



Mudanças climáticas: inteligência artificial na previsão de eventos hidrológicos extremos no estado do Amazonas

Gilberto Massulo Neto¹, Emerson Leão B. do Nascimento².

¹Venturus – Centro de Inovação Tecnológica, Rua Pará, 500, Nossa Senhora das Graças, Manaus, AM, Brasil.

²Fundação, Centro de Análise Pesquisa e Inovação Tecnológica – FUCAPI, Av. Gov. Danilo de Matos Areosa, 381 – Distrito Industrial I, Manaus-AM.

gilbertomassulo@gmail.com

eng.emersonleaoobrito@gmail.com

Palavras-Chave: Inteligência Artificial; Mudanças Climáticas; Previsão Hidrológica; Amazonas.

Introdução

As mudanças climáticas têm provocado alterações significativas no regime hidrológico da Amazônia, intensificando a ocorrência de eventos extremos, como secas e cheias severas. Esses fenômenos impactam diretamente a população, a economia regional e os ecossistemas. Nesse contexto, torna-se essencial desenvolver mecanismos preditivos capazes de antecipar tais eventos, fornecendo subsídios para a tomada de decisão em políticas públicas e estratégias de mitigação. A aplicação de técnicas de inteligência artificial (IA) representa um avanço relevante para a previsão hidrológica, pois permite analisar grandes volumes de dados climáticos e identificar padrões complexos. O uso desses métodos possibilita maior precisão na estimativa de cenários futuros, contribuindo para a redução de riscos e para o planejamento sustentável no estado do Amazonas.

Material e Métodos

O estudo utilizou séries históricas de dados hidrológicos e climáticos do estado do Amazonas, contemplando variáveis como nível dos rios, precipitação e temperatura. Os dados foram obtidos em bases públicas de monitoramento ambiental e tratados por meio de técnicas de normalização e análise estatística. Para a modelagem preditiva, foram aplicados algoritmos de inteligência artificial, com destaque para redes neurais recorrentes do tipo Long Short-Term Memory (LSTM), adequadas para séries temporais. O processo envolveu três etapas principais: (i) preparação e limpeza dos dados, (ii) treinamento e validação dos modelos em diferentes cenários de simulação, e (iii) avaliação do desempenho com métricas de erro, como RMSE e MAE. As análises foram realizadas em ambiente computacional com suporte a bibliotecas específicas de aprendizado de máquina em Python, o que possibilitou maior precisão e agilidade no processamento dos dados.

Resultados e Discussão

Para avaliar o desempenho dos três modelos de aprendizado de máquina — Random Forest, Gradient Boosting e Long Short-Term Memory (LSTM) — foram realizados experimentos com diferentes configurações de hiperparâmetros. O objetivo foi identificar a configuração mais eficiente em termos de Erro Quadrático Médio (RMSE), métrica utilizada para mensurar a precisão das previsões. De

modo geral, observou-se que, embora todos os modelos tenham conseguido capturar tendências hidrológicas, houve variação significativa nos níveis de acurácia. O LSTM apresentou desempenho superior, destacando-se em relação aos algoritmos baseados em árvores. O modelo Random Forest apresentou desempenho razoável, com melhor resultado no experimento RFOREST-1, atingindo RMSE = 72.28. A maior profundidade (Max Depth = 120) associada ao critério sqrt para seleção de variáveis mostrou-se adequada para capturar padrões mais complexos. Em contrapartida, configurações inadequadas, como no experimento RFOREST-5, elevaram o erro (RMSE = 83.79), evidenciando a sensibilidade do modelo aos hiperparâmetros escolhidos (Tabela 3).

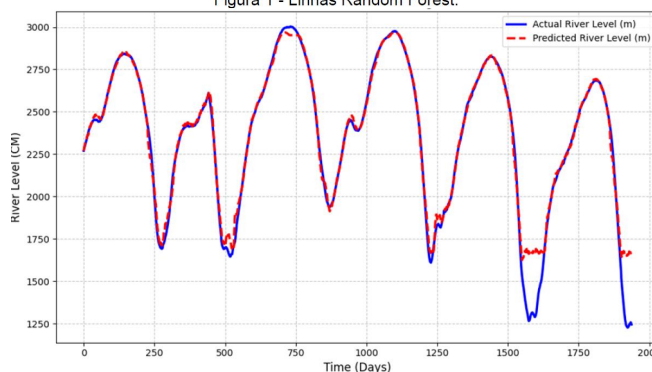
Tabela 3. Resultados - Random Forest.

