



# Utilizando Redes Neurais Recorrentes Para Previsão do Processo Chuva Vazão

TRABALHO DE CONCLUSÃO DE CURSO

ENGENHARIA DA COMPUTAÇÃO

ESTYVISON LINHARES DE ALBUQUERQUE

ORIENTADOR: PROF. DR. MÊUSER JORGE SILVA VALENÇA



UNIVERSIDADE  
DE PERNAMBUCO



Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Graduação em Engenharia de Computação

Estyvison Linhares de Albuquerque

# **Utilizando Redes Neurais Recorrentes Para Previsão do Processo Chuva Vazão**

Monografia apresentada como requisito parcial  
para obtenção do diploma de Bacharel em Enge-  
nharia de Computação pela Escola Politécnica  
de Pernambuco – Universidade de Pernambuco.

Recife, Outubro de 2022

Albuquerque, Estyvison Linhares de  
Utilizando Redes Neurais Recorrentes Para Previsão do  
Processo Chuva Vazão / Estyvison Linhares de Albuquerque. – Recife  
- PE, 2022.

xv, 45 f. : il. ; 29 cm.

Trabalho de Conclusão de Curso (Graduação em Engenharia de  
Computação) Universidade de Pernambuco, Escola Politécnica de  
Pernambuco, Recife, 2022.

Orientador: Prof<sup>ª</sup>. Dr. Mêuser Jorge Silva

Valença.

1. Redes Neurais Recorrentes. 2. Previsão de Vazão. 3. Usinas  
Hidrelétricas. I. Utilizando Redes Neurais Recorrentes Para Previsão  
do Processo Chuva Vazão. II. Valença, Mêuser Jorge Silva. III.  
Universidade de Pernambuco.

## MONOGRAFIA DE FINAL DE CURSO

### Avaliação Final (para o presidente da banca)\*

No dia 26/10/2022, às 09h00min, reuniu-se para deliberar sobre a defesa da monografia de conclusão de curso do(a) discente **ESTYVISON LINHARES DE ALBUQUERQUE**, orientado(a) pelo(a) professor(a) **MÊUSER JORGE SILVA VALENÇA**, sob título Utilizando Redes Neurais Recorrentes Para Previsão do Processo Chuva Vazão, a banca composta pelos professores:

**SÉRGIO MÁRIO LINS GALDINO (PRESIDENTE)**

**MÊUSER JORGE SILVA VALENÇA (ORIENTADOR)**

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada       Aprovada com Restrições\*       Reprovada

e foi-lhe atribuída nota: 9,5 (nove e cinco)

\*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O(A) discente terá 7 dias para entrega da versão final da monografia a contar da data deste documento.

*Sérgio Galvão*

\_\_\_\_\_  
AVALIADOR 1: Prof (a) **SÉRGIO MÁRIO LINS GALDINO**

*Mêuser Jorge Silva Valença*

\_\_\_\_\_  
AVALIADOR 2: Prof (a) **MÊUSER JORGE SILVA VALENÇA**

\_\_\_\_\_  
AVALIADOR 3: Prof (a)

\* Este documento deverá ser encadernado juntamente com a monografia em versão final.

*Este trabalho é dedicado à toda minha família e amigos  
que me ajudaram e apoiaram durante toda minha jornada neste curso.*

# Agradecimentos

Agradeço aos meus pais, Gedenilson Pereira e Tacimeire Campelo, pelo apoio, carinho e amor durante todo o percurso e caminhada até aqui.

Agradeço aos meus tios, Gerlane Pereira e Paulo Henrique, pelo apoio na elaboração deste trabalho, carinho e amor.

Agradeço aos meus avós, Genilda Cecília e Elvison Wanderley, pelo apoio, carinho e amor durante todo o percurso e caminhada até aqui.

Agradeço aos meus amigos, em especial Diego Alves, Kelly Mota, Guilherme Teixeira, Paulo Guilherme, Daniel Almeida e Arthur Moura, obrigado pelo companheirismo durante toda a jornada acadêmica.

Agradeço ao meu professor orientador, Mêuser Valença, por todos os ensinamentos, pela ajuda e por todas as oportunidades que me concedeu ao decorrer desta graduação.

A todos que direta ou indiretamente fizeram parte de minha formação, o meu muito obrigado.

*“Com o mundo tão empenhado em se despedaçar,  
não me parece uma ideia tão ruim que eu queira  
reconstruí-lo nem que seja um pouco...”*

***Soldado Desmond Doss***

## Autorização de publicação de PFC

Eu, **Estyvison Linhares de Albuquerque** autor(a) do projeto de final de curso intitulado: **Utilizando Redes Neurais Recorrentes Para Previsão do Processo Chuva Vazão**; autorizo a publicação de seu conteúdo na internet nos portais da Escola Politécnica de Pernambuco e Universidade de Pernambuco.

O conteúdo do projeto de final de curso é de responsabilidade do autor.

*Estyvison Linhares de Albuquerque*

**Estyvison Linhares de Albuquerque**

*Meuser Jorge Silva Valença*

Orientador(a): **Meuser Jorge Silva Valença**

\_\_\_\_\_  
Coorientador(a):



\_\_\_\_\_  
Prof. de TCC: **Daniel Augusto Ribeiro Chaves**

\_\_\_\_\_  
Data: 26/10/2022

# Resumo

A previsão de vazões afluentes a reservatórios hidrelétricos é fundamental para garantir a sua eficácia operacional com objetivos econômicos e de segurança. Muitos modelos são utilizados nessas previsões, entre eles há os modelos físicos, estatísticos e aqueles baseados na técnica de Redes Neurais Artificiais. A Usina Hidrelétrica Serra do Facão já foi objeto de estudo de outras pesquisas utilizando esses modelos de previsão, que teve por destaque, o modelo de Rede Neural MLP com Random Forest. O objetivo deste trabalho foi calibrar um modelo de previsão de vazões naturais, utilizando Redes Neurais Recorrentes, mais especificamente as redes LSTM, BiLSTM e GRU, para no final escolher dos três o que possuir melhor desempenho na previsão de vazão. Essa escolha se dará por meio do teste t-Student e será comparando com os resultados de outros modelos de previsão utilizados na barragem Serra do Facão. Os resultados obtidos nessa pesquisa mostram que as Redes Neurais Recorrentes foram um pouco inferiores à metodologia MLP com Random Forest, porém foram superiores aos demais modelos utilizados. Os valores do Erro Percentual Médio Absoluto foram aproximadamente 2,4% inferiores aos divulgados pelo ONS e 4,6% superiores ao modelo MLP com Random Forest.

**Palavras-chave:** redes neurais recorrentes, usinas hidrelétricas, previsão de vazões, modelos de previsão.

# Abstract

Forecasting flows into hydroelectric reservoirs is essential to ensure their operational effectiveness with economic and safety objectives. Many models are used in these predictions, among them there are physical, statistical models and those based on the technique of Artificial Neural Networks. The Serra do Facão Hydroelectric Power Plant has already been the object of study of other researches using these forecast models, which had the highlight, the MLP Neural Network model with Random Forest. The objective of this work was to calibrate a model for forecasting natural flows, using Recurrent Neural Networks, more specifically the LSTM, BiLSTM and GRU networks, in order to choose the one that has the best performance in the flow prediction of the three. This choice will be made through the t-Student test and will be compared with the results of other forecast models used in the Serra do Facão dam. The results obtained in this research show that the Recurrent Neural Networks were a little inferior to the MLP methodology with Random Forest, but they were superior to the other models used. The Mean Absolute Percentage Error values were approximately 2.4% lower than those published by the ONS and 4.6% higher than the MLP model with Random Forest.

**Keywords:** recurrent neural networks, hydroelectric plants, flow forecasting, forecast models.

# Lista de Figuras

Figura 1 – Usina Serra do Facão . . . . .	17
Figura 2 – Localização do reservatório da UHE Serra do Facão, GO . . . . .	18
Figura 3 – Na bacia hidrográfica do rio São Marcos: (a) Mapa de altitude; (b) Mapa de declividade . . . . .	19
Figura 4 – Modelo de um neurônio biológico. . . . .	22
Figura 5 – Modelo de um neurônio artificial . . . . .	23
Figura 6 – Estrutura de uma MLP com apenas uma camada oculta . . . . .	25
Figura 7 – Rede Neural Recorrente (a) e diagrama computacional, onde o quadrado azul indica um atraso de passo de tempo (b). . . . .	26
Figura 8 – Estrutura simples de uma célula da camada oculta de uma RNR (b). . . . .	26
Figura 9 – Estrutura de um bloco LSTM. . . . .	27
Figura 10 – Estrutura de um bloco GRU. . . . .	29
Figura 11 – Gráfico do MAPE do modelo LSTM . . . . .	36
Figura 12 – Ocorrido x previsto 1 do modelo LSTM . . . . .	37
Figura 13 – Ocorrido x previsto 2 do modelo LSTM . . . . .	37
Figura 14 – Ocorrido x previsto 3 do modelo LSTM . . . . .	38
Figura 15 – Gráfico do MAPE do modelo BiLSTM . . . . .	38
Figura 16 – Ocorrido x previsto 1 do modelo BiLSTM . . . . .	39
Figura 17 – Ocorrido x previsto 2 do modelo BiLSTM . . . . .	39
Figura 18 – Ocorrido x previsto 3 do modelo BiLSTM . . . . .	40
Figura 19 – Gráfico do MAPE do modelo GRU . . . . .	40
Figura 20 – Ocorrido x previsto 1 do modelo GRU . . . . .	41
Figura 21 – Ocorrido x previsto 2 do modelo GRU . . . . .	41
Figura 22 – Ocorrido x previsto 3 do modelo GRU . . . . .	42

## Lista de Tabelas

Tabela 1 – Relação de entradas e saídas calculadas nos modelos de previsão . . . . .	32
Tabela 2 – Resultados de menor erro MAPE, semente e média do erro MAPE para 30 rodadas de treino das RNR . . . . .	34
Tabela 3 – Valores do MAPE para cada modelo computacional. . . . .	36

# Lista de Abreviaturas e Siglas

ANEEL	Agência Nacional de Energia Elétrica
BP	<i>Back Propagation</i>
BPTT	<i>Back Propagation Through Time</i>
BiLSTM	<i>Deep Bidirectional Long Short-Term Memory</i>
GRU	<i>Gated Recurrent Unit</i>
INMET	Instituto Nacional de Meteorologia
IA	Inteligência Artificial
LSTM	<i>Long Short-Term Memory</i>
MLP	<i>Multi-Layer Perceptron</i>
ONS	Operador Nacional do Sistema Elétrico
RF	<i>Random Forest</i>
RNA	Redes Neurais Artificiais
RNN	<i>Recurrent Neural Network</i>
RNR	Redes Neurais Recorrentes
SIN	Sistema Interligado Nacional
UHE	Usina Hidrelétrica

