

#### INTRODUÇÃO:

Este trabalho propõe uma contribuição significativa para o estudo das Séries Temporais Multivariadas Irregulares (IMTS), com foco no monitoramento do Porto de Paranaguá, maior porto graneleiro da América Latina. O conjunto de dados *Paranaguá Port Meteorological and Oceanographic Dataset* (P<sup>2</sup>MOD) preenche uma lacuna na literatura ao oferecer dados reais que refletem as irregularidades comuns em sensoriamento ambiental, permitindo o desenvolvimento de modelos mais robustos para tarefa de *IMTS-forecasting* usando AI/ML.

#### OBJETIVOS:

Os objetivos deste trabalho são, primeiramente, disponibilizar o *Paranaguá Port Meteorological and Oceanographic Dataset* (P<sup>2</sup>MOD), um conjunto de dados que reflete as irregularidades típicas das séries temporais reais. Em seguida, será realizada uma análise exploratória (EDA) para identificar as principais irregularidades presentes nos dados e, a partir disso, demonstrar a aplicabilidade do *dataset* em cenários práticos, utilizando dois modelos de previsão, um básico e um com técnicas específicas para lidar com irregularidades. Por fim, o estudo visa contribuir para a pesquisa em IMTS, oferecendo um recurso valioso para o desenvolvimento de métodos mais eficazes no tratamento dessas irregularidades.

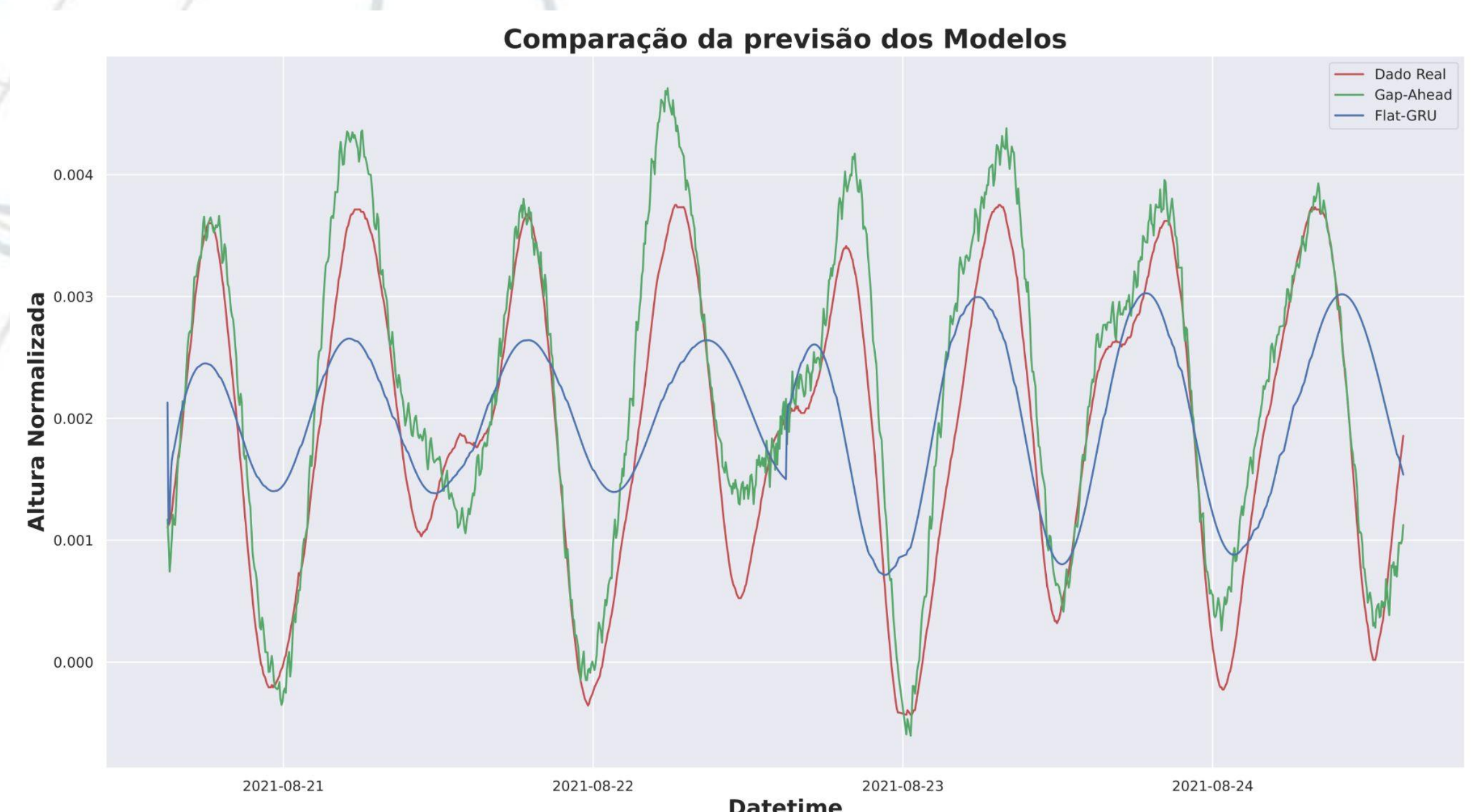
#### METODOLOGIA:

