

Previsão de nível e vazão de água de um rio usando rede Perceptron Multi-Camada: um estudo de caso do Rio Itapemirim

Eduardo Henrique Próspero Souza¹, Vinicius Marques de Oliveira²,
Jefferson Oliveira Andrade², Karin Satie Komati²

¹ Coordenação de Sistemas de Informação

²Programa de Pós-graduação em Computação Aplicada (PPComp)
Instituto Federal do Espírito Santo (IFES)
Av. dos Sabiás, 330 – 29.166-630 – Serra – ES – Brasil

duvr dx@gmail.com, vinicius.armlock@hotmail.com,
{kkomati, jefferson.andrade}@ifes.edu.br

Abstract. *This article presents a study on the use of multivariate time series models to predict the water level or flow of the Itapemirim River in Brazil. Two datasets were used in the experiments; one where the target variable is the river level based on flow data from 3 river gauge stations (“Sensors and Level”) and another where the target variable is the flow from a station near the urban area based on the flow data from the same 3 stations of the first dataset (“Sensors Only”). Multi-Layer Perceptron (MLP) network were developed using the Neural Architecture Search (NAS) algorithm to find the best architecture for each dataset. A comparison of the results, based on the R^2 coefficient, of the MLP models was made with the mathematical regression model used in the Itapemirim River Alert System (SARI), a system that monitors and predicts the river’s water level. The conclusion is that the results with the “Sensors Only” dataset achieved an R^2 similar to the mathematical model, suggesting it could be used as an alternative method for monitoring the flow of the Itapemirim River.*

Resumo. *Este artigo apresenta um estudo sobre o uso de modelos de séries temporais para prever o nível ou a vazão da água do Rio Itapemirim no Brasil. Foram usadas duas bases de dados nos experimentos, uma em que a variável alvo é o nível do rio através de dados de vazão de 3 estações fluviométricas (“Sensores e Nível”) e outra em que a variável alvo é a vazão de uma estação próxima da área urbana através dos dados da vazão das mesmas 3 estações da primeira base (“Somente Sensores”). Foram criados modelos do tipo Perceptron Multi-Camadas (MLP) com o uso de algoritmo de Neural Architecture Search (NAS) para encontrar a melhor arquitetura para cada base de dados. Foi realizada a comparação dos resultados, baseados no coeficiente R^2 , dos modelos MLP com o modelo matemático de regressão usado no Sistema de Alerta do Rio Itapemirim (SARI), que é um sistema que monitora e prevê o nível da água do rio. A conclusão é que o resultado com a base de dados “Somente Sensores” obteve R^2 similar ao modelo matemático, podendo ser usado como um segundo método de acompanhamento do Rio Itapemirim, através de sua vazão.*

1. Introdução

Diversas cidades no Brasil enfrentam desafios ambientais significativos devido à sua proximidade com rios, e um dos problemas é a ocorrência de inundações quando o nível dos rios sobe, resultando em perdas materiais e até mesmo perdas de vidas humanas [Saito et al. 2023]. As inundações têm um impacto negativo, tanto do ponto de vista socioeconômico, quanto em termos de qualidade de vida das populações afetadas [Farias and Mendonça 2022]. Portanto, a capacidade de prever com antecedência a possibilidade de inundações é importante, permitindo que as autoridades tomem medidas preventivas para reduzir danos.

A previsão do nível da água em um rio pode ser modelada como um conjunto de séries temporais ou por séries temporais multivariadas [Lima et al. 2023]. A previsão de séries temporais consiste em analisar e modelar dados sequenciais ao longo do tempo, com o objetivo de fazer previsões futuras com base em padrões passados. Uma série temporal (univariada) é uma sequência de pontos de dados que ocorrem em ordem sucessiva ao longo de algum período de tempo [Morettin and Toloï 2018]. Uma série temporal multivariada é uma série temporal que apresenta mais de uma variável dependente do tempo, e cada variável depende não apenas de seus valores anteriores, mas também dos valores anteriores das outras variáveis [Morettin and Toloï 2020]. Frequentemente, uma série temporal é uma sequência de amostragens tomada em pontos sucessivos igualmente espaçados n tempo, e portanto, é uma sequência de dados em tempo discreto [Oliveira et al. 2021].