Experiment	RMSE	Window Size	Max Depth	Max Features	N Estimators
RFOREST-1	72.2828	30	120	sqrt	100
RFOREST-2	81.8818	30	100	sqrt	100
RFOREST-3	81.8820	30	50	sqrt	100
RFOREST-4	82.0530	30	10	sqrt	100
RFOREST-5	83.7919	30	150	log2	100

Fonte: Autoria própria, 2024.

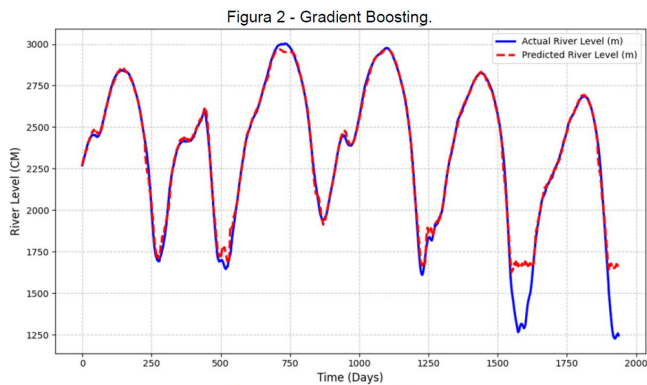
A Figura 1 ilustra a comparação entre valores reais e previstos pelo Random Forest. Nota-se que o modelo consegue acompanhar tendências históricas, mas apresenta desvios consideráveis nos anos mais recentes, especialmente em 2023 e 2024, quando não conseguiu capturar adequadamente a queda dos níveis do rio.

Figura 1 - Linhas Random Forest.



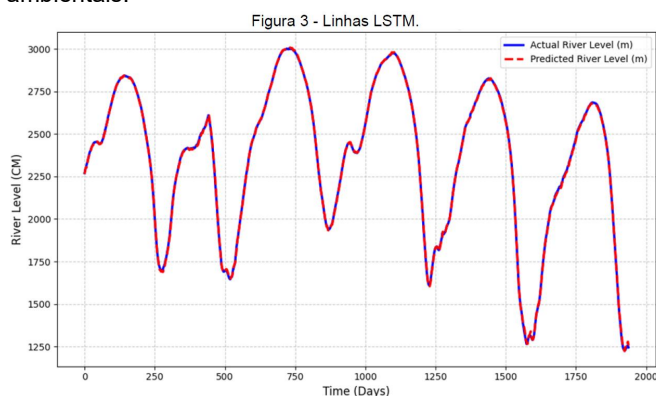
Fonte: Autoria própria, 2024.

O Gradient Boosting alcançou desempenho semelhante ao Random Forest, com melhor configuração no experimento GBOOSTING-1, obtendo RMSE = 78.25 (Tabela 4). O ajuste fino dos hiperparâmetros foi determinante: profundidade de 100, taxa de aprendizado de 0.1 e subsample de 0.3 permitiram melhor generalização.



Fonte: Autoria própria, 2024.

A Figura 2 mostra que o modelo conseguiu acompanhar de maneira consistente as oscilações nos níveis do rio, capturando picos e estiagens ao longo da série histórica. Entretanto, observa-se desvio crescente nos últimos anos, indicando limitação do algoritmo em lidar com padrões mais recentes ou mudanças abruptas associadas a eventos extremos. Comparativamente, o Gradient Boosting apresentou erro ligeiramente maior que o Random Forest, reforçando que, embora eficiente, sua performance depende de ajustes criteriosos e pode ser instável diante de séries temporais longas. O modelo LSTM obteve os melhores resultados gerais, com destaque para o experimento LSTM-1, que alcançou RMSE = 4.43 (Tabela 5). Essa diferença é expressiva quando comparada ao Random Forest e Gradient Boosting, que apresentaram erros acima de 70. Isso demonstra a superioridade do LSTM na captura de padrões temporais e dependências sequenciais em séries históricas ambientais.



Fonte: Autoria própria, 2024.

A Figura 3 evidencia a proximidade entre valores reais e previstos, mostrando que o LSTM acompanha com precisão tanto os ciclos de cheias quanto os de secas. Diferentemente dos modelos de árvores, o LSTM manteve robustez mesmo diante de variações nos hiperparâmetros, o que reforça sua capacidade de generalização e estabilidade no aprendizado.

Agradecimentos

Primeiramente à Deus, pois sem ele nada seria possível. Minha família e amigos e ao Prof. Emerson, por todo apoio.

