

Bárbara Fernandes Madera - 11915032

Relatório de Estágio Obrigatório

Orientador: Prof. Dr. Maxmilliam Luppe
Coorientador: Dra. Renata de Paris

São Carlos
2025

Resumo

Este relatório apresenta as atividades desenvolvidas durante o estágio obrigatório realizado no Instituto de Pesquisa Eldorado, sede de Brasília, no período de 6 de janeiro a 7 de março de 2025. O objetivo principal foi aprimorar conhecimentos em análise e modelagem de séries temporais utilizando a linguagem Python, aplicando tais técnicas a dados meteorológicos da cidade de Brasília (INMET). O estágio também envolveu a utilização de ferramentas corporativas para versionamento e integração de resultados. O relatório contempla introdução, apresentação da empresa, objetivos, cronograma, atividades desenvolvidas, metodologia de trabalho, análise crítica e conclusão, conforme as normas da EESC/USP. **Palavras-chave:** estágio obrigatório; séries temporais; Python; INMET; modelagem.

Resumo

This report describes the activities developed during the mandatory internship at the Eldorado Research Institute, Brasília unit, from January 6 to March 7, 2025. The main objective was to improve knowledge in time series analysis and modeling using Python, applying these techniques to meteorological data from Brasília (INMET). The internship also involved the use of corporate tools for version control and results integration. The report includes introduction, company presentation, objectives, schedule, developed activities, work methodology, critical analysis, and conclusion, according to EESC/USP guidelines. **Keywords:** mandatory internship; time series; Python; INMET; modeling.

Sumário

1	INTRODUÇÃO	5
2	APRESENTAÇÃO DA EMPRESA	6
3	OBJETIVOS	7
4	CRONOGRAMA	8
5	METODOLOGIA DE TRABALHO	9
6	ATIVIDADES DESENVOLVIDAS	10
6.1	Análise e Modelagem de Séries Temporais	10
6.2	Integração com Ferramentas Corporativas	10
6.3	Relacionamento com a Empresa e Aspectos Administrativos	10
6.4	Disciplinas Relacionadas	11
6.5	Aspectos Éticos e Profissionais	11
7	IMPLEMENTAÇÃO DE ALGORITMOS	12
8	ANÁLISE CRÍTICA	13
8.1	Comprovante de Estágio	13
9	CONCLUSÃO	15
10	REFERÊNCIAS	16

1 Introdução

Este relatório tem como objetivo apresentar as atividades realizadas durante o estágio obrigatório supervisionado no Instituto de Pesquisa Eldorado, localizado na sede de Brasília, no período de 06 de janeiro a 07 de março de 2025. O estágio foi realizado no contexto do curso de Engenharia de Computação da Escola de Engenharia de São Carlos (EESC) da Universidade de São Paulo (USP).

O estágio foi importante para a minha formação profissional, dado que me permitiu aplicar conhecimentos adquiridos ao longo da graduação, especificamente nas áreas de modelagem de séries temporais e análise de dados utilizando Python. Além disso, o modelo híbrido de expediente me permitiu ter uma vivência de trabalhar de home-office 3 vezes na semana e também poder vivenciar o ambiente corporativo durante 2 vezes na semana. Durante o estágio, tive a oportunidade de trabalhar com dados reais do Instituto Nacional de Meteorologia (INMET), envolvendo técnicas avançadas de previsão de séries temporais, como ARIMA, SARIMA, N-BEATS e XGBoost, e o desenvolvimento de uma feature específica para tratamento de dados de temperatura, considerando casos de outliers.

Os principais objetivos deste estágio foram:

- Desenvolver habilidades em análise e modelagem de séries temporais.
- Aplicar técnicas de previsão para dados meteorológicos utilizando Python.
- Trabalhar com ferramentas corporativas para versionamento de código, criação de dashboards e integração de resultados.
- Implementar modelos preditivos como N-BEATS e XGBoost e contribuir com o estudo de ARIMA e SARIMA, desenvolvidos pela minha dupla de estágio.
- Melhorar a qualidade das previsões por meio do tratamento de outliers nos dados de temperatura e ajuste de hiperparâmetros.