A proposta deste trabalho é comparar os resultados de modelos de redes neurais artificiais (RNA) do tipo Perceptron Multi-Camadas (MLP, do inglês Multilayer Perceptrons) com o resultado do modelo matemático de regressão usado no SARI (Sistema de Alerta do Rio Itapemirim)¹ para prever o nível ou vazão de água do Rio Itapemirim. O Rio Itapemirim se localiza na cidade de Cachoeiro de Itapemirim, que fica no estado do Espírito Santo, e que passou por problemas de alagamento [De Souza and dos Santos 2018].

Redes neurais dependem da seleção da arquitetura mais adequada ao problema. A maior parte dos trabalhos recentes nessa área, se debruça em comparar e propor diferentes arquiteturas para os mais diversos problemas. Embora seja possível se otimizar parâmetros através da experimentação, Miikkulainen e colegas [Miikkulainen et al. 2019] defendem que a abordagem de tentativa e erro é muito dispendiosa para redes neurais que podem chegar a centenas de milhares de hiper-parâmetros. Uma solução é o AutoML (aprendizado de máquina automatizado), que é o processo de automação total ou parcial da aplicação do aprendizado de máquina a um problema [Telikani et al. 2021]. O AutoML pode ser implementado através de dois tipos de tarefas [Hutter et al. 2019]: o NAS (do inglês *Neural Architecture Search*) e a otimização de hiper-parâmetros. O NAS é uma abordagem para automatizar o ajuste de topologia de RNAs [Zoph and Le 2016] e a otimização de hiper-parâmetros pode ser definida como o problema de buscar o conjunto de hiper-parâmetros que levem ao modelo de aprendizado de máquina com maior desempenho [Feurer and Hutter 2019].

Para os experimentos deste trabalho, usa-se a proposta de NAS feita por

¹<https://servicos.agerh.es.gov.br/sari/>

[Oliveira 2022] para encontrar a melhor arquitetura de MLP para o problema. Os experimentos serão executadas em duas bases de dados: a primeira é a mesma do SARI composta por dados telemétricos de vazão de 3 estações fluviométricas e a segunda é a união da primeira base de dados com o dado do nível disponibilizado pela Defesa Civil. Os resultados são baseados no coeficiente R^2 , que é a mesma métrica usada pelo SARI.

Esse artigo está organizado da seguinte maneira: na Seção 2 são apresentados os trabalhos relacionados, a Seção 3 descreve os materiais e métodos utilizados ao decorrer do trabalho. Os resultados obtidos estão presentes na Seção 4. E por fim, a conclusão se encontra na Seção 5.

2. Trabalhos Correlatos

No trabalho de [Lima et al. 2023], foram comprados diversos modelos de aprendizado de máquina clássicos, incluindo redes neurais recorrentes (RNN, do inglês Recurrent Neural Network), floresta aleatória e Support Vector Regression (SVR), e duas propostas de modelos o ARIMA-RF e o ARIMA-RNN, para prever o nível de água em um caso de estudo envolvendo o Canal São Gonçalo. A base de dados um total de 14.369 registros, sendo dividido em 70% para treino e 30% para testes. O modelo ARIMA-RF (erro médio absoluto de 0,02) apresentou melhores resultados que o ARIMA-RNN (erro médio absoluto de 0,058).

O trabalho de [Brandão et al. 2023] apresenta uma análise comparativa de duas abordagens de RNAs para prever os níveis de cota dos rios usando dados de séries temporais de quatro estações hidrológicas na bacia do Rio Madeira, com o objetivo de antecipar os níveis do rio em períodos de previsão de 15 e 30 dias. As abordagens são: RNAs do tipo MLP e Redes Neurais Recorrentes do tipo LSTM (Long Short-Term Memory). Foram exploradas diversas configurações arquiteturais para as MLP, enquanto as LSTM consistiam de uma única camada oculta recorrente. A métrica de desempenho adotada foi o coeficiente de determinação (R^2), que quantifica a capacidade preditiva das RNAs em relação aos dados reais. Para avaliar o desempenho dos modelos, realizou-se validação cruzada com divisão de séries temporais. Os resultados apontam que, em algumas situações, as LSTM superaram as MLP, principalmente em horizontes de previsão mais curtos. No entanto, para períodos de previsão mais longos, não foram observadas diferenças estatisticamente significativas no desempenho. Em resumo, a pesquisa avalia a aplicação de Redes Neurais Artificiais, especificamente MLPs e LSTMs, para a previsão dos níveis do rio ao longo do Rio Madeira, no Brasil. Os resultados sugerem que as LSTM podem ser mais eficazes em previsões de curto prazo, mas não apresentam uma vantagem evidente em horizontes de previsão mais longos.