Referências

- [1] Aff, I. R.; Nawirah, N. Design of financial statement preparation based on SAK EMKM using Microsoft Excel. *Journal of Islamic Accounting and Finance Research*, v.2, n.1, p.141-152, 2020.
- [2] Ageng, D.; Huang, C. Y.; Cheng, R. G. A short-term household load forecasting framework using LSTM and data preparation. *IEEE Access*, v.9, p.167911-167919, 2021.
- [3] Agência Nacional de Águas – ANA. Dados sobre recursos hídricos. Disponível em: <https://www.gov.br/pt-br/orgaos/agencia-nacional-de-aguas>. Acesso em: 23 nov. 2024.

[4] Araújo, D. N. et al. Impactos da sujidade e efeitos da limpeza no desempenho de módulos fotovoltaicos. In: *Anais Congresso Brasileiro de Energia Solar – CBENS*, 2020.

[5] Artaxo, P. As três emergências que nossa sociedade enfrenta: saúde, biodiversidade e mudanças climáticas. *Estudos Avançados*, v.34, p.53-66, 2020.

[6] CEMADEN. Centro Nacional de Monitoramento e Alertas de Desastres Naturais. Disponível em: <https://www.cemaden.gov.br>. Acesso em: 23 nov. 2024.

[7] Chen, W. et al. Deep learning reservoir porosity prediction based on multilayer long short-term memory network. *Geophysics*, v.85, n.4, p.WA213–WA225, 2020.

[8] Das, D.; Semaan, B. Quoras: A Python API for Quora Data Collection to Increase Multi-Language Social Science Research. *CSCW Companion Publication*, 2020.

[9] Ferreira, B. C. C.; Valverde, M. C. Análise dos índices extremos de precipitação em cenários futuros na bacia do rio Ribeira de Iguape – São Paulo. *Revista Brasileira de Meteorologia*, v.37, p.41-55, 2022.

[10] Gadde, H. AI-Enhanced Data Warehousing: Optimizing ETL Processes for Real-Time Analytics. *Revista de Inteligência Artificial em Medicina*, v.11, n.1, p.300-327, 2020.

[11] GOV.BR. Seca histórica na Amazônia em 2023 foi 30 vezes mais provável devido à mudança do clima. Disponível em: <https://www.gov.br/mcti>. Acesso em: 23 nov. 2024.

[12] Gunawan, A. E. K.; Wibowo, A. Stock Price Movement Classification Using Ensembled Model of LSTM and Random Forest. *JOIV: International Journal on Informatics Visualization*, v.7, n.4, p.2255-2262, 2023.

[13] INMET. Instituto Nacional de Meteorologia. Disponível em: <https://portal.inmet.gov.br>. Acesso em: 23 nov. 2024.

[14] Izza, Y.; Ignatiev, A.; Marques-Silva, J. On tackling explanation redundancy in decision trees. *Journal of Artificial Intelligence Research*, v.75, p.261-321, 2022.

[15] Kotsiantis, S.; Konstaninidis, K.; Tsougrinis, T. Data mining techniques in the prediction of academic performance. *Artificial Intelligence Review*, v.30, n.4, p.233-252, 2006.

[16] Laudon, K. C.; Laudon, J. P. *Management Information Systems: Managing the Digital Firm*. 15.ed. Upper Saddle River: Pearson Education, 2020.

[17] López, X. et al. A short-term household load forecasting framework using LSTM and data preparation. *IEEE Access*, v.9, p.167911-167919, 2021.

[18] Oracle. RMSE: Root Mean Square Error. Disponível em: <https://docs.oracle.com/cloud>. Acesso em: 30 nov. 2024.

[19] Oruh, J.; Viriri, S.; Adegun, A. Long short-term memory recurrent neural network for automatic speech recognition. *IEEE Access*, v.10, p.30069-30079, 2022.

[20] Pölsterl, S. scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn. *Journal of Machine Learning Research*, v.21, n.212, p.1-6, 2020.

[21] Punia, S. et al. Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. *International Journal of Production Research*, v.58, n.16, p.4964-4979, 2020.

[22] Qiao, M. Y. et al. Bearing fault diagnosis based on natural adaptive moment estimation algorithm and improved octave convolution. *IEEE Access*, v.8, p.196790-196803, 2020.

[23] Schonlau, M.; Zou, R. Y. The random forest algorithm for statistical learning. *The Stata Journal*, v.20, n.1, p.3-29, 2020.

[24] Szega, M. Methodology of advanced data validation and reconciliation application in industrial thermal processes. *Energy*, v.198, p.117326, 2020.

[25] Xie, J. et al. Estimating gridded monthly baseflow from 1981 to 2020 for the contiguous US using LSTM networks. *Water Resources Research*, v.58, n.8, p.e2021WR031663, 2022.