# Sumário

1	INTRODUÇÃO . . . . .	14
1.1	Motivação e Contextualização . . . . .	14
1.2	Objetivos . . . . .	15
1.2.1	Objetivo Geral . . . . .	15
1.2.2	Objetivos Específicos . . . . .	15
1.3	Estrutura do Documento . . . . .	16
2	REFERENCIAL TEÓRICO . . . . .	17
2.1	Caracterização da Área de Estudo . . . . .	17
2.1.1	Previsão de Vazões . . . . .	19
2.2	<i>Machine Learning</i> . . . . .	20
2.3	Redes Neurais Artificiais . . . . .	21
2.3.1	Características Principais . . . . .	21
2.3.2	Nerônio biológico . . . . .	22
2.3.3	Neurônio Artificial . . . . .	22
2.3.4	Como uma rede neural aprende? . . . . .	24
2.3.5	Arquiteturas de RNA . . . . .	24
2.3.5.1	Redes Neurais MLP . . . . .	24
2.3.5.2	Rede Neural Recorrente (RNR) . . . . .	25
2.3.5.2.1	<i>Long Short-Term Memory (LSTM)</i> . . . . .	26
2.3.5.2.2	<i>Deep Bidirectional LSTM (BiLSTM)</i> . . . . .	28
2.3.5.2.3	<i>Gated Recurrent Unit (GRU)</i> . . . . .	28
2.4	<i>Teste t-Student</i> . . . . .	30
3	MATERIAIS E MÉTODOS . . . . .	31
3.1	Ferramentas utilizadas . . . . .	31
3.2	Base de dados . . . . .	31
3.3	Aplicação das arquiteturas LSTM, BiLSTM e GRU . . . . .	32
3.4	Aplicação do teste t-Student . . . . .	33
4	RESULTADOS E DISCUSSÃO . . . . .	34
4.1	Resultados de média MAPE . . . . .	34
4.2	Resultados da análise do teste t-Student . . . . .	34
4.2.1	Entre LSTM e BiLSTM . . . . .	34
4.2.2	Entre BiLSTM e GRU . . . . .	35
4.2.3	Entre LSTM e GRU . . . . .	35

<b>4.3</b>	<b>Análise gráfica das previsões . . . . .</b>	<b>36</b>
<b>4.3.1</b>	<b>LSTM: gráficos de MAPE e ocorrido vs previsto . . . . .</b>	<b>36</b>
<b>4.3.2</b>	<b>BiLSTM: gráficos de MAPE e ocorrido vs previsto . . . . .</b>	<b>38</b>
<b>4.3.3</b>	<b>GRU: gráficos de MAPE e ocorrido vs previsto . . . . .</b>	<b>40</b>
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS . . . . .</b>	<b>43</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>44</b>

# Capítulo 1

## Introdução

Neste capítulo serão apresentadas a motivação e contextualização deste trabalho, além de seus objetivos, estrutura e organização geral.

### 1.1 Motivação e Contextualização

Segundo a Agência Nacional de Energia Elétrica (ANEEL), o Brasil tem em operação 219 usinas hidrelétricas de grande porte (UHE), além de 425 pequenas centrais hidrelétricas e 739 centrais geradoras hidrelétricas (ENGIE, 2022). Com isso, a previsão de vazões afluentes a reservatórios hidrelétricos é fundamental para garantir a sua eficácia operacional com objetivos econômicos e de segurança (MINE; TUCCI, 1999). Para que isso seja possível, é necessária a elaboração de um plano razoável e eficaz de funcionamento do reservatório para a implementação de suas funções, tais como o controle de inundações, abastecimento de água, produção de energia, transporte e fins ecológicos.

A operação do reservatório em tempo real apresenta requisitos mais elevados para a solução da função operacional, o que requer não só a alta precisão dos resultados da solução, mas também a alta velocidade do processo de resolução. Embora existam diagramas de funcionamento e função operacional que sirvam de base e referência, ainda requerem que o pessoal de gestão tenha fortes qualidades profissionais e uma rica experiência na própria operação do reservatório, de maneira que cumpram os objetivos e reduzam a incerteza permitida por várias condições extremas de afluência. O estabelecimento de um modelo eficaz de previsão de escoamento é de grande importância para reduzir a pressão de trabalho dos operadores e para a realização do trabalho científico e gestão eficiente (ZHANG et al., 2019).

Diversos modelos são utilizados na determinação dessas previsões, entre os quais podem ser citados modelos físicos, estatísticos e aqueles baseados na técnica de Redes Neurais Artificiais (RNA) (FERREIRA et al., 2012). Esta última se inspira no comportamento do cérebro humano, possuindo uma ou mais camadas de tratamento de dados simbolizados por neurônios (VALENÇA, 2005b) e possui muitos trabalhos que mostram a evolução de sua aplicação em diversas áreas e,

particularmente, na previsão de vazões naturais para determinados intervalos de tempo (PAIVA; BARROS; VALENÇA, 2009). Existe o uso da topologia *Multi-Layer Perceptron* (MLP) com *Random Forest* (RF) para prever vazões, a qual traz algumas vantagens na sua utilização, pois elimina os parâmetros irrelevantes do aprendizado supervisionado, técnica utilizada no treinamento das RNAs, filtrando os melhores parâmetros de entrada da rede e evitando o fenômeno de *overfitting* – um ajustamento excessivo dos modelos ao conjunto de treinamento, que compromete a acurácia na classificação de novos casos (MOTA, 2021).

O algoritmo de aprendizagem da máquina em epígrafe possa completar a simulação da operação do reservatório, existem ainda muitos problemas no processo de aplicação prática. A maioria dos algoritmos de Inteligência Artificial (IA) tradicionais têm falhas no treino ou na atração do mínimo local. Eles precisam ser otimizados para obterem a solução ótima global (YANG et al., 2017). Ao mesmo tempo, uma vez que tais algoritmos são, na sua maioria, arquiteturas pouco profundas, os modelos não podem obter eficazmente as características profundas dos dados.

Nos últimos anos, o rápido desenvolvimento de algoritmos de aprendizagem profunda melhorou muito o desempenho do algoritmo de aprendizagem de máquina. Entre eles, destacam-se as Redes Neurais Recorrentes (RNR). Elas, podem preservar, lembrar, e lidar com sinais complexos a longo prazo. As RNRs mapeiam a sequência de entrada para a sequência de saída, prevendo a saída do próximo passo. Elas também possuem forte capacidade de resolver problemas complexos de *timing*. Em resumo, a operação do reservatório é um problema de temporização complexo afetado por muitos fatores, mas a aplicação da RNR na operação do reservatório ainda precisa ser explorada (ZHANG et al., 2019).

Tendo por base tudo o que foi exposto até o momento, este trabalho visa substituir as topologias MLP com Random Forest, por três arquiteturas de RNR: *Long Short-Term Memory* (LSTM), *Deep Bidirectional LSTM* (BiLSTM) e *Gated Recurrent Unit* (GRU).

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

O principal objetivo deste trabalho é realizar uma análise de previsão de vazão utilizando arquiteturas de RNR com dados da barragem nacional Serra do Facão e comparar os resultados da previsão com outros modelos preditivos já presentes na literatura utilizados para esta barragem. O intuito é o de compreender a eficácia desse modelo de rede neural no uso preditivo de vazões em usinas hidrelétricas e barragens.

### 1.2.2 Objetivos Específicos

Com relação às metas específicas, destacam-se:

- Preparar a base de dados com informações sobre a vazão e média de precipitação da barragem nacional Serra do Facão;
- Utilizar as arquiteturas BiLSTM, LSTM e GRU para realização de treino e teste de previsão da vazão;
- Realizar 30 execuções de treinamento para cada uma das três arquiteturas RNR para obter a média do *Mean Absolute Percentage Error* (MAPE);
- Realizar o teste t-Student para selecionar a melhor RNR entre a BiLSTM, LSTM e GRU;
- Comparar o resultado do modelo selecionado com o modelo que utilizou MLP com Random Forest, dentre outros modelos preditivos existentes para essa mesma barragem na literatura;
- Expor as conclusões referentes à pesquisa realizada e possíveis trabalhos futuros que poderão ser realizados.

### 1.3 Estrutura do Documento

Esta monografia está dividida em cinco capítulos, integrando este que conta com uma introdução a respeito do tema e objetivos do projeto. Em seguida a organização do trabalho é descrita como:

- **Capítulo 2 - Referencial Teórico:** Serão apresentados conceitos essenciais para o entendimento deste trabalho, apresentando uma fundamentação teórica sobre a UHE Serra do Facão, *Deep Learning*, Redes Neurais Artificiais e teste t-Student;
- **Capítulo 3 - Materiais e Métodos:** Neste capítulo será apresentado como foi conduzida a pesquisa, as ferramentas utilizadas, a base de dados e os procedimentos de treinamento das RNRs escolhidas para esta pesquisa;
- **Capítulo 4 - Resultados e Discussão:** Neste capítulo serão expostos os resultados obtidos do treinamento para cada arquitetura de RNR, a escolha da melhor entre elas pelo teste t-Student e a sua comparação com outros modelos de previsão de vazão para a UHE Serra do Facão presentes na literatura;
- **Capítulo 5 - Conclusão e Trabalhos Futuros:** Descreve as conclusões finais do trabalho, além de relatar possíveis trabalhos futuros inspirados com esta pesquisa.

## Capítulo 2

### Referencial Teórico

Este capítulo tem como objetivo fornecer a fundamentação teórica, apresentando conceitos e aspectos fundamentais para o entendimento deste trabalho.

#### 2.1 Caracterização da Área de Estudo

A Usina Hidrelétrica de Serra do Facão – "UHE Serra do Facão", localiza-se no Rio São Marcos, entre os municípios de Davinópolis e Catalão (GO), a Usina Hidrelétrica Serra do Facão tem uma potência instalada de 212,58 MW, suficiente para suprir a demanda de energia de uma cidade com 1,2 milhão de habitantes. As obras foram iniciadas em fevereiro de 2007 e a usina inaugurada em outubro de 2010, dois meses antes do previsto. Durante a implantação do empreendimento foram gerados cerca de 1,6 mil empregos diretos e outros 3,2 mil indiretos (FURNAS, 2018).

Figura 1 – Usina Serra do Facão



Fonte: Furnas (2018)

O reservatório de Serra do Facão tem uma área de 218,84 km<sup>2</sup> e cota máxima de inundação de 756 km<sup>2</sup>. Abrange parcialmente áreas de cinco municípios goianos (Catalão, com 72,8%; Campo Alegre de Goiás, com 22,4%; Cristalina, com 0,6%; Davinópolis, com 0,4% e Ipameri, com 0,1%); e um município de Minas Gerais (Paracatu com 3,7%). (FURNAS, 2018)

Figura 2 – Localização do reservatório da UHE Serra do Facão, GO



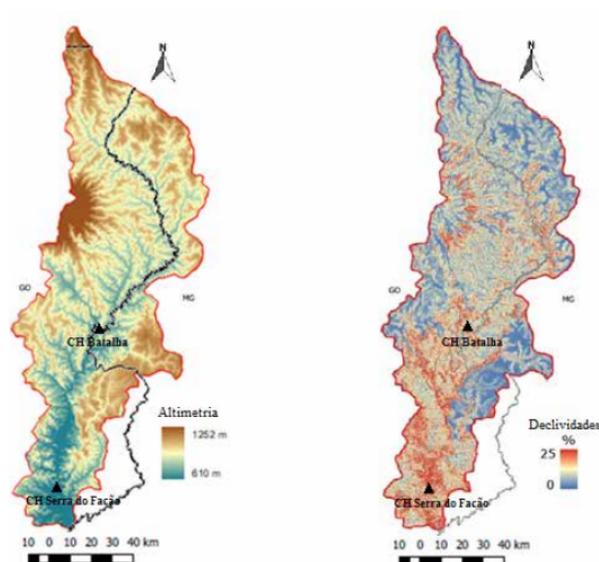
Fonte: SEFAC (2018)

Na região predomina o clima tropical úmido, com os períodos seco e úmido bem definidos. Conforme o Instituto Nacional de Meteorologia (INMET), a precipitação anual na bacia varia entre 1330 e 1550 mm, sendo a média de aproximadamente 1426 mm. Já a evapotranspiração média anual é de cerca de 900 mm. As temperaturas médias anuais predominantes ficam entre 18 e 22°C, com máxima de 37°C entre janeiro e março e mínima de 0°C entre maio e julho.