Na nota técnica [Bof 2022] disponibilizada pela Agência Estadual de Recursos Hídricos do Espírito Santo (AGERH), foi desenvolvido um modelo matemático de regressão para prever inundações na área urbana de Cachoeiro de Itapemirim, esse modelo é utilizado no SARI. O modelo utiliza dados de vazão de estações fluviométricas fornecidos pela Agência Nacional de Águas (ANA). O modelo logarítmico de regressão é baseada na soma das vazões de três estações e possui um alto coeficiente de determinação (R^2) de 0,9595, demonstrando sua eficácia. Durante testes práticos em eventos de chuva em dezembro de 2022, o modelo foi calibrado para melhorar sua precisão, com um ajuste na equação que reduziu a cota estimada em 30 centímetros. O modelo emite estados de

criticidade com base no nível do rio, classificando-os como “Estado de Observação” (menor ou igual a 2 metros), “Estado de Atenção” (entre 2 e 3 metros) e “Estado de Alerta” (maior que 3 metros).

No trabalho de [Lima et al. 2023], são utilizadas RNNs entre outras técnicas tradicionais assim como o ARIMA, diferente do nosso trabalho que utiliza apenas MLPs. No trabalho de [Brandão et al. 2023] temos o uso de MLPs, mas em um estudo de caso com dados do Rio Madeira. Já em [Bof 2022] temos o uso de regressão junto com os dados do Rio Itapemirim disponibilizados pela ANA e Defesa Civil, e no nosso trabalho utilizamos os dados dos sensores da ANA e MLPs.

3. Materiais e Métodos

3.1. Bases de Dados

O primeiro conjunto de dados dos experimentos é a mesma base de dados usada pelo SARI, que é mantido e desenvolvido pela AGERH, e utiliza um conjunto de dados processado dos dados dos sensores das estações que são disponibilizados pela Agência Nacional de Águas e Saneamento Básico (ANA) [ANA 2023]. Daqui por diante, denominaremos esta base como “Sensores e Nível”.

Tabela 1. Amostra dos dados da base “Sensores e Nível”

Data	Hora	Nível (m)	Vazão:	Vazão:	Vazão:	Soma de vazões das 3 estações
			Francisco Gross (m ³ /s)	São João (m ³ /s)	Pacotuba (m ³ /s)	
19/02/2021	17	1,05	145,37	25,67	46,83	217,87
19/02/2021	18	1,23	163,12	25,88	51,33	240,33
19/02/2021	19	1,33	177,59	26,72	56,07	260,38
19/02/2021	20	1,44	192,64	28	62,08	282,72
19/02/2021	21	1,55	206	28,87	71,74	306,61

Tabela 2. Amostra dos dados da base “Somente Sensores”

Data e Hora	Vazão:	Vazão:	Vazão:	Vazão:
	Francisco Gross (m ³ /s)	Pacotuba (m ³ /s)	São João (m ³ /s)	Ilha da Luz (m ³ /s)
2023-10-02 22:00:00	86,63	103,38	13,48	260,72
2023-10-02 21:00:00	88,21	107,19	13,70	265,70
2023-10-02 20:00:00	92,99	113,03	13,70	271,98
2023-10-02 19:00:00	97,83	117,01	14,15	279,60
2023-10-02 18:00:00	101,08	121,05	14,15	286,04

