019	19
451	27	mar	2019	7
452	28	mar	2019	20
453	29	mar	2019	19
454	30	mar	2019	18
455	31	mar	2019	0
456	1	abr	2019	13

457	2	abr	2019	13
458	3	abr	2019	20
459	4	abr	2019	0
460	5	abr	2019	21
461	6	abr	2019	23
462	7	abr	2019	0
463	8	abr	2019	29
464	9	abr	2019	17
465	10	abr	2019	11
466	11	abr	2019	18
467	12	abr	2019	24
468	13	abr	2019	15
469	14	abr	2019	0
470	15	abr	2019	35
471	16	abr	2019	30
472	17	abr	2019	33
473	18	abr	2019	36
474	19	abr	2019	0
475	20	abr	2019	0
476	21	abr	2019	0
477	22	abr	2019	27
478	23	abr	2019	21
479	24	abr	2019	28
480	25	abr	2019	23
481	26	abr	2019	18
482	27	abr	2019	13
483	28	abr	2019	0
484	29	abr	2019	30
485	30	abr	2019	26
486	1	mai	2019	0
487	2	mai	2019	33
488	3	mai	2019	29
489	4	mai	2019	18
490	5	mai	2019	0
491	6	mai	2019	32
492	7	mai	2019	20
493	8	mai	2019	29
494	9	mai	2019	5
495	10	mai	2019	10
496	11	mai	2019	17
497	12	mai	2019	0
498	13	mai	2019	25

499	14	mai	2019	22
500	15	mai	2019	28
501	16	mai	2019	32
502	17	mai	2019	29
503	18	mai	2019	15
504	19	mai	2019	0
505	20	mai	2019	27
506	21	mai	2019	27
507	22	mai	2019	14
508	23	mai	2019	38
509	24	mai	2019	27
510	25	mai	2019	2
511	26	mai	2019	0
512	27	mai	2019	12
513	28	mai	2019	16
514	29	mai	2019	22
515	30	mai	2019	13
516	31	mai	2019	26
517	1	jun	2019	26
518	2	jun	2019	0
519	3	jun	2019	35
520	4	jun	2019	21
521	5	jun	2019	22
522	6	jun	2019	18
523	7	jun	2019	19
524	8	jun	2019	9
525	9	jun	2019	0
526	10	jun	2019	21
527	11	jun	2019	27
528	12	jun	2019	16
529	13	jun	2019	18
530	14	jun	2019	24
531	15	jun	2019	19
532	16	jun	2019	0
533	17	jun	2019	10
534	18	jun	2019	16
535	19	jun	2019	25
536	20	jun	2019	24
537	21	jun	2019	12
538	22	jun	2019	15
539	23	jun	2019	0
540	24	jun	2019	0

541	25	jun	2019	38
542	26	jun	2019	38
543	27	jun	2019	22
544	28	jun	2019	26
545	29	jun	2019	0
546	30	jun	2019	0
547	1	jul	2019	15
548	2	jul	2019	19
549	3	jul	2019	23
550	4	jul	2019	11
551	5	jul	2019	29
552	6	jul	2019	15
553	7	jul	2019	0
554	8	jul	2019	28
555	9	jul	2019	18
556	10	jul	2019	24
557	11	jul	2019	41
558	12	jul	2019	37
559	13	jul	2019	17
560	14	jul	2019	0
561	15	jul	2019	27
562	16	jul	2019	22
563	17	jul	2019	0
564	18	jul	2019	30
565	19	jul	2019	23
566	20	jul	2019	9
567	21	jul	2019	0
568	22	jul	2019	29
569	23	jul	2019	24
570	24	jul	2019	12
571	25	jul	2019	30
572	26	jul	2019	29
573	27	jul	2019	21
574	28	jul	2019	0
575	29	jul	2019	36
576	30	jul	2019	26
577	31	jul	2019	25
578	1	ago	2019	38
579	2	ago	2019	18
580	3	ago	2019	17
581	4	ago	2019	0
582	5	ago	2019	0