Conforme exibida na Figura 3a, as altitudes na bacia hidrográfica variam de pouco mais de 640 m, na foz do rio São Marcos, a mais de 1200 m no domo de Cristalina, localizado no município de Cristalina-GO. O domo de Cristalina possui diâmetro de 30 a 40 km, tendo sido originado a partir do arqueamento de corrente tectônica que ocasionou dobramentos em toda área circunvizinha. Já na Figura 3b é possível notar quão heterogêneo é o terreno, com declividade máxima de 25%. Nos trechos superiores, o relevo é plano a ondulado; já na parte inferior, o relevo é acidentado e montanhoso. As áreas com chapadões e vales abertos e suaves favorecem a prática da agricultura irrigada.

Inserida no bioma cerrado, a bacia hidrográfica apresenta vegetação variando de campos limpos até cerradões, relativamente mais densos. Cabe destacar que o elevado índice de cobertura vegetal na bacia deve-se especialmente à presença de amplos terrenos com maior declividade, estando também mais suscetíveis a processos de erosão e assoreamento. Os solos que predominam na região são de baixa fertilidade natural, sendo então necessária a aplicação

Figura 3 – Na bacia hidrográfica do rio São Marcos: (a) Mapa de altitude; (b) Mapa de declividade



Fonte: Basin (2020)

de fertilizantes e corretivos de solo. Outra importante característica da bacia hidrográfica é a presença de solos profundos e bem drenados, principalmente nos chapadões. Esses solos possuem alta capacidade de infiltração e de armazenamento de água, favorecendo a alimentação de córregos, rios e ribeirões durante longos períodos de estiagem (BASIN, 2020).

### 2.1.1 Previsão de Vazões

As previsões de vazões naturais, para todos os locais de aproveitamentos hidroelétricos do Sistema Interligado Nacional (SIN), são obtidas, com a participação dos agentes de geração responsáveis por essas instalações, para dois horizontes temporais: de curtíssimo prazo, até 12 dias, com discretização em valores médios diários, para fins de programação diária; e de curto prazo, até seis semanas, por valores médios semanais, para o planejamento de curto prazo da operação eletro energética (GUILHON; ROCHA; MOREIRA, 2007).

Existem algumas metodologias de modelos de previsão de vazão que estão consolidadas em pesquisas preliminares referentes à UHE Serra do Facão, sendo elas:

- **Modelo PREVIVAZ:** desenvolvido pelo Centro de Pesquisas de Energia Elétrica (CEPEL);
- **Modelo SMAP:** modelo conceitual de simulação hidrológica, do tipo transformação chuva-vazão. Foi desenvolvido em 1981 por Lopes, J.E.G., Braga, B.P.F e Conejo, J.G.L., apre-

sentado no International Symposium on Rainfall-Runoff Modeling realizado no Mississipi, Estados Unidos, e publicado pela Water Resources Publications. (ONS, NT 136/2012);

- **Modelo MLP com RF:** Modelo que utiliza técnicas de Inteligência Artificial baseadas em RNA e Árvores de Decisão (MOTA, 2021).

## 2.2 *Machine Learning*

Segundo Zaccone, Karim e Menshawy (2017), *machine learning* é uma área de pesquisa da ciência da computação que trata de métodos para identificar e implementar sistemas e algoritmos pelos quais um computador pode aprender, com base nos exemplos fornecidos na entrada. O desafio do aprendizado de máquina é permitir que um computador aprenda a reconhecer automaticamente padrões complexos e tomar decisões tão inteligentes quanto possível. Todo o processo de aprendizagem requer um conjunto de dados da seguinte forma:

- Conjunto de treinamento: esta é a base de conhecimento usada para treinar o algoritmo de aprendizado de máquina. Durante esta fase, os parâmetros do modelo de aprendizado de máquina podem ser alternados conforme o desempenho obtido;
- Conjunto de teste: é utilizado apenas para avaliar o desempenho do modelo em dados não vistos.

A teoria da aprendizagem usa ferramentas matemáticas derivadas das teorias da probabilidade e informação. Isso permite a otimização de alguns métodos em relação a outros.

Existem basicamente três paradigmas de aprendizagem que serão brevemente discutidos:

- **Aprendizado supervisionado:** é a tarefa de aprendizagem automática mais simples e mais conhecida. Baseia-se num conjunto de exemplos pré-classificados, nos quais, nomeadamente, se conhece inicialmente a categoria a que deve pertencer cada uma das entradas utilizadas como exemplos. Neste caso, a questão crucial é o problema da generalização. Após a análise de uma amostra (geralmente pequena) de exemplos, o sistema deve produzir um modelo que funcione bem para todas as entradas possíveis;
- **Aprendizado não supervisionado:** um conjunto de entradas é fornecido ao sistema durante a fase de treinamento que, no entanto, ao contrário do caso do aprendizado supervisionado, não é rotulado com a categoria de pertencimento relacionada. Esse tipo de aprendizado é importante porque no cérebro humano é mais comum de ocorrer do que o supervisionado;
- **Aprendizado com reforço:** é uma abordagem de inteligência artificial que enfatiza o aprendizado do sistema por meio de suas interações com o ambiente. Com o aprendizado

por reforço, o sistema adapta seus parâmetros com base no feedback recebido do ambiente, que dá então feedback sobre as decisões tomadas.

## 2.3 Redes Neurais Artificiais

As redes neurais artificiais são modelos computacionais inspirados no sistema nervoso dos seres vivos. Elas conseguem adquirir e manter conhecimento (baseado em informações) e podem ser definidas como um conjunto de unidades de processamento, representadas por neurônios artificiais, interligados por muitas interconexões (sinapses artificiais), implementado por vetores e matrizes de pesos sinápticos. O primeiro modelo de neurônio artificial foi proposto em 1943 por McCulloch e Pitts em um modelo computacional de atividade nervosa. Este modelo foi seguido por outro, proposto por John von Neumann, Marvin Minsky, Frank Rosenblatt e muitos outros (ZACCONE; KARIM; MENSRAWY, 2017).

### 2.3.1 Características Principais

Segundo (SILVA et al., 2017), as características mais relevantes sobre aplicações neurais artificiais são as seguintes:

- **Adaptação da experiência:** Os parâmetros internos da rede, geralmente seus pesos sinápticos, são ajustados com o exame de exemplos sucessivos (padrões, amostras ou medições) relacionadas ao comportamento do processo, permitindo assim a aquisição de conhecimento pela experiência;
- **Capacidade de aprendizagem:** Através do uso de um método de aprendizagem, a rede pode extrair relação entre as diversas variáveis da aplicação;
- **Capacidade de generalização:** Uma vez concluído o processo de aprendizagem, a rede pode generalizar o conhecimento adquirido, possibilitando a estimativa de soluções até então desconhecidas.
- **Organização de dados:** Com base em informações inatas de um determinado processo, a rede pode organizar essas informações, permitindo assim o agrupamento de padrões com características;
- **Tolerância ao erro:** Graças ao elevado número de interconexões entre neurônios artificiais, a rede neural torna-se um sistema tolerante à falha se parte de sua estrutura interna está corrompido em algum grau;
- **Armazenamento distribuído:** O conhecimento sobre o comportamento de um processo particular aprendido por uma rede é armazenada em cada uma das várias sinapses entre

os neurônios, melhorando assim a robustez da arquitetura no caso da perda de alguns neurônios;

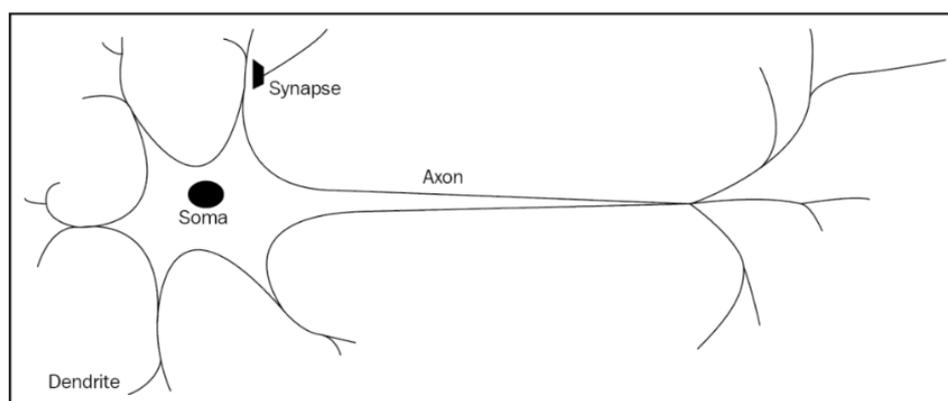
- **Prototipagem facilitada:** Dependendo das particularidades da aplicação, a maioria das arquiteturas neurais podem ser facilmente prototipadas em hardware ou software, pois seus resultados, após o treinamento do processo, são geralmente obtidos com algumas operações matemáticas fundamentais.

### 2.3.2 Neurônio biológico

Na Figura 4, um neurônio biológico é composto por:

- Um corpo celular ou soma;
- Um ou mais dendritos, cuja responsabilidade é receber sinais de outros neurônios;
- Um axônio, que transmite os sinais gerados pelo mesmo neurônio para os outros neurônios conectados.

Figura 4 – Modelo de um neurônio biológico.



Fonte: Zaccone, Karim e Menshawy (2017)

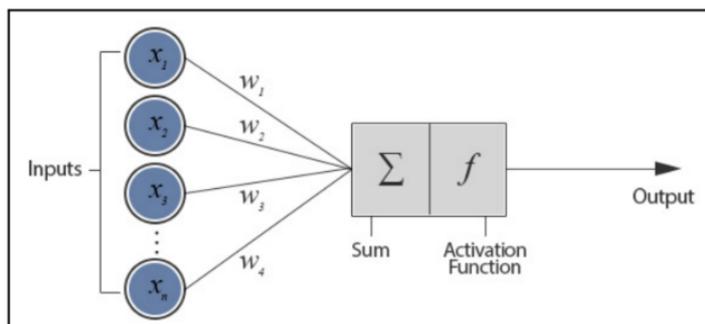
### 2.3.3 Neurônio Artificial

Na figura 5, semelhante ao biológico, o neurônio artificial consiste no seguinte:

- Uma ou mais conexões de entrada, com a tarefa de coletar sinais numéricos de outros neurônios; cada conexão recebe um peso que será usado para considerar cada sinal enviado;
- Uma ou mais conexões de saída que transportam o sinal para os demais neurônios;

- Uma função de ativação que determina o valor numérico do sinal de saída, com base nos sinais recebidos das conexões de entrada com outros neurônios, sendo adequadamente coletados dos pesos associados a cada sinal captado e do limiar de ativação do próprio neurônio.

Figura 5 – Modelo de um neurônio artificial



Fonte: Zaccone, Karim e Menshawy (2017)

Na saída o sinal pelo qual o neurônio transmite sua atividade para fora, é calculada aplicando a função de ativação, também chamada de função de transferência, à soma ponderada das entradas. Essas funções têm uma dinâmica entre -1 e 1, ou entre 0 e 1.

Há um conjunto de funções de ativação que diferem em complexidade e saída:

- **Função de passo:** corrige o valor limite  $x$  (por exemplo,  $x = 10$ ). A função retornará 0 ou 1 se a soma matemática das entradas estiver acima ou abaixo do valor limite;
- **Combinação linear:** em vez de gerenciar um valor limite, a soma ponderada dos valores de entrada é extraída de um valor padrão; haverá um resultado binário, mas será expresso por uma saída positiva ( $+b$ ) ou negativa ( $-b$ ) da subtração.
- **Sigmoide:** produz uma curva sigmoide com tendência  $S$ . Muitas vezes, a função sigmoide refere-se a um caso especial da função logística.