Na Tabela 1, retirada de [Bof 2022], temos uma amostra dos dados, apenas as cinco primeiras linhas, que foram utilizados para treinar o modelo de regressão do SARI. A primeira coluna é a data da coleta, seguida pela hora, o nível medido em metros (no ponto do rio nas coordenadas -20,8441 e -41,1220 coletado pela Defesa Civil Municipal²) e 3 colunas com os dados dos sensores de vazão em m^3/s das estações de Francisco Gross (marcado na cor amarela na Figura 1), São João (marcado na cor vermelha na Figura 1) e Pacotuba (marcado na cor azul na Figura 1). A última coluna é a soma da vazão das 3 estações. O período de coleta inicia no dia 19-02-2021 às 17h e vai até o dia 10-02-2023 às 1h.

²<https://www.cachoeiro.es.gov.br/defesa-civil/plantao/>

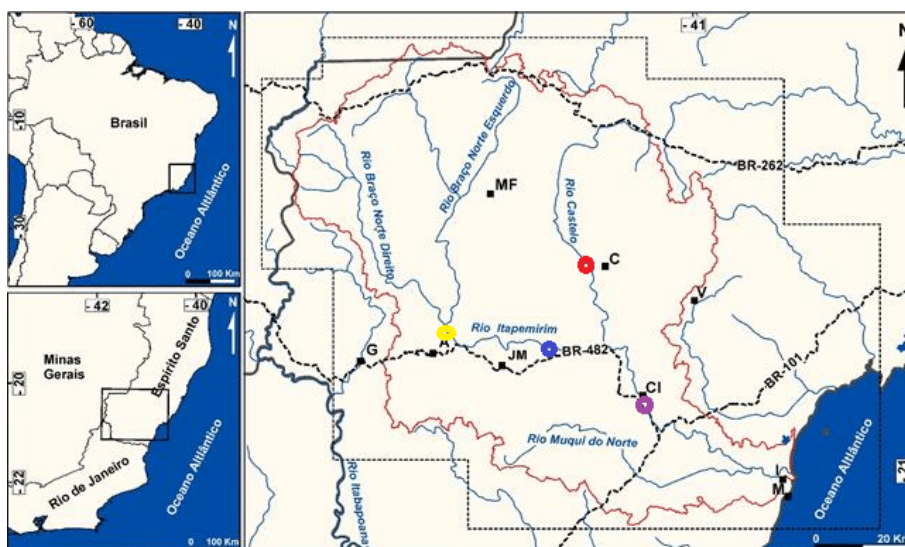


Figura 1. Localização das estações. Adaptado de [Neves et al. 2016]

Para construirmos nosso segundo conjunto de dados, daqui por diante denominado por “Somente Sensores”, utilizamos os dados de vazão das mesmas 3 estações do SARI, porém ao invés de utilizarmos o nível em metros coletado pela Defesa Civil Municipal (a terceira coluna da Tabela 1), utilizamos os dados de vazão da estação Ilha da Luz (marcado na cor lilás na Figura 1). Esta estação é a mais próxima da área urbana do município de Cachoeiro de Itapemirim (marcado como CI na Figura 1). Assim, esta base de dados é composta por: data da coleta, hora da coleta, dados das vazões das 3 estações (Francisco Gross, São João e Pacotuba), e a variável alvo que é a vazão da estação Ilha da Luz. O período de coleta inicia no dia 19-02-2021 à zero hora, e vai até 10-02-2023 às 22h. Uma amostra de 5 linhas é apresentada na Tabela 2.

Para a construção da base de dados foram realizados os seguintes passos: (i) descarregamento dos dados históricos de cada estação do *WebService*; (ii) execução de um *script* na linguagem de programação Python que realiza a conversão do arquivo proveniente do *WebService* da extensão *Extensible Markup Language* (XML) para a extensão *Comma-separated values* (CSV); (iii) concatenação dos dados via código usando a biblioteca Pandas. Todos os arquivos e os *scripts* desenvolvidos estão disponíveis no GitHub³.

Na Tabela 3 apresenta-se a quantidade final de dados obtidos para cada base de dados. Como a base de dados “Sensores e Nível” contém a leitura realizada pela Defesa Civil, a quantidade de registros é bem menor que a base composta apenas por dados de sensores.

