

# ALEX SOARES PRESTES

Engenheiro de Machine Learning | Computação Científica | HPC

Bauru, SP | [alex.prestes@outlook.com.br](mailto:alex.prestes@outlook.com.br) | (14) 99728-8656

[linkedin.com/in/alexsprestes](https://linkedin.com/in/alexsprestes) | [github.com/AlexPrestes](https://github.com/AlexPrestes)

## Resumo Profissional

Engenheiro de Machine Learning com trajetória técnica construída a partir de formação científica e atuação contínua em projetos acadêmicos e aplicados. Especialista em engenharia de pipelines de dados, otimização de performance em Python e computação científica, com foco em problemas de alta complexidade e escala. Experiência prática em construção de pipelines de Machine Learning, validação estatística, modelagem em grafos e integração com sistemas corporativos (SAP Business One). Bacharel em Física (USP), com base sólida em álgebra linear, processamento eficiente e métodos numéricos. Atualmente cursando MBA em Data Science & Analytics (USP/ESALQ), atuo no desenvolvimento de soluções ponta a ponta — da concepção metodológica à implementação, validação e documentação reproduzível.

## Experiência Profissional

### Engenheiro de Machine Learning — Consultoria Técnica (Nov/2025)

Remoto, Brasil

- Atuação na otimização de pipeline do mestrado de **link prediction em grafos biológicos (PPI)**.
- Modelagem do pipeline (ingestão, amostragem, treino e validação), obtendo ganho de **163x em performance e 329x em escalabilidade**.
- Otimização vetorial e uso de **álgebra linear esparsa** para cálculo de features estruturais em **34 milhões de pares de nós**, com tempo total de execução de **4,4 segundos**, respeitando restrições de hardware (notebook i5, 10GB RAM).
- Desenvolvimento de **K-Fold customizado para grafos**, garantindo ausência de data leakage e avaliação metodologicamente correta.
- Implementação de validação estatística (teste t pareado), documentação integral do pipeline e garantia de **reprodutibilidade e auditabilidade científica**.

### Analista de Dados – Gemmini (Dez/2024 – Fev/2025)

- Atuação no suporte técnico à transição para o **SAP Business One**, garantindo consistência e integridade dos dados logísticos.
- Desenvolvimento de automações em **Python e Excel** para relatórios operacionais e indicadores de estoque.
- Suprimento operacional à geração de relatórios e acompanhamento de dados logísticos.

### Analista de Dados – Completa Atacadista (Set/2010 – Nov/2013)

- Liderança técnica na migração de processos manuais em Excel para o **ERP SAP Business One**.
- Automação de relatórios com **SQL e VBA**, reduzindo em até 70% o tempo de execução de rotinas.
- Criação de **dashboards estratégicos** de vendas e estoque para suporte à tomada de decisão.
- Desenvolvimento de integrações de sistemas usando **Python e Shell Script**, aumentando a confiabilidade dos dados corporativos.

## Formação Acadêmica

### MBA em Data Science & Analytics — USP/ESALQ

(Mai/2025 – Dez/2026) (em andamento)

### Bacharelado em Física — USP/IFSC

(Jan/2017 – Jul/2023)

TCC: Análise de séries temporais financeiras por meio de **grafos de visibilidade**, com foco em extração de métricas topológicas e modelagem supervisionada.

## Atividades Relevantes

### Análise de Séries Temporais Financeiras via Grafos de Visibilidade e Machine Learning (TCC)

USP – Instituto de Física de São Carlos

- Desenvolvimento de um pipeline completo para modelagem de séries temporais financeiras utilizando **Visibility Graphs**.
- Extração automatizada de métricas topológicas (grau, densidade, clustering, conectividade) para composição de vetores de features.
- Preparação de dataset tabular e treinamento de modelos supervisionados (**Random Forest Regressor**), com validação cruzada e análise de importância de variáveis.
- Investigação da relação entre propriedades estruturais dos grafos e o comportamento real dos fundos, gerando insights interpretáveis a partir de métricas topológicas.

### Kenshi Translator Toolkit — Biblioteca Python publicada no PyPI (2025)

Projeto independente

- Engenharia reversa completa do formato binário **.mod** do jogo Kenshi, com desenvolvimento de parser e encoder para leitura, manipulação e reconstrução segura dos arquivos.
- Implementação de arquitetura modular com separação entre domínio e infraestrutura, empacotamento e publicação no **PyPI**.
- Biblioteca ativa com usuários externos, distribuída publicamente via pip.

### GEPAC — Grupo de Estudos em Programação Científica (2019 – 2021)

USP – Instituto de Física de São Carlos

- Co-fundador e organizador de grupo técnico voltado à computação científica e programação de alto desempenho.
- Elaboração e apresentação de palestras técnicas com material autoral, incluindo: [Computação Científica com Julia](#) e [Simulação Numérica de Oscilações em Membranas](#).
- Produção de materiais públicos e projetos educacionais em computação científica, com código e conteúdo publicados em [gepac.github.io](https://gepac.github.io).
- Instrutor de Python científico e contribuidor do [Grupy-Sanca](#).

### Atuação Técnica Colaborativa em Projetos Acadêmicos (2017 – 2024)

USP – IFSC e colaborações informais

- Apoio técnico recorrente a projetos acadêmicos em **Machine Learning, computação científica e visão computacional**.
- Implementação e ajuste de código para TCCs e projetos de pesquisa, incluindo análise de erros, experimentação computacional e validação de resultados.
- Colaborações técnicas em contextos de Física Computacional, biofotônica e modelagem em grafos, sem vínculo formal ou autoria científica.

## Competências Técnicas

- **Engenharia de Dados:** pipelines ETL/ELT, integração com sistemas transacionais, versionamento e auditoria de dados.
- **Python & Performance:** NumPy, SciPy, álgebra linear esparsa, profiling e otimização vetorial.
- **Machine Learning Aplicado:** validação cruzada, métricas, feature engineering, modelagem em grafos.
- **Computação Científica:** simulação numérica, métodos matriciais, análise de desempenho.
- **Ferramentas:** Git, Docker, GitHub Actions, SQL, SAP B1, Power BI.