Dessas formas mais simples, utilizadas na prototipagem dos primeiros neurônios artificiais, passam depois para formas mais complexas que permitem uma maior caracterização do funcionamento do neurônio. Sendo elas:

- Função tangente hiperbólica;
- Função de base radial;
- Função de seção cônica;

- Função *Softmax*.

Deve-se lembrar que a rede e, em seguida, os pesos nas funções de ativação, serão treinados. Como a seleção da função de ativação é uma tarefa importante na implementação da arquitetura de rede, estudos indicam diferenças marginais em qualidade de saída se a fase de treinamento for realizada adequadamente (ZACCONE; KARIM; MENSRAWY, 2017).

### **2.3.4 Como uma rede neural aprende?**

O processo de aprendizagem de uma rede neural configura-se como um processo iterativo de otimização dos pesos, sendo, portanto, do tipo supervisionado. Os pesos são modificados com base no desempenho da rede em um conjunto de exemplos pertencentes ao conjunto de treinamento, onde a categoria a que pertencem é conhecida. O objetivo é minimizar uma função de perda, que indica o grau em que o comportamento da rede se desvia do desejado. O desempenho da rede é então verificado em um conjunto de teste que consiste em objetos diferentes daqueles do conjunto de treinamento (ZACCONE; KARIM; MENSRAWY, 2017).

### **2.3.5 Arquiteturas de RNA**

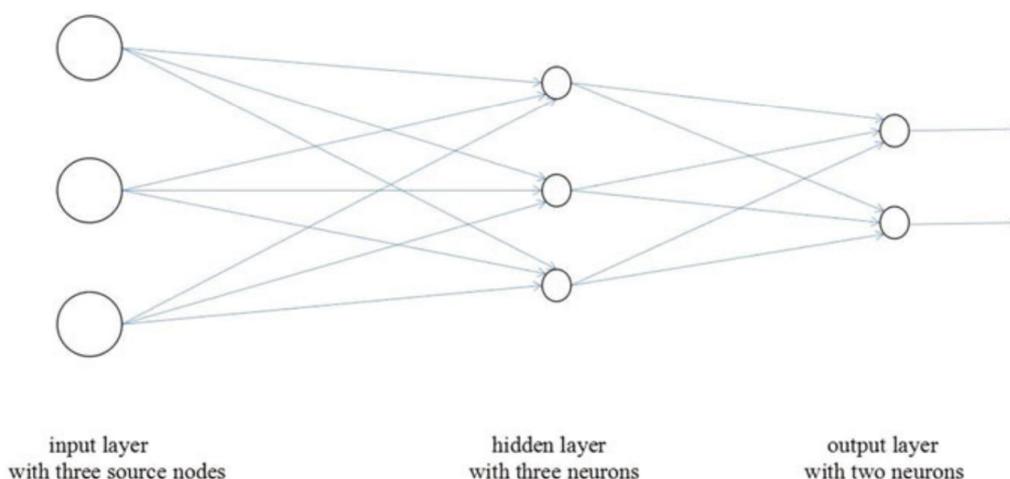
A RNA possui uma estrutura de unidades interconectadas de grande número de neurônios. Cada neurônio da rede consegue receber sinais de entrada, processá-los e enviar um sinal de saída. Consiste em um conjunto de sinapses ponderadas, um somador para somar os dados de entrada ponderados pela respectiva força sináptica, e uma função de ativação para limitar a amplitude da saída do neurônio. As arquiteturas de rede têm duas classes fundamentalmente diferentes, são elas as redes MLP e as recorrentes (ZHANG, 2018).

#### **2.3.5.1 Redes Neurais MLP**

As MLP estão sendo usadas atualmente em uma variedade de aplicações climáticas e ambientais com grande sucesso. Sua estrutura se resume a um número de neurônios organizados em camadas. Cada neurônio em uma camada está conectado com todos os neurônios da camada anterior. Essas conexões não são todas iguais; cada conexão pode ter uma força diferente ou peso.

Uma MLP consiste em uma camada de entrada de nós de origem, uma ou mais camadas ocultas e uma camada de saída de neurônios (Figura 6). As camadas ocultas na rede não são vistas diretamente da camada de entrada ou saída da rede. Elas permitem que a rede neural extraia os recursos estatísticos de ordem superior de sua entrada. Neurônios em camadas ocultas são correspondentemente chamados de neurônios ocultos. Esses neurônios ocultos têm a função de intervir entre entrada externa e a saída de rede de alguma maneira útil.

Figura 6 – Estrutura de uma MLP com apenas uma camada oculta



Fonte: Zhang (2018)

### 2.3.5.2 Rede Neural Recorrente (RNR)

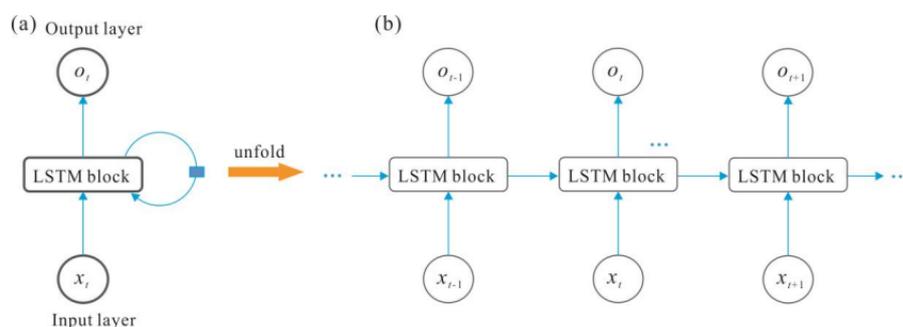
A característica fundamental de uma RNR é que ela contém pelo menos uma conexão de feedback, para que as ativações possam fluir em um loop. Ela permite que as redes façam processamento temporal e aprendam sequências, por exemplo, executem reconhecimento/reprodução de sequência ou associação/previsão temporal (ZACCONE; KARIM; MENSRAWY, 2017). Sua estrutura consiste em uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída (Figura 7). A camada oculta treina continuamente a rede conforme o tempo de entrada da sequência (Figura 8).

A entrada de dados  $X (x_1, x_2, \dots, x_i, \dots, x_{n_1})$  e saída de dados  $Y (y_1, y_2, \dots, y_k, \dots, y_{n_3})$  são conectados pela camada oculta  $H (h_1, h_2, \dots, h_j, \dots, h_{n_2})$ , onde  $n_1, n_2$  e  $n_3$  representam o número total de entradas, neurônios ocultos e saídas. O número de nós das camadas de entrada e saída são determinados pelos dados de entrada e saída do modelo. Como um parâmetro importante que afeta o desempenho do modelo, o número de nós da camada oculta é geralmente determinado por um método de tentativa e erro no processo de treinamento. A matriz de coeficiente de peso é representada por  $W$  (por exemplo,  $W_{xh}$  representa a matriz de coeficiente de peso da camada de entrada para a camada oculta), o vetor de deslocamento é representado por  $b$  ( $b_h$  representa o vetor de deslocamento da camada oculta) e a função de ativação é representada por  $f()$ . Com isso em mente, o processo de aprendizado e treinamento da RNR é o seguinte:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (2.1)$$

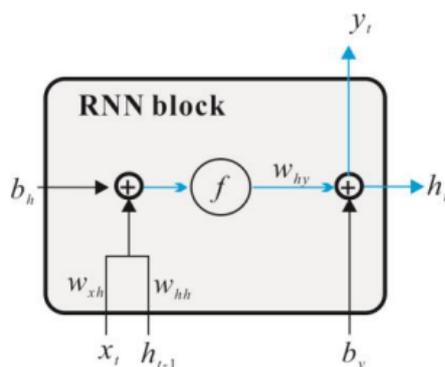
$$y_t = W_{hy}h_t + b_y \quad (2.2)$$

Figura 7 – Rede Neural Recorrente (a) e diagrama computacional, onde o quadrado azul indica um atraso de passo de tempo (b).



Fonte: Zhang et al. (2019)

Figura 8 – Estrutura simples de uma célula da camada oculta de uma RNR (b).



Fonte: Zhang et al. (2019)

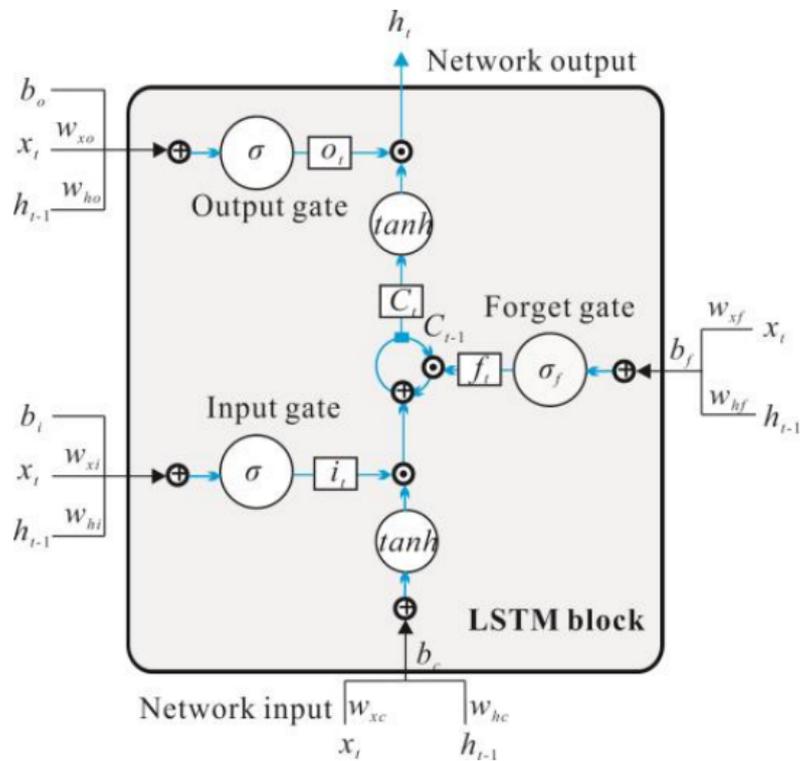
### 2.3.5.2.1 Long Short-Term Memory (LSTM)

Embora uma RNR possa lidar efetivamente com séries temporais não lineares, ainda existem algumas observações: devido aos problemas de desaparecimento e explosão de gradiente, uma RNR não pode capturar as características temporais, especialmente as dependências de longo prazo (HOCHREITER et al., 2001). O treinamento de modelos RNR precisa determinar o comprimento da sequência de entrada com antecedência, mas é difícil obter o valor ideal deste parâmetro na aplicação prática. Para resolver esses problemas, o modelo LSTM foi inventado. A diferença entre o modelo LSTM e o modelo RNR é que o bloco RNR na camada oculta é substituído pelo bloco LSTM. Isso permite modelar uma capacidade de memória de longo prazo (SAK; SENIOR; BEAUFAYS, 2014). O bloco LSTM foi obtido de Gers, Schmidhuber e Cummins (2000), e consiste numa porta de entrada, uma célula de memória, uma porta de esquecimento, e uma porta de saída, como mostrado na Figura 9.

O primeiro passo do LSTM é decidir quais dados devem ser removidos da célula Estado. Essa decisão é feita pela camada citada anteriormente: o portão de esquecimento. Ele irá gerar um valor de  $f_t$  entre 0 e 1, baseado na saída do momento anterior  $h_{t-1}$  e na entrada atual  $x_t$  para decidir se deve deixar passar a informação inteira ou parcialmente  $C_{t-1}$  que foi produzida no momento anterior, representado pela Equação (2.3):

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (2.3)$$

Figura 9 – Estrutura de um bloco LSTM.