Tabela 3. Quantidade de amostras de cada base de dados

Base de dados	Quantidade de amostras
Sensores e Nível	608
Somente Sensores	15.800

³https://github.com/duvrDX/dataset_rio

3.2. Métricas

Para avaliarmos o desempenho dos nossos resultados, utilizamos o coeficiente de determinação R^2 , que é a métrica adotada pelo sistema SARI. O R^2 é uma medida estatística que representa a proporção da variabilidade nos dados dependentes (variável alvo) que é explicada pelas variáveis independentes (variáveis preditoras) em um modelo de regressão, conforme Equação (2). Ele varia de 0 a 1, onde 0 indica que o modelo não explica nenhuma variabilidade e 1 indica que o modelo explica toda a variabilidade nos dados, o R^2 é frequentemente usado para avaliar o quão bem um modelo se ajusta aos dados. Um valor elevado de R^2 (próximo a 1) sugere que o modelo é capaz de explicar uma grande parte da variabilidade nos dados, enquanto um valor baixo (próximo a 0) indica que o modelo não é eficaz na explicação da variabilidade.

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (1)$$

3.3. Técnicas de NAS

O trabalho de [Oliveira 2022] teve como objetivo desenvolver um modelo de previsão de consumo de gás por meio de técnicas de aprendizado de máquina, visando superar as previsões dos métodos estatísticos convencionais. Para alcançar esse objetivo, foram empregados métodos de seleção de características, redução de dimensionalidade e exploração de parâmetros e arquiteturas de redes neurais (MLP, RNN e LSTM), utilizando diversas metodologias de AutoML (usando AutoKeras, framework DEAP – Distributed Evolutionary Algorithm e metodologia NEAT – NeuroEvolution of Augmenting Topologies). Os resultados das previsões dos modelos resultantes foram avaliados com base na métrica RMSE (do inglês, Root Mean Squared Error, em português Raiz Quadrada do Erro Quadrático Médio) e superaram os resultados dos métodos estatísticos.

Segundo [Oliveira 2022], o destaque foi o resultado do modelo MLP gerado pelo NAS usando AutoKeras, que apresentou um tempo de predição bem menor, de apenas 0,8s, sendo assim o modelo com maior potencial para uma aplicação numa rotina de utilização no ambiente de produção. [Haykin 2001] cita que o MLP destaca-se como modelo não-linear geral com capacidade de aprender padrões complexos e não apresenta viés indutivo. Para este trabalho, decidimos seguir os resultados do trabalho de [Oliveira 2022] e os experimentos foram realizados com o modelo MLP. O MLP é uma RNA que opera com neurônios organizados em camadas de entrada e camadas escondidas, conectados por sinapses ponderadas. A aprendizagem desse tipo de rede é facilitada pela retropropagação do erro, onde o *feedback* das camadas subsequentes informa e aprimora as camadas anteriores para futuras classificações.

A abordagem de NAS selecionada por este trabalho foi pela implementação de NAS por algoritmo de otimização, o *hill climbing* combinado a um critério de *greedy search*, baseada no trabalho de Elsken e colegas [Elsken et al. 2017] sem o uso de uma biblioteca específica. Os códigos desenvolvidos por [Oliveira 2022] estão disponíveis no GitHub⁴. O código foi desenvolvido utilizando a linguagem de programação Python em conjunto com as bibliotecas Sci-kit Learn, Tensorflow, Keras, Pandas e algumas outras

⁴github.com/Vinicius42-00

Tabela 4. Parâmetros e valores via NAS com AutoKeras.

Parâmetros	Valores
Quantidade de neurônios por camada	[10, 20, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500]
Função de Ativação	['identity', 'logistic', 'relu', 'tanh']
Alpha	[0.001, 0.0001, 0.00001, 0.000001, 0.0000001, 0.00000001]
Otimizador	[SGD, RMSprop, Adam]
Taxa de aprendizado	[0.1, 0.01, 0.001, 0.0001, 0.00001, 0.000001]
Qtde máxima de iterações	[50, 100, 150, 200, 250, 300, 350, 400, 450, 500]

bibliotecas em partes menores do código. São exploradas diversos parâmetros da arquitetura da MLP, com variação de valores apresentados na Tabela 4. Os parâmetros são: a quantidade de camadas escondidas, diferentes funções de ativação, diversos algoritmos otimizadores, taxas de aprendizado, Alpha (é um parâmetro para termo de regularização, também conhecido como termo de penalidade, que combate o *overfitting* restringindo o tamanho dos pesos) e números máximos de iterações.