A metodologia deste trabalho envolve a coleta, preparação e análise de dados do Porto de Paranaguá, com foco em identificar as irregularidades típicas das Séries Temporais Multivariadas Irregulares (IMTS). Inicialmente, o *dataset* P<sup>2</sup>MOD será estruturado, documentado e disponibilizado para a comunidade acadêmica. A seguir, será realizada uma análise exploratória para caracterizar as irregularidades nos dados. Para validar o uso do *dataset*, serão aplicados dois modelos de previsão, tanto uma GRU-standard quanto um modelo de Gap-Ahead + Time Encoding, a fim de testar a eficácia e os desafios de lidar com dados irregulares para a tarefa de *IMTS-Forecasting*.

**Integrantes:** - Enzo Bustos Da Silva  
**Orientador(a):** - Prof. Dra. Anna Reali Costa  
**Co-orientador(a):** - Mestre Marcel de Barros

#### CONCLUSÕES:

Os experimentos realizados demonstraram que abordar diretamente as irregularidades inerentes às séries temporais multivariadas, sem recorrer a regularizações, é significativamente mais eficaz para a tarefa de previsão em dados ambientais reais, como os do *dataset* P<sup>2</sup>MOD. O modelo Gap-Ahead com Time Encoding superou a GRU-standard na métrica analisada, destacando sua capacidade de capturar as dinâmicas complexas dos dados irregulares e reafirmando a hipótese de que tratar as irregularidades como uma característica intrínseca, e não como algo a ser regularizado, resulta em maior precisão preditiva e robustez. Além disso, o P<sup>2</sup>MOD se mostrou um benchmark valioso e desafiador para a pesquisa em *IMTS-forecasting*, fornecendo um cenário realista que fomenta o desenvolvimento de modelos avançados e reforça a relevância da preservação da integridade dos dados em análises futuras.



#### REFERÊNCIAS:

- [1] B. M. Marlin, "A Survey on Principles, Models and Methods for Learning from Irregularly Sampled Time Series", *arXiv preprint arXiv:2012.00168*, 2020, doi: [10.48550/arxiv.2012.00168](https://doi.org/10.48550/arxiv.2012.00168).
- [2] S. N. Shukla, "Deep Learning Models for Irregularly Sampled and Incomplete Time Series", Article, University of Massachusetts Amherst, 2021. doi: [10.7275/24475265](https://doi.org/10.7275/24475265).
- [3] M. Barros *et al.*, "Early Detection of Extreme Storm Tide Events Using Multimodal Data Processing", *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, nº 20, p. 21923–21931, mar. 2024, doi: [10.1609/aaai.v38i20.30194](https://doi.org/10.1609/aaai.v38i20.30194).