Fonte: Zhang et al. (2019)

A segunda etapa é decidir quais novos dados serão salvos no estado da célula. Isso tem duas partes: primeiro, a porta de entrada determina quais valores serão atualizados. Em seguida, a célula de memória cria um vetor de novos valores candidatos  $C_t$ , que podem ser adicionados ao estado. Mais tarde, os valores produzidos por essas duas partes serão combinados para atualizar os dados:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2.4)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (2.5)$$

O passo final é determinar a saída do modelo. Primeiro, uma saída inicial é obtida através da camada sigmoide (porta de saída) e, em seguida, o valor  $C_t$  é redimensionado para -1 a 1 pela função  $\tanh$  e multiplicado pela saída da porta sigmoide para serem produzidas apenas as partes de destino:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (2.6)$$

$$h_t = o_t \times \tanh(C_t) \quad (2.7)$$

Nas Equações (2.6) e (2.7),  $f$ ,  $i$ ,  $C$ ,  $o$  e  $h$  são a porta de esquecimento, porta de entrada, célula, porta de saída e porta oculta, respectivamente, e  $\sigma$  e  $\tanh$  representam a função de ativação do portão (a sigmoide logística neste papel) e a função de ativação tangente hiperbólica, respectivamente,  $W$  representa o peso correspondente da matriz de coeficiente.

### 2.3.5.2.2 *Deep Bidirectional LSTM (BiLSTM)*

Os LSTMs bidirecionais profundos (SCHUSTER; PALIWAL, 1997) são uma extensão do modelo LSTM descrito anteriormente, nos quais, dois LSTMs são aplicados aos dados de entrada. Na primeira rodada, um LSTM é aplicado na sequência de entrada (ou seja, camada direta). Na segunda rodada, a forma inversa da sequência de entrada é alimentada no modelo LSTM (ou seja, camada inversa). Aplicar o LSTM duas vezes leva a melhorar o aprendizado de dependências de longo prazo e, assim, conseqüentemente irá melhorar a precisão do modelo (BALDI et al., 1999).

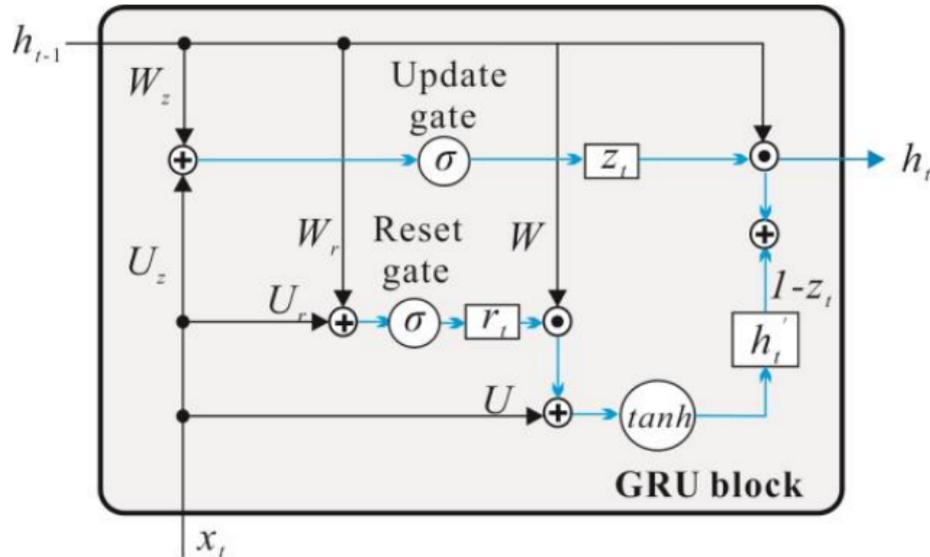
### 2.3.5.2.3 *Gated Recurrent Unit (GRU)*

O modelo LSTM evoluiu em muitas variantes, onde uma das mais bem sucedidas é o GRU, proposto por Chung et al. (2014). GRU é uma versão simplificada do LSTM, e seu princípio é semelhante ao LSTM, que controla a entrada, memória, saída e outros dados de informação através de um mecanismo de *gating*. A principal alteração corresponde à substituição das portas de entrada, esquecimento e saída da célula LSTM, pelos portões de atualização e de reinicialização combinados com o estado da célula e a saída (Figura 10). O principal processo de cálculo é o seguinte: primeiro, a situação das células da camada oculta é atualizada pelo portão de atualização e pelo portão de redefinição:

$$z_t = \sigma(x_t U_z + h_{t-1} W_z) \quad (2.8)$$

$$r_t = \sigma(x_t U_r + h_{t-1} W_r) \quad (2.9)$$

Figura 10 – Estrutura de um bloco GRU.



Fonte: Zhang et al. (2019)

Em seguida, as células de memória geram o conteúdo de memória do momento atual da camada oculta:

$$h'_t = \tanh(x_t U + r_t W h_{t-1}) \quad (2.10)$$

Finalmente, a saída da camada oculta do momento atual é calculada:

$$h_t = z_t h_{t-1} + (1 - z_t) h'_t \quad (2.11)$$

Nas Equações (2.8) e (2.11),  $z$ ,  $r$ ,  $h'$  e  $h$  são as portas de atualização, reinicialização, memória oculta e saída, respectivamente. Com isso,  $\sigma$  e  $\tanh$  representam a função de ativação do portão (o sigmoide logístico) e a função de ativação da tangente hiperbólica, respectivamente.  $W$  e  $U$  representam o peso correspondente da matriz de coeficiente.

O processo de treinamento dos três modelos vistos anteriormente adota o *Back Propagation Through Time* (BPTT), um princípio que é semelhante ao *Back Propagation* (BP) (WERBOS, 1990). Ele funciona da seguinte forma: Em primeiro lugar, os valores de saída das células da camada oculta são calculados conforme o método de cálculo direto. Em segundo lugar, os valores de erro da camada oculta são calculados com base na retro propagação do passo de tempo e da estrutura da rede. Além disso, os gradientes de cada peso são calculados segundo os termos de erro correspondentes. Finalmente, os coeficientes de peso são atualizados pelo *adaptive moment estimation algorithm*. Portanto, para as três redes neurais selecionadas, o

número de iterações é um dos principais parâmetros que afetam o desempenho do modelo como uma rede neural com BP.

Além disso, para o modelo de aprendizado profundo, o tamanho do lote também é um parâmetro importante que afeta o seu desempenho. Sendo representado pelo tamanho do volume de dados, ou seja, o número de amostras de treinamento em uma iteração. Em seguida, o modelo atualiza os parâmetros da rede (peso e viés) por calcular o erro entre o valor de saída real e o valor de saída esperado do conjunto de amostra (ZHANG et al., 2019).

## 2.4 *Teste t-Student*

Quando a análise de variância de um experimento mostra que as médias de tratamento não são estatisticamente iguais, é apenas lógico perguntar quais são as médias que diferem entre si. O pesquisador em geral gostaria de aplicar um teste para comparar médias, duas a duas. Considerando um experimento para comparar três tratamentos, A, B e C. Se a análise de variância mostrar que as médias desses tratamentos não são estatisticamente iguais, é bastante possível que o pesquisador procure um teste estatístico para comparar as médias de A e B, A e C, B e C (OLIVEIRA, 2008). O teste t Student pode ser usado para comparar médias, sendo duas médias A e B obtidas de grupos experimentais  $r_A$  e  $r_B$ , respectivamente, podem ser comparadas pela relação:

$$t = \frac{A - B}{\left(\frac{S_e^2}{r_A}\right) + \left(\frac{S_e^2}{r_B}\right)}, \quad (2.12)$$

onde  $S_e^2$  é a variância do resíduo estimada pela análise de variância. As médias comparadas por este teste diferirão estatisticamente se o valor calculado de t for maior que aquele tabelado segundo os graus de liberdade do erro (SAMPAIO, 2002). Caso contrário, serão consideradas estatisticamente iguais.

# Capítulo 3

## Materiais e Métodos

Neste capítulo se encontra detalhada a metodologia desta pesquisa, com detalhes a respeito da base de dados e a manipulação desses dados; aplicação das arquiteturas LSTM, BiLSTM e GRU; e elaboração do teste t-Student para análise estatística.

### 3.1 Ferramentas utilizadas

A construção das arquiteturas de RNR foi feita utilizando a linguagem de programação Python, a qual permite a utilização das principais bibliotecas *open source* voltadas para o estudo de *Deep Learning* e redes neurais: o TensorFlow e o Keras. O local de criação e execução do código foi o Google colab, uma plataforma online que permitiu a utilização das bibliotecas citadas anteriormente sem precisar realizar *download* no computador pessoal. Isso permite que o trabalho possa ser desenvolvido em qualquer máquina desde que se tenha uma boa conexão com a internet.

### 3.2 Base de dados

A base de dados possui valores diários de vazão e precipitação em seis afluentes da barragem Serra do Facão (Planaltina, Cristalina, Fazenda Campo Grande, Engenho Amarin, Fazenda São Domingos e Porto dos Pereiras), desde o ano de 1995 até 2001. A partir dos valores de precipitação, foi realizada uma média aritmética simples dos afluentes supracitados da barragem para cada dia.

Foi feita a divisão dos dados em treinamento, validação e teste dos modelos. Utilizou-se dos anos de 1995 até 1999 para treinamento, em que 20% desse período foi reservado para validação, e os últimos dois anos, 2000 a 2001, para realização dos testes.

Antes da realização do treinamento, as arquiteturas de RNR exigem que os dados de entrada sejam modelados para um formato tridimensional. Com isso, para cada dia de vazão a ser prevista, foi utilizado como entrada: o total de dias entre 31 de janeiro de 1995 até 31 de

dezembro de 1999 menos 60 *timesteps*, a soma dos 30 dias de vazão passada com 30 dias da média de precipitação e o *timesteps* para cada dia de vazão passada e média de precipitação. Já na saída dos modelos, utilizou-se o total de dias entre 1 de janeiro de 2000 até 20 de dezembro de 2001 menos 60 *time steps* e 12 dias de vazão futura para realizar o aprendizado supervisionado.

Tabela 1 – Relação de entradas e saídas calculadas nos modelos de previsão

Entradas			Saídas	
Total de dias entre 31/01/1995 e 31/12/1999 - time steps = 1736 dias	30 dias de vazão passada + 30 dias da média de precipitação = 60 dias	time steps = 60 dias	Total de dias entre 01/01/2000 e 20/12/2001 - time steps = 660 dias	Vazão futura dos 12 dias seguintes

Fonte: Autoria própria

### 3.3 Aplicação das arquiteturas LSTM, BiLSTM e GRU

Neste trabalho, foi feita a previsão de vazões diárias da barragem Serra do Facão, de 1 até 12 dias a frente. Os dados foram pré-processados para dentro de uma escala de 0.1 e 0.9. A fase de configuração das redes BiLSTM, LSTM e GRU, cujos parâmetros foram alterados até que fossem obtidos resultados satisfatórios para a previsão. Ao final desta etapa, todas as redes foram configuradas com duas camadas escondidas, a primeira com 256 neurônios e a segunda com 128 neurônios, utilizando-se da tangente hiperbólica como função de ativação. Em ambas as camadas foi realizado um *dropout* de 0.1 dos neurônios, para evitar *overfitting*. Por fim, na camada de saída foram utilizados 12 neurônios com a função de ativação Sigmoid.