4. Resultados e Discussão

As duas bases de dados passam por uma divisão de 75% para a base de treino e 25% para a base de teste, essa divisão foi aplicada no trabalho original e mantida nos nossos exemplos [Oliveira 2022]. Nos experimentos definimos o valor do hiper-parâmetro de época igual a 10. Durante o treinamento são mostrados alguns resultados decorrentes da época, ao final são mostradas as informações das métricas que foram obtidas no melhor modelo.

No nosso experimento com a base de dados “Somente Sensores”, o melhor modelo obteve R^2 de 0,947 para a base de dados de treinamento. No nosso experimento com a base de dados “Sensores e Nível”, o melhor modelo obteve um R^2 de 0.840 para a base de dados de treinamento. A Tabela 5 apresenta os hiper-parâmetros encontrados pelo algoritmo NAS para as duas bases de dados.

Tabela 5. Tabela com a arquitetura dos melhores modelos

Hiperparâmetro	Somente Sensores	Sensores e Nível
Quantidade de neurônios por camada	450	450
Função de Ativação	logistic	relu
Alpha	1e-07	1e-04
Otimizador	Adam	Adam
Taxa de Aprendizado	0.1	0.1
Qtde máxima de iterações	400	200

Para avaliar os resultados, o algoritmo executa a melhor arquitetura encontrada pelo algoritmo NAS e faz a predição sobre a base de dados de testes. Para a base de dados “Sensores e Nível”, o resultado foi de R^2 de 0,864. Na Figura 2 é apresentada a curva dos dados reais do nível do rio em metros (com nível máximo de 3 metros) em vermelho e a curva da predição em azul. A melhor arquitetura obtida para a base de dados “Somente Sensores” teve um desempenho de R^2 de 0,955. O gráfico da Figura 3 apresenta a a curva dos dados reais da vazão da estação Ilha da Luz em m^3/s (com vazão máxima próximo de $800 m^3/s$) em vermelho e a curva da predição em azul.

A Tabela 6 apresenta um resumo das métricas, na primeira linha a métrica R^2 do SARI, na segunda linha as métricas do melhor modelo usando os dados da base “Senso-

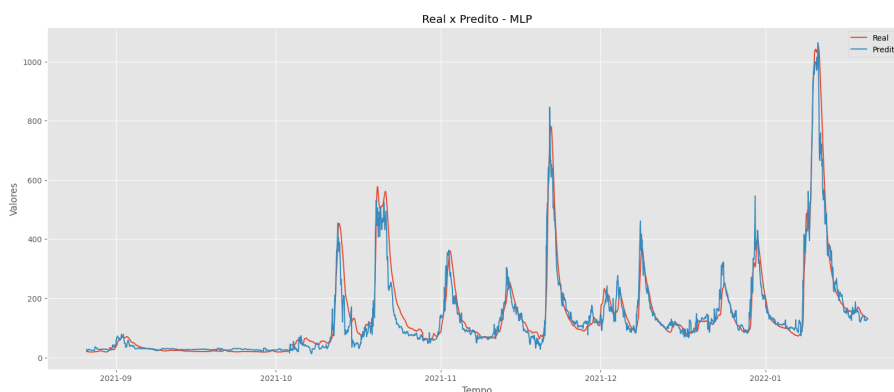


Figura 2. Gráfico comparando a curva de dados de nível do Rio Itapemirim e o valor predito pelo melhor modelo da base de dados “Sensores e Nível”