583	6	ago	2019	35
584	7	ago	2019	24
585	8	ago	2019	27
586	9	ago	2019	43
587	10	ago	2019	24
588	11	ago	2019	0
589	12	ago	2019	31
590	13	ago	2019	27
591	14	ago	2019	14
592	15	ago	2019	3
593	16	ago	2019	13
594	17	ago	2019	11
595	18	ago	2019	0
596	19	ago	2019	24
597	20	ago	2019	23
598	21	ago	2019	40
599	22	ago	2019	33
600	23	ago	2019	47
601	24	ago	2019	23
602	25	ago	2019	0
603	26	ago	2019	31
604	27	ago	2019	22
605	28	ago	2019	16
606	29	ago	2019	11
607	30	ago	2019	18
608	31	ago	2019	21
609	1	set	2019	0
610	2	set	2019	26
611	3	set	2019	27
612	4	set	2019	30
613	5	set	2019	40
614	6	set	2019	37
615	7	set	2019	0
616	8	set	2019	0
617	9	set	2019	31
618	10	set	2019	28
619	11	set	2019	24
620	12	set	2019	28
621	13	set	2019	32
622	14	set	2019	16
623	15	set	2019	0
624	16	set	2019	32

625	17	set	2019	5
626	18	set	2019	17
627	19	set	2019	18
628	20	set	2019	13
629	21	set	2019	18
630	22	set	2019	0
631	23	set	2019	22
632	24	set	2019	16
633	25	set	2019	30
634	26	set	2019	29
635	27	set	2019	23
636	28	set	2019	16
637	29	set	2019	0
638	30	set	2019	21
639	1	out	2019	7
640	2	out	2019	10
641	3	out	2019	21
642	4	out	2019	33
643	5	out	2019	24
644	6	out	2019	0
645	7	out	2019	31
646	8	out	2019	9
647	9	out	2019	15
648	10	out	2019	23
649	11	out	2019	30
650	12	out	2019	0
651	13	out	2019	0
652	14	out	2019	21
653	15	out	2019	44
654	16	out	2019	26
655	17	out	2019	35
656	18	out	2019	26
657	19	out	2019	24
658	20	out	2019	0
659	21	out	2019	17
660	22	out	2019	42
661	23	out	2019	5
662	24	out	2019	13
663	25	out	2019	22
664	26	out	2019	12
665	27	out	2019	0
666	28	out	2019	23

667	29	out	2019	8
668	30	out	2019	16
669	31	out	2019	30
670	1	nov	2019	0
671	2	nov	2019	0
672	3	nov	2019	0
673	4	nov	2019	0
674	5	nov	2019	0
675	6	nov	2019	19
676	7	nov	2019	23
677	8	nov	2019	23
678	9	nov	2019	21
679	10	nov	2019	0
680	11	nov	2019	38
681	12	nov	2019	30
682	13	nov	2019	10
683	14	nov	2019	16
684	15	nov	2019	13
685	16	nov	2019	0
686	17	nov	2019	0
687	18	nov	2019	33
688	19	nov	2019	19
689	20	nov	2019	15
690	21	nov	2019	25
691	22	nov	2019	29
692	23	nov	2019	19
693	24	nov	2019	0
694	25	nov	2019	28
695	26	nov	2019	16
696	27	nov	2019	7
697	28	nov	2019	18
698	29	nov	2019	32
699	30	nov	2019	16
700	1	dez	2019	0
701	2	dez	2019	14
702	3	dez	2019	6
703	4	dez	2019	17
704	5	dez	2019	13
705	6	dez	2019	23
706	7	dez	2019	3
707	8	dez	2019	0
708	9	dez	2019	10

709	10	dez	2019	18
710	11	dez	2019	12
711	12	dez	2019	15
712	13	dez	2019	13
713	14	dez	2019	11
714	15	dez	2019	0
715	16	dez	2019	29
716	17	dez	2019	14
717	18	dez	2019	11
718	19	dez	2019	12
719	20	dez	2019	20
720	21	dez	2019	3
721	22	dez	2019	0
722	23	dez	2019	2
723	24	dez	2019	0
724	25	dez	2019	0
725	26	dez	2019	6
726	27	dez	2019	3
727	28	dez	2019	0
728	29	dez	2019	0
729	30	dez	2019	6
730	31	dez	2019	0

Fonte: O autor (2023).

APÊNDICE B — Algoritmo em linguagem de programação python utilizado para a previsão de redes neurais recorrentes LSTM

```

import math
import numpy as np
import seaborn as sns
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
import matplotlib.pyplot as plt
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from keras import callbacks
from keras.callbacks import EarlyStopping, ReduceLROnPlateau

def plot_predictions(model):
    predictions = model1.predict(x_test)
    predictions = scaler.inverse_transform(predictions)
    stock_data = data.filter(['Quantidade'])
    train = stock_data[:training_data_len]
    validation = stock_data[training_data_len:]
    validation['Predictions'] = predictions
    plt.figure(figsize=(16,8))
    plt.title('Model')
    plt.xlabel('Date')
    plt.ylabel('Close Price USD ($)')
    plt.plot(train)
    plt.plot(validation[['Quantidade', 'Predictions']])
    plt.show()
    rmse = np.sqrt(np.mean(predictions - y_test)**2)
    print('                rmse=',rmse)