Na compilação da LSTM, GRU e BiLSTM, foi utilizada a função de perda de Huber Loss, o otimizador RMSprop e o MAPE como para métrica de avaliação dos modelos, conforme a equação 3.1. Por fim, foram utilizadas 25 iterações e *batch size* de 128 para o treinamento das redes.

$$MAPE = \frac{1}{N} \left( \sum_{p=1}^N \frac{Z_p - Z_0}{Z_0} \right) * 100 \quad (3.1)$$

O procedimento de treinamento da LSTM, BiLSTM e GRU foi realizado de maneira semelhante. Para cada modelo foram realizadas 30 execuções de treino, utilizando as mesmas configurações descritas anteriormente, sendo que em cada execução foi utilizada uma função randômica de números inteiros com valores entre 1 e 1234 para a semente de atualização de pesos das três RNR. Isso se deve pelo fato de que com sementes diferentes, o treinamento das redes também diferirá, implicando diferentes valores de MAPE ao final de cada execução.

### **3.4 Aplicação do teste t-Student**

Para os 30 valores de MAPE de cada execução foi feita a média aritmética desses valores com o intuito de realizar o teste t-Student e comprovar se os modelos LSTM, BiLSTM e GRU são estatisticamente iguais ou diferentes. Caso sejam iguais, escolhe-se o que possui arquitetura mais simples e atual. Caso difiram, deve-se escolher o modelo que possui a menor média de MAPE ao longo das 30 execuções.

O modelo escolhido ao final destes procedimentos será comparado com a topologia de rede neural utilizando MLP com Random Forest, que utiliza a mesma base de dados da barragem nacional Serra do Facão e os mesmos parâmetros. Também foi analisado o desempenho com outros modelos utilizados para essa mesma barragem.

## Capítulo 4

### Resultados e Discussão

Neste capítulo estão dispostos os resultados obtidos, além de suas análises correspondentes para cada modelo de RNR utilizado neste trabalho.

#### 4.1 Resultados de média MAPE

O resultado de médias do MAPE para LSTM, BiLSTM e GRU foram 16%, 17.5% e 2.

Tabela 2 – Resultados de menor erro MAPE, semente e média do erro MAPE para 30 rodadas de treino das RNR

Nome do modelo	LSTM	BiLSTM	GRU
Valor do menor MAPE ao longo de 30 rodadas de treino (%)	15.7%	16.2%	14.6%
Valor da(s) semente(s) do respectivo MAPE de menor valor ao longo de 30 rodadas de treino (inteiro)	12, 287, 358 e 173	695, 414 e 718	480
Média do MAPE para 30 rodadas de treino (%)	16%	17.5%	16.3%

Fonte: Autoria própria

#### 4.2 Resultados da análise do teste t-Student

A regra de decisão utilizada na rejeição de hipótese nula diz o seguinte: rejeitar a hipótese nula se o valor t for maior do que 2.0086 ou menor do que -2.0086. Esses valores são obtidos diretamente da tabela t para determinado número de amostras e graus de liberdade.

##### 4.2.1 Entre LSTM e BiLSTM

- **Hipótese nula:** Média de MAPE BiLSTM igual à média MAPE de LSTM;

- **Hipótese 1:** Média de MAPE BiLSTM diferente da média MAPE de LSTM.

Através de cálculos foi obtido o valor t de 5.6 para as médias de MAPE entre os modelos LSTM e BiLSTM. Portanto, como foi definida a regra de decisão: rejeitar a hipótese nula se t for maior que 2.0086 ou menor que -2.0086, e 5.6 preenche esses critérios, pode-se rejeitar a hipótese nula de que a média do MAPE de BiLSTM é igual da LSTM.

#### 4.2.2 Entre BiLSTM e GRU

- **Hipótese nula:** Média de MAPE BiLSTM igual à média MAPE de GRU;
- **Hipótese 1:** Média de MAPE BiLSTM diferente da média MAPE de GRU.

Através de cálculos foi obtido o valor t de 4.0 para as médias de MAPE entre os modelos GRU e BiLSTM. Portanto, como foi definida a regra de decisão: rejeitar a hipótese nula se t for maior que 2.0086 ou menor que -2.0086, e 4.0 preenche esses critérios, pode-se rejeitar a hipótese nula de que a média do MAPE de BiLSTM é igual da GRU.

#### 4.2.3 Entre LSTM e GRU

- **Hipótese nula:** Média de MAPE LSTM igual à média MAPE de GRU;
- **Hipótese 1:** Média de MAPE LSTM diferente da média MAPE de GRU.

Através de cálculos foi obtido o valor t de 1.1 para as médias de MAPE entre os modelos LSTM e BiLSTM. Portanto, como foi definida a regra de decisão: rejeitar a hipótese nula se t for maior que 2.0086 ou menor que -2.0086, e 1.1 não preenche esses critérios, não se pode rejeitar a hipótese nula de que a média do MAPE de LSTM é igual da GRU.

O modelo BiLSTM teve um desempenho pior do que os modelos LSTM e GRU, com média de MAPE estatisticamente diferente de ambos. Por isso, será descartado o modelo BiLSTM da comparação final com os demais modelos utilizados anteriormente na barragem nacional Serra do Facão.

Resta saber qual modelo escolher entre o LSTM e GRU visto que ambos, conforme o teste t-Student, foram considerados estatisticamente iguais. Nesse caso, o modelo a ser escolhido para a análise final será o modelo GRU. Ele foi mais otimizado e possuiu maior velocidade de processamento ao longo das execuções de treino.

Para o modelo GRU, será escolhido o treinamento que apresentou o menor valor do MAPE, o qual foi o de 14,6% para a previsão de vazão utilizando a vazão passada com a média de precipitação da UHE Serra do Facão. Com isso, foi possível observar que este modelo demonstra resultados inferiores ao modelo MLP com Random Forest para um mesmo período observado.

Porém, este modelo apresentou resultados melhores com os demais modelos também utilizados. Essa comparação está explicitada na Tabela 3.

Tabela 3 – Valores do MAPE para cada modelo computacional.

Este modelo (vazão passada e chuva futura)	MLP com RF (vazão passada e chuva futura)	SMAP	PREVIVAZ
14.6%	10%	17%	18%

Fonte: (MOTA, 2021) (MLP com RF), (ONS, NT 136/2012) (SMAP e PREVIVAZ).

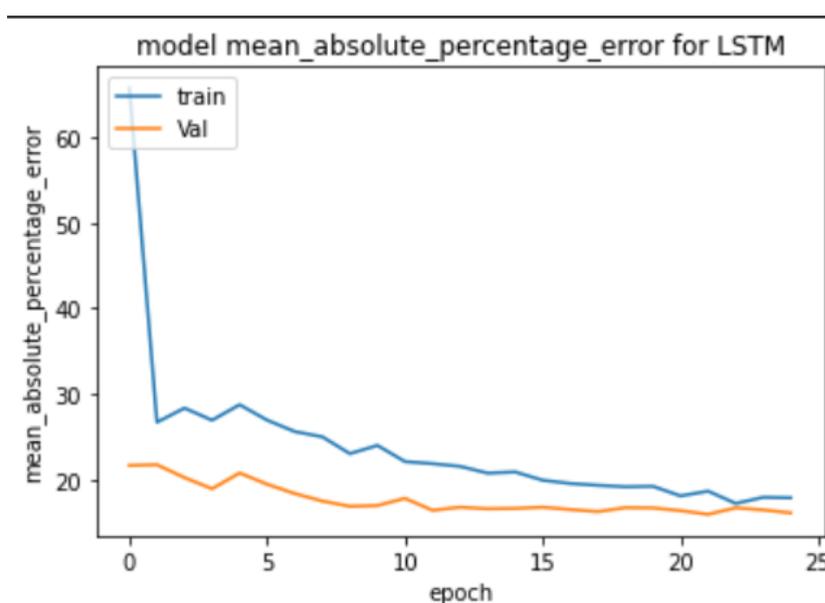
### 4.3 Análise gráfica das previsões

Agora será apresentado as análises gráficas de cada um dos modelos de RNR utilizados neste trabalho para mostrar a capacidade delas em modelos de previsão de vazão.

#### 4.3.1 LSTM: gráficos de MAPE e ocorrido vs previsto

O modelo LSTM apresentou, aproximadamente, menor MAPE igual a 16%. As respectivas sementes de atualização de pesos foram as de valores inteiro 12, 287, 358 e 173. O gráfico do MAPE está representado logo abaixo:

Figura 11 – Gráfico do MAPE do modelo LSTM

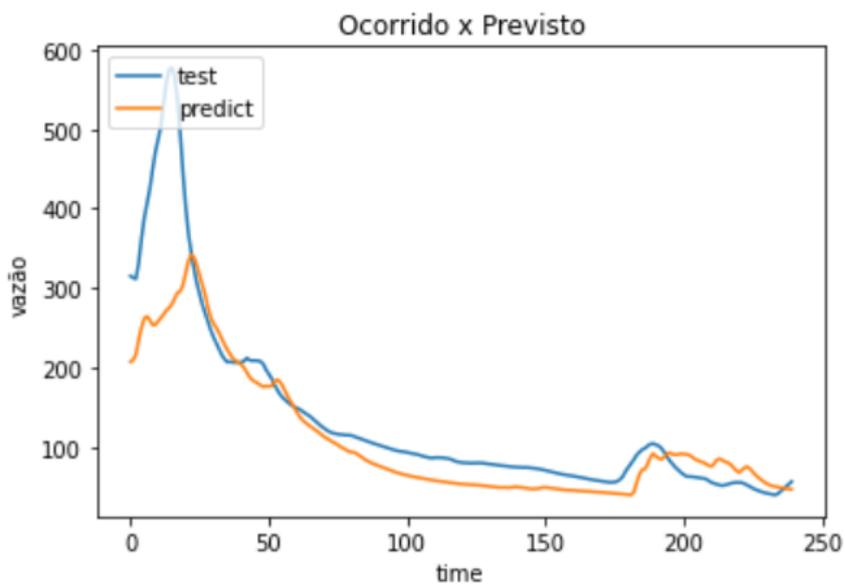


Fonte: Autoria própria.

O gráfico acima indica uma queda no valor MAPE do LSTM de aproximadamente 22% até 15,7%.

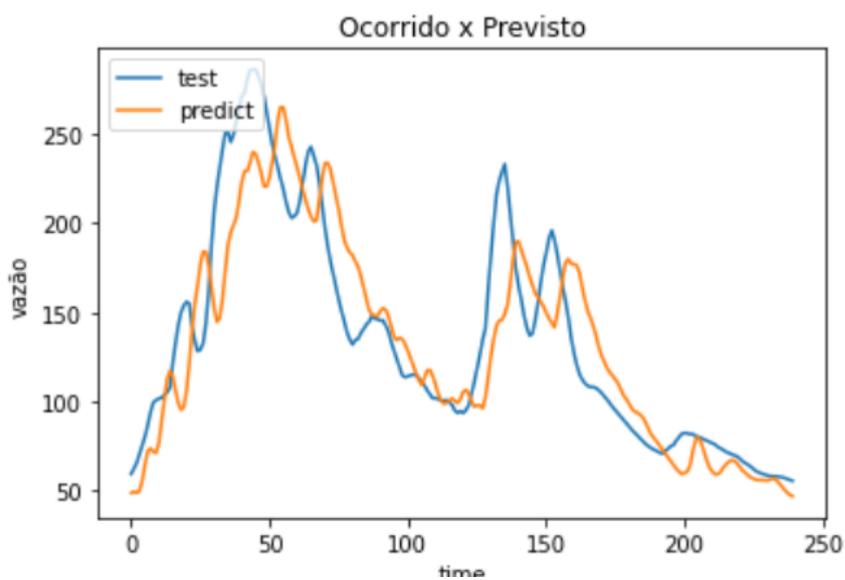
As previsões de vazões comparando o valor ocorrido com o valor previsto pelo LSTM estão expressos, respectivamente, nas Figuras 12, 13 e 14.