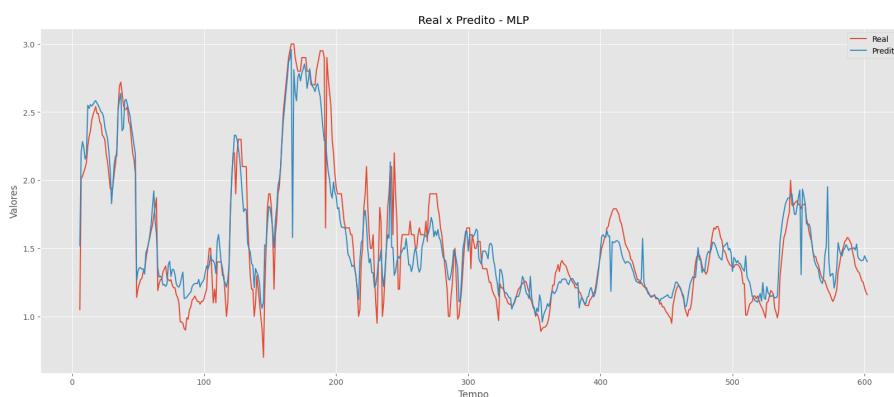


Figura 3. Gráfico comparando a curva de dados de nível do Rio Itapemirim e o valor predito pelo melhor modelo da base de dados “Somente Sensores”

res e Nível” e na terceira linha as métricas do melhor modelo usando os dados da base “Somente Sensores”.

Tabela 6. Tabela com as métricas dos melhores modelos

Modelo	R^2	#amostras	variável alvo	máximo
Nota Técnica - SARI	0,959	608	nível do rio	3m
“Sensores e Nível”	0,864	608	nível do rio	3m
“Somente Sensores”	0,955	15.800	vazão Ilha da Luz	$800m^3/s$

É interessante observar que o modelo MLP gerado para a base de dados “Sensores e Nível” apresentou resultados piores, com menor valor de R^2 , que o modelo matemático. Lembrando que R^2 maior indica que o modelo explica melhor a variabilidade nos dados.

Com a base de dados “Somente Sensores”, o valor de R^2 é próximo ao resultado do modelo matemático. Um pressuposto deste melhor resultado do MLP com relação à outra base de dados é que o modelo de aprendizado de máquina pode ter se saído melhor, por ter uma quantidade maior de dados para o treinamento. A vantagem desta abordagem usando somente sensores é que a base de dados não depende da medição manual feita pela Defesa Civil. Entendemos que ter as duas abordagens funcionando de forma concomitante é interessante para se ter um sistema de alertas duplo, que é mais robusto e

seguro.

5. Conclusão

A proposta do trabalho foi o de comparar os resultados de modelos de MLP com o resultado modelo matemático de regressão usado no SARI para prever o nível ou vazão de água do Rio Itapemirim. Para se gerar o modelo MLP foi usada a técnica NAS implementada por [Oliveira 2022]. Os experimentos foram realizados com duas bases de dados, uma com os dados do SARI em que a variável alvo é o nível do rio e outra com os dados de vazão de 4 estações fluviométricas.

Verificamos que o resultado do modelo matemático de regressão para a previsão do nível do Rio Itapemirim apresentou métrica R^2 melhor, mas podemos considerar que o tamanho da base de dados não era suficiente para o treinamento do MLP. E que o resultado com a base de dados somente de sensores obteve R^2 similar ao modelo matemático, podendo ser usado como um segundo método de acompanhamento do Rio Itapemirim, através de sua vazão. A base de dados somente com sensores possui mais dados e portanto, mais adequado ao uso de métodos de aprendizado de máquina.

Como trabalhos futuros, podemos incluir os dados de outras estações fluviométricas do Rio Itapemirim, bem como incluir informações dos dados de chuva. Além de gerar modelos para outros rios que tenham estações que colem informações similares. Isso nos permitirá explorar diferentes modelos e algoritmos para encontrar a melhor solução possível para prever inundações com antecedência e, assim, contribuir para as agências públicas na mitigação desse problema ambiental crítico.

6. Agradecimentos

Todos os autores agradecem pelo projeto FAPES nº1023/2022 P:2022-8TZV6, EDITAL FAPES Nº 20/2022, sendo que o primeiro autor agradece pela bolsa deste projeto. A professora Karin Komati agradece ao CNPq pela Bolsa de Produtividade DT-2 (308432/2020-7) e pelo projeto 407742/2022-0, também agradece à FAPES pelo Auxílio Taxa de Pesquisa (nº 293/2021).