```

```

print("\n\n")

def plot_loss(history):
    training_loss = history
    epoch_count = range(1, len(training_loss) + 1)
    plt.plot(epoch_count, training_loss, 'r--')
    plt.legend(['Training Loss', 'Test Loss'])
    plt.xlabel('Épocas')
    plt.ylabel('Perda')
    plt.show();

data= pd.read_csv('pasta1.csv', sep=";")

plt.figure(figsize=(15, 8))
plt.title('Stock Prices History')
plt.plot(data['Quantidade'])
plt.xlabel('Date')
plt.ylabel('Prices ($)')

# Preparação para treinamento

quantidade = data['Quantidade']

values = quantidade.values

training_data_len = math.ceil(len(values)* 0.8)

scaler = MinMaxScaler(feature_range=(0,1))

scaled_data = scaler.fit_transform(values.reshape(-1,1))

train_data = scaled_data[0: training_data_len, :]

```

```
# Criando as variáveis de previsão
```

```
x_train = []
```

```
y_train = []
```

```
for i in range(7, len(train_data)):
```

```
    x_train.append(train_data[i-7:i, 0])
```

```
    y_train.append(train_data[i, 0])
```

```
x_train, y_train = np.array(x_train), np.array(y_train)
```

```
x_train.shape
```

```
x_train = np.reshape(x_train , newshape = (x_train.shape[0],
                                             x_train.shape[1], 1))
```

```
x_train.shape
```

```
test_data = scaled_data[training_data_len-7: , : ]
```

```
x_test = []
```

```
y_test = values[training_data_len:]
```

```
for i in range(7, len(test_data)):
```

```
    x_test.append(test_data[i-7:i, 0])
```

```
x_test = np.array(x_test)
```

```
x_test = np.reshape(x_test, (x_test.shape[0], x_test.shape[1], 1))
```

```
### Criando o modelo de previsão

model1= keras.Sequential()

## Adiciona a primeira camada recorrente
model1.add(layers.LSTM(units = 100,
                        return_sequences = True,
                        input_shape = (x_train.shape[1], 1)))

## Adiciona a segunda camada recorrente
model1.add(layers.LSTM(units = 50, return_sequences = True))

## Adiciona a terceira camada recorrente

model1.add(layers.LSTM(units = 50, return_sequences = False))

## Adiciona a quarta camada

model1.add(layers.Dense(units = 50))

## Adiciona a camada de saída

model1.add(layers.Dense(units = 1, activation = 'sigmoid'))

## Compila o modelo
model1.compile(optimizer = 'adam',
               loss = 'mean_squared_error',
               metrics =['accuracy'])

## métricas de paradas
```

```
es = EarlyStopping(monitor = 'loss',  
                    min_delta = 0.0001,  
                    patience = 20,  
                    mode = 'auto',  
                    verbose = 1)
```

```
rlr = ReduceLROnPlateau(monitor = 'loss',  
                         factor = 0.5,  
                         patience = 20,  
                         mode = 'auto',  
                         verbose = 1)
```

```
## Treinamento da rede neural
```

```
history = model1.fit(x_train,  
                     y_train,  
                     batch_size = 32,  
                     epochs = 200,  
                     callbacks =[es, rlr])
```

```
## Plota a perda do treinamento  
plot_loss( history.history['loss'])
```

```
## Plota a acurácia do treinamento
```

```
plot_loss( history.history['accuracy'])
```

```
## Plota a previsão em todo o período
```

```
plot_predictions(model1)
```

```
## Plota somente a previsão e teste
```

```
predictions = model1.predict(x_test)
predictions = scaler.inverse_transform(predictions)
```

```
plt.plot(quantidade.values[584:730])
plt.plot(predictions)
plt.legend(['Teste', 'Previsão'])
plt.rcParams.update({'font.size': 15})
plt.show()
```

```
plt.plot(quantidade.values[584:657])
plt.plot(predictions[0:73])
plt.legend(['Teste', 'Previsão'])
plt.rcParams.update({'font.size': 15})
plt.show()
```

```
plt.plot(quantidade.values[657:730])
plt.plot(predictions[73:146])
plt.legend(['Teste', 'Previsão'])
plt.rcParams.update({'font.size': 15})
plt.xlim(70,150)
plt.show()
```