Figura 12 – Ocorrido x previsto 1 do modelo LSTM



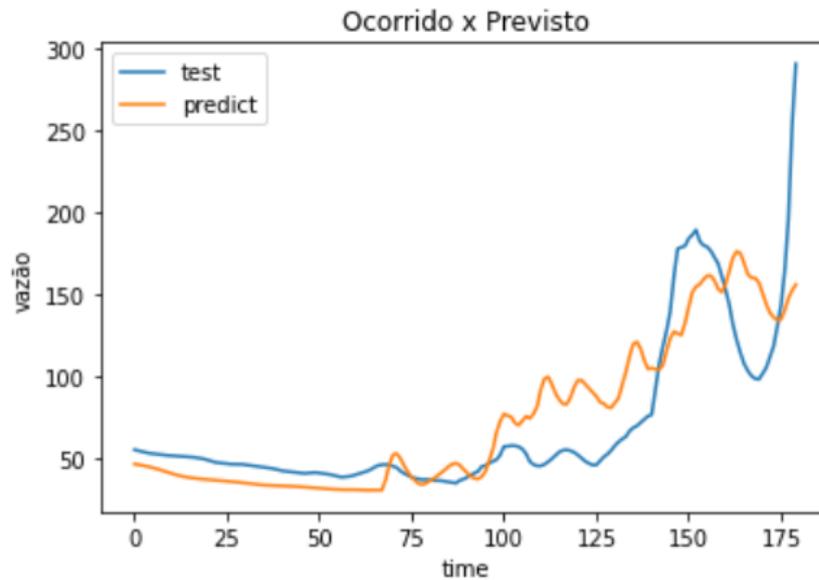
Fonte: Autoria própria.

Figura 13 – Ocorrido x previsto 2 do modelo LSTM



Fonte: Autoria própria.

Figura 14 – Ocorrido x previsto 3 do modelo LSTM

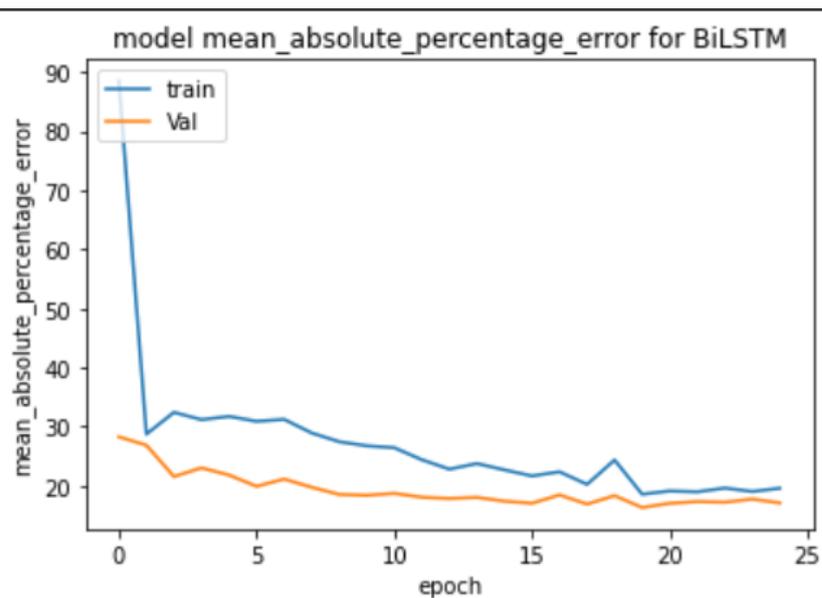


Fonte: Autoria própria.

#### 4.3.2 BiLSTM: gráficos de MAPE e ocorrido vs previsto

O modelo BiLSTM apresentou, aproximadamente, menor MAPE igual a 16.2%. A respectiva semente de atualização de pesos foi a de valor inteiro 695. O gráfico do MAPE está representado logo abaixo:

Figura 15 – Gráfico do MAPE do modelo BiLSTM

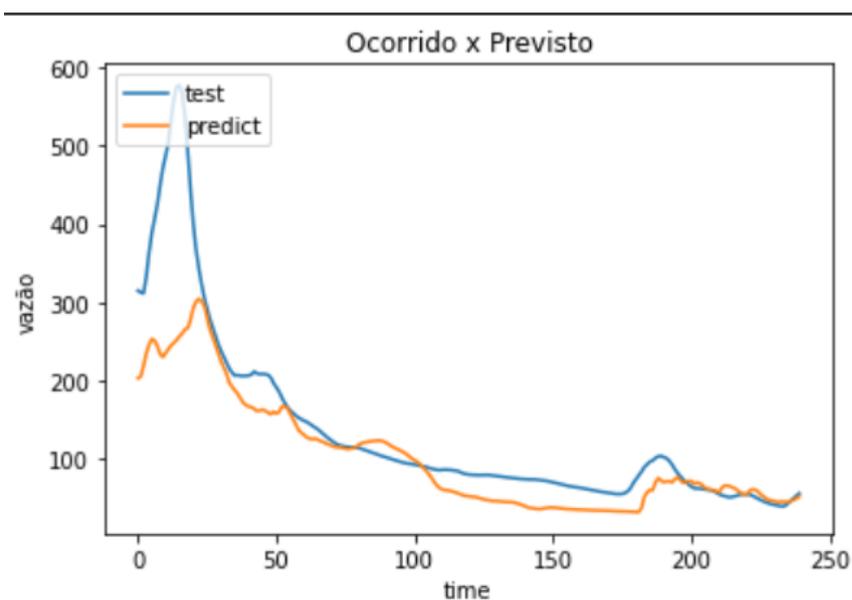


Fonte: Autoria própria.

O gráfico acima indica uma queda no valor MAPE do LSTM de aproximadamente 29% até 16,2%.

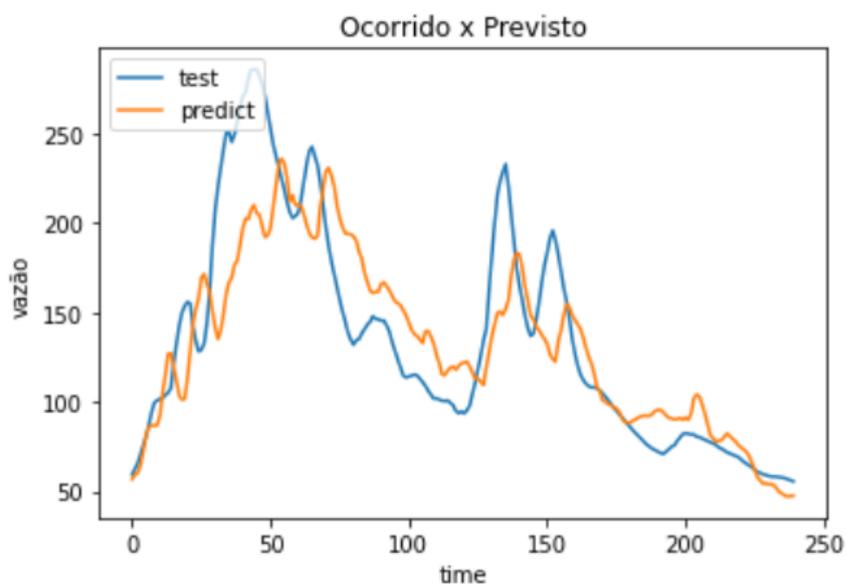
As previsões de vazões comparando o valor ocorrido com o valor previsto pelo BiLSTM estão expressos, respectivamente, nas Figuras 16, 17 e 18.

Figura 16 – Ocorrido x previsto 1 do modelo BiLSTM



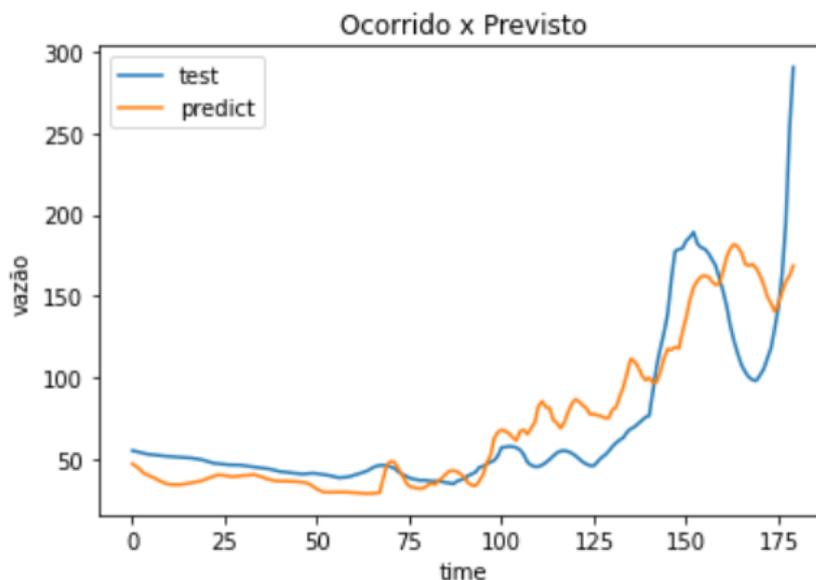
Fonte: Autoria própria.

Figura 17 – Ocorrido x previsto 2 do modelo BiLSTM



Fonte: Autoria própria.

Figura 18 – Ocorrido x previsto 3 do modelo BiLSTM

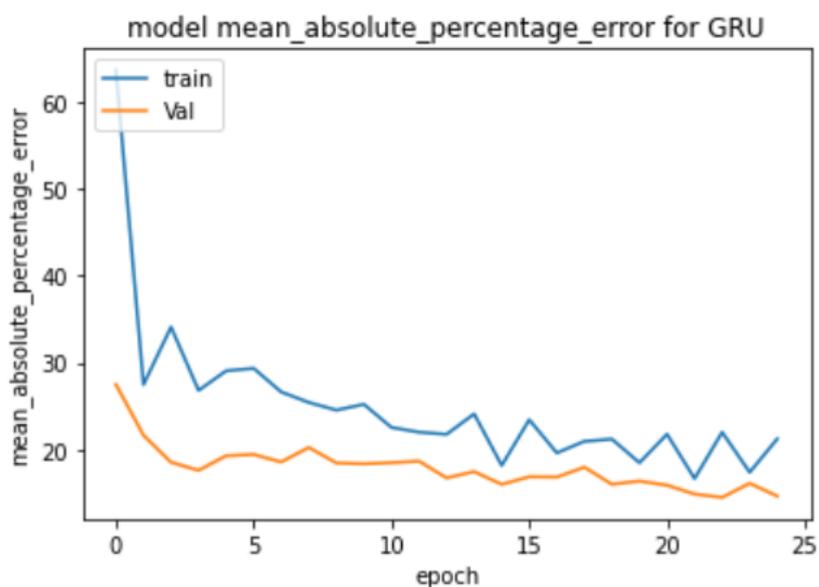


Fonte: Autoria própria.