Referências

- ANA (2023). Agência nacional de Águas e saneamento básico - webservice. <https://telemetriaws1.ana.gov.br/ServiceANA.aspx>. Acesso em: 10 out. 2023.
- Bof, L. (2022). Sari - sistema de alerta do rio itapemirim. <https://servicos.agerh.es.gov.br/sari/>. Acesso em: 10 out. 2023.
- Brandão, J., Correa, F., and Guedes, E. (2023). A comparative analysis of artificial neural networks on river level forecasting for the Rio Madeira Basin. In *Anais do XX Encontro Nacional de Inteligência Artificial e Computacional*, pages 141–155, Porto Alegre, RS, Brasil. SBC.
- De Souza, A. C. and dos Santos, C. T. (2018). Identificação dos problemas de alagamento na cidade de Cachoeiro de Itapemirim-ES: estudo de caso sobre a área central. In *16º Congresso Brasileiro de Geologia de Engenharia e Ambiental*.

- Elsken, T., Metzen, J.-H., and Hutter, F. (2017). Simple and efficient architecture search for convolutional neural networks. *arXiv preprint arXiv:1711.04528*.
- Farias, A. and Mendonça, F. (2022). Riscos socioambientais de inundação urbana sob a perspectiva do sistema ambiental urbano. *Sociedade & Natureza*, 34(1).
- Feurer, M. and Hutter, F. (2019). Hyperparameter optimization. In *Automated machine learning*, pages 3–33. Springer, Cham.
- Haykin, S. S. (2001). *Redes neurais: princípios e prática*.
- Hutter, F., Kotthoff, L., and Vanschoren, J. (2019). *Automated machine learning: methods, systems, challenges*. Springer Nature.
- Lima, P., Marques, F., and Orth, S. (2023). Predição do nível de água utilizando os modelos ARIMA e random forest: Um estudo de caso da barragem eclusa do São Gonçalo. In *Anais do L Seminário Integrado de Software e Hardware*, pages 24–35, Porto Alegre, RS, Brasil. SBC.
- Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., et al. (2019). Evolving deep neural networks. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pages 293–312. Elsevier.
- Morettin, P. A. and Toloi, C. M. (2018). *Análise de séries temporais – Vol. 1*. Editora Blucher, São Paulo, Brasil, 3 edition.
- Morettin, P. A. and Toloi, C. M. (2020). *Análise de séries temporais – Vol. 2*. Editora Blucher.
- Neves, M., Ferreira, R., Da, M., Trigo, S., Serri, M., Oliveira, M., Pessoa, A., Luis, and Mancini, L. (2016). Assinatura isotópica das Águas pluviais e subterrâneas na bacia hidrográfica do Rio Itapemirim, estado do Espírito Santo. In *XIX Congresso Brasileiro de Águas Subterrâneas*.
- Oliveira, V. M. (2022). *Previsão do consumo de gás no processo de pelletização através de séries temporais multivariadas*. PhD thesis, Mestrado em Computação Aplicada, Instituto Federal do Espírito Santo (IFES).
- Oliveira, V. M., Komati, K. S., and Andrade, J. O. (2021). Seleção de características de séries temporais multivariadas do consumo de gás na pelletização de minério de ferro. In *Anais do XXVIII SIMPEP (Simpósio Brasileiro de Engenharia de Produção 2021)*. UNESP.
- Saito, S. M., Dias, M. C. d. A., Alvalá, R. C. d. S., Stenner, C., Franco, C. d. O., Ribeiro, J. V. M., Souza, P. A. d., and Santana, R. A. S. d. M. (2023). População urbana exposta aos riscos de deslizamentos, inundações e enxurradas no brasil. *Sociedade & Natureza*, 31:e46320.
- Telikani, A., Tahmassebi, A., Banzhaf, W., and Gandomi, A. H. (2021). Evolutionary machine learning: A survey. *ACM Comput. Surv.*, 54(8).
- Zoph, B. and Le, Q. V. (2016). Neural architecture search with reinforcement learning. *arXiv preprint arXiv:1611.01578*.