### 4.3.3 GRU: gráficos de MAPE e ocorrido vs previsto

O modelo GRU apresentou, aproximadamente, menor MAPE igual a 14.6%. A respectiva semente de atualização de pesos foi a de valor inteiro 480. O gráfico do MAPE está representado logo abaixo:

Figura 19 – Gráfico do MAPE do modelo GRU

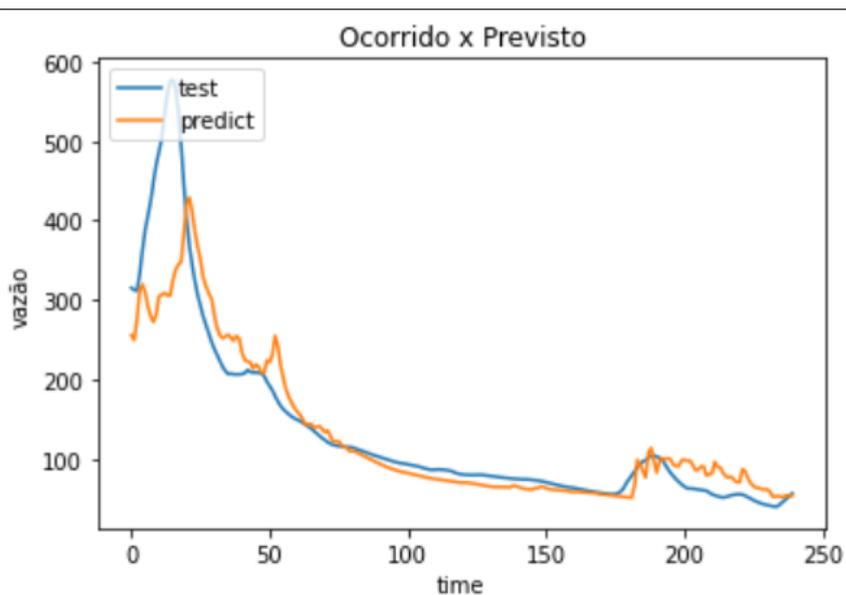


Fonte: Autoria própria.

O gráfico acima indica uma queda no valor MAPE do LSTM de aproximadamente 29% até 14,6%.

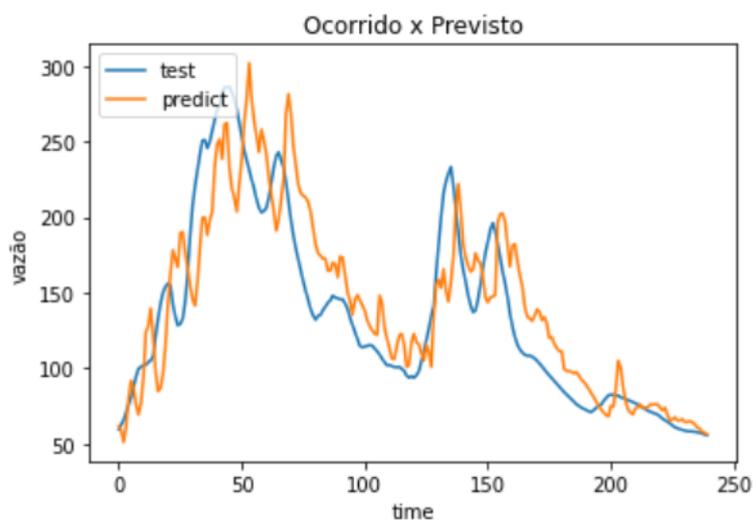
As previsões de vazões comparando o valor ocorrido com o valor previsto pelo GRU estão expressos, respectivamente, nas Figuras 20, 21 e 22.

Figura 20 – Ocorrido x previsto 1 do modelo GRU



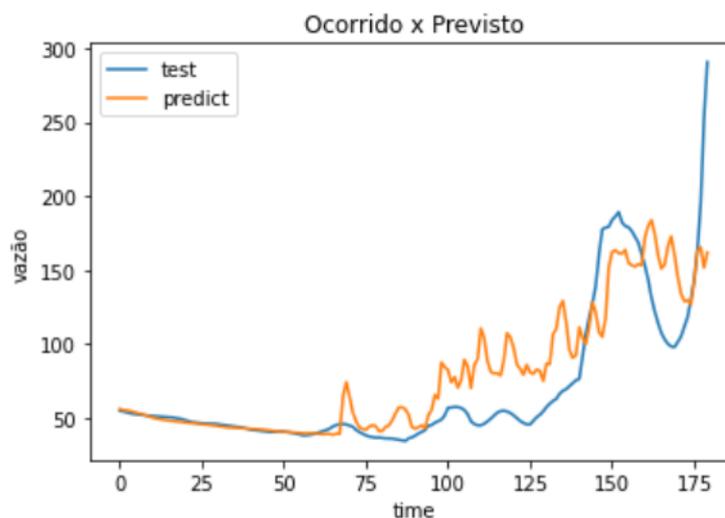
Fonte: Autoria própria.

Figura 21 – Ocorrido x previsto 2 do modelo GRU



Fonte: Autoria própria.

Figura 22 – Ocorrido x previsto 3 do modelo GRU



Fonte: Autoria própria.

Os gráficos de ocorrido vs previsto também mostram a perda de eficiência na previsão quando ocorrem pontos de acentuada subida ou descida nos dados de teste. Além da previsão ser feita para curtíssimo prazo, outro fator que prejudica as RNR, visto que foram feitas essencialmente para previsões de longo prazo. Porém a capacidade de previsão das Redes Neurais Recorrentes demonstraram valores bem promissores mesmo sendo um pouco inferiores ao modelo MLP com Random Forest utilizando a mesma base dados.

## Capítulo 5

### Conclusões e Trabalhos Futuros

Pode-se concluir que as Redes Neurais Recorrentes não foram melhores quando comparados com a pesquisa que utilizou MLP com Random Forest para previsão de vazões de curtíssimo prazo dos afluentes da UHE Serra do Facão. O que indica para dados de longo prazo a LSTM, BiLSTM e GRU provavelmente consigam obter melhores resultados porque elas são projetadas para esse tipo de situação. Porém elas obtiveram resultados melhores que os modelos publicados pelo ONS: PREVIVAZ e SMAP, o que demonstra o potencial de se utilizar esse tipo de Rede Neural para previsão de vazões naturais de modo geral, permitindo novas possibilidades de pesquisas a serem realizadas futuramente. Entre elas podem ser listados:

- A utilização de uma análise de previsão de longo prazo tanto para a UHE Serra do Facão, quanto para outras usinas do SIN;
- A combinação das redes neurais recorrentes com Random Forest para seleção de melhores parâmetros;
- A aplicação de outras arquiteturas de redes neurais recorrentes para esta usina hidrelétrica do SIN;
- Extensão da metodologia utilizada no presente trabalho para as demais usinas hidrelétricas do SIN.

## Referências

- BALDI, P.; BRUNAK, S.; FRASCONI, P.; SODA, G.; POLLASTRI, G. Exploiting the past and the future in protein secondary structure prediction. **Bioinformatics**, Oxford University Press, v. 15, n. 11, p. 937–946, 1999.
- BASIN, U. P. R. Modelagem de qualidade da água no rio são marcos, bacia hidrográfica do alto paraná. **A Revista DAE é classificada pelo QUALI/CAPES e está adicionada/indexada nas seguintes bases**, p. 20, 2020.
- CHUNG, J.; GULCEHRE, C.; CHO, K.; BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. **arXiv preprint arXiv:1412.3555**, 2014.
- ENGIE. **Hidrelétricas são fundamentais para o Brasil**. 2022. [Online; accessed 13-July-2022]. Disponível em: <<https://umsoplaneta.globo.com/patrocinado/engie/noticia/2022/02/17/hidreletricas-sao-fundamentais-para-o-brasil.ghtml>>.
- FERREIRA, C. d. C. et al. Previsão de vazões naturais diárias afluentes ao reservatório da uhe tucuruí utilizando a técnica de redes neurais artificiais. **Universidade Federal de Goiás**, 2012.
- FURNAS, E. **Usina Serra do Facão**. 2018. [Online; accessed 29-August-2022]. Disponível em: <<https://www.furnas.com.br/subsecao/130/usina-serra-do-facao?culture=pt>>.
- GERS, F. A.; SCHMIDHUBER, J.; CUMMINS, F. Learning to forget: Continual prediction with lstm. **Neural computation**, MIT Press, v. 12, n. 10, p. 2451–2471, 2000.
- GUILHON, L. G. F.; ROCHA, V. F.; MOREIRA, J. C. Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. **Revista Brasileira de Recursos Hídricos**, v. 12, n. 3, p. 13–20, 2007.
- HOCHREITER, S.; BENGIO, Y.; FRASCONI, P.; SCHMIDHUBER, J. et al. **Gradient flow in recurrent nets: the difficulty of learning long-term dependencies**. [S.l.]: A field guide to dynamical recurrent neural networks. IEEE Press In, 2001.
- MINE, M. R. M.; TUCCI, C. E. M. Previsão em tempo real de vazões afluentes a reservatórios de usinas hidrelétricas. **Rbrh: Revista Brasileira de Recursos Hídricos. Porto Alegre, RS. vol. 4, n. 2 (abr./jun. 1999), p. 73-95**, 1999.
- MOTA, K. B. Utilizando redes neurais e random forest para previsão do processo chuva vazão. **Trabalho de Conclusão de Curso, Engenharia da Computação, Universidade de Pernambuco. Recife. p. 43**, 2021.
- OLIVEIRA, A. F. G. Testes estatísticos para comparação de médias. **Revista Eletrônica Nutritime**, v. 5, n. 6, p. 777–788, 2008.

ONS. **Aplicação de modelos de previsão de vazões a curto prazo na sub bacia do alto do rio Paranaíba Relatório final**, NT 136/2012.

PAIVA, L.; BARROS, A.; VALENÇA, M. Aplicação do modelo neuro3m para previsão de vazões naturais afluentes ao reservatório de três marias. **In Anais do XVIII Simpósio Brasileiro de Recursos Hídricos, Campo Grande**, 2009.

SAK, H.; SENIOR, A.; BEAUFAYS, F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. **arXiv preprint arXiv:1402.1128**, 2014.

SAMPAIO, I. B. M. Estatística aplicada à experimentação animal. **Belo Horizonte: Fundação de Estudo e Pesquisa em Medicina Veterinária e Zootecnia**, v. 2, p. 265, 2002.

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

SEFAC. **Serra do Facão Energia S.A: Nossa História**. 2018. [Online; accessed 29-August-2022]. Disponível em: <<https://sefac.com.br/empresa/nossa-historia/>>.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A.; LIBONI, L. H. B.; ALVES, S. F. dos R. **Artificial Neural Networks**. [S.l.]: Springer International Publishing, 2017. ISBN 978-3-319-43161-1.

VALENÇA, M. J. S. **Redes neurais: princípios e prática**. [S.l.]: Editora Livro rápido. Olinda, PE: Ed. Do Autor, 284p, 2005b.

WERBOS, P. J. Backpropagation through time: what it does and how to do it. **Proceedings of the IEEE**, IEEE, v. 78, n. 10, p. 1550–1560, 1990.

YANG, T.; ASANJAN, A. A.; FARIDZAD, M.; HAYATBINI, N.; GAO, X.; SOROOSHIAN, S. An enhanced artificial neural network with a shuffled complex evolutionary global optimization with principal component analysis. **Information Sciences**, Elsevier, v. 418, p. 302–316, 2017.

ZACCONE, G.; KARIM, M. R.; MENSRAWY, A. **Deep learning with TensorFlow**. [S.l.]: Packt Publishing Ltd, 2017.

ZHANG, D.; PENG, Q.; LIN, J.; WANG, D.; LIU, X.; ZHUANG, J. Simulating reservoir operation using a recurrent neural network algorithm. **Water**, MDPI, v. 11, n. 4, p. 865, 2019.

ZHANG, Z. **Artificial Neural Network**. Springer International Publishing, 2018. 1-35 p. Disponível em: <[http://link.springer.com/10.1007/978-3-319-67340-0\\_1](http://link.springer.com/10.1007/978-3-319-67340-0_1)>.