Este estágio também proporcionou uma visão prática do trabalho em um ambiente corporativo de pesquisa, envolvendo não apenas aspectos técnicos de modelagem, mas também desafios relacionados ao gerenciamento de projetos e comunicação de resultados em um ambiente colaborativo.

2 Apresentação da Empresa

O Instituto de Pesquisa Eldorado é uma instituição privada de pesquisa, desenvolvimento e inovação com forte presença no mercado brasileiro. Fundado com a missão de fornecer soluções tecnológicas de ponta, o instituto se destaca pela sua atuação em diversas áreas, incluindo engenharia, ciência de dados, inteligência artificial e tecnologias emergentes. Seu portfólio de atuação inclui desde a pesquisa básica até o desenvolvimento de soluções aplicadas, com foco em inovação e excelência técnica.

A sede do Instituto de Pesquisa Eldorado em Brasília, por sua vez, é especializada em projetos voltados para desenvolver soluções tecnológicas de software para grandes empresas tais como Tim, Lenovo, por exemplo. Para isso, a empresa possui uma infraestrutura significativa, com equipes multidisciplinares que trabalham de forma colaborativa para desenvolver soluções inteligentes para diferentes empresas sejam elas nacionais ou internacionais.

O setor no qual desenvolvi o programa de estágio foi o de Ciência de Dados e Análise de Dados Meteorológicos, um departamento dedicado ao processamento e análise de séries temporais para previsão do clima e estudos ambientais. Esse setor é responsável por lidar com grandes bases de dados ambientais, como os fornecidos pelo Instituto Nacional de Meteorologia (INMET), e aplicar modelos estatísticos e de aprendizado de máquina para gerar previsões e insights que podem ser aplicados para setores múltiplos, são eles: agro-negócio, auxiliando no planejamento de atividades agrícolas; gestão de recursos hídricos, otimização de rotas no transporte e logística; segurança e defesa civil, saúde pública, turismo e eventos ao ar livre, além de otimizar o uso de energias renováveis como solar e eólica. Ao permitir a tomada de decisões informadas e reduzir riscos, a previsão do tempo desempenha um papel na melhoria da eficiência e segurança em múltiplos setores da sociedade. Neste sentido, durante o estágio, tive a oportunidade de trabalhar diretamente com essas ferramentas, utilizando técnicas avançadas de modelagem preditiva, como ARIMA, SARIMA, XGBoost e N-BEATS, sempre com foco em melhorar a precisão das previsões climáticas e a qualidade dos dados utilizados.

3 Objetivos

- Desenvolver habilidades em análise e modelagem de séries temporais utilizando Python.
- Aplicar modelos como ARIMA, SARIMA, XGBoost e N-BEATS a dados meteorológicos do INMET.
- Integrar as ferramentas corporativas para versionamento de código, visualização de resultados e documentação.
- Produzir dashboards interativos e relatórios técnicos para suporte a decisões.
- Aplicar técnicas de Engenharia de Software durante o projeto para gerir o projeto.

4 Cronograma

Período	Atividades Desenvolvidas
06/01 a 20/01	Coleta, análise exploratória e limpeza dos dados meteorológicos do INMET.
21/01 a 10/02	Estudo de modelos ARIMA e SARIMA.
11/02 a 28/02	Desenvolvimento de modelos avançados (XGBoost e N-BEATS).
01/03 a 05/03	Integração dos resultados, elaboração de relatórios e dashboards interativos.
06/03 a 07/03	Apresentação de resultados e entrega de materiais fornecidos pela empresa.

5 Metodologia de Trabalho

Foi adotado um ciclo ágil de desenvolvimento, com as seguintes etapas:

1. Coleta e pré-processamento dos dados;
2. Análise exploratória e identificação de padrões;
3. Modelagem preditiva com diferentes algoritmos;
4. Avaliação de desempenho com métricas estatísticas;
5. Documentação e apresentação dos resultados.

Foram utilizados repositórios GitHub para controle de versão e integração com as ferramentas corporativas para dashboards e relatórios assim como foram fornecidos cursos de Udemey para adquirir conhecimento para aplicação correta de Análise Exploratória em Python, foram eles:

- Master Time Series Analysis and Forecasting with Python 2025. Disponível em: <https://github.com/barbarafernandesmadera/Master-Time-Series-Analysis-and-Forecasting-with-Python-2025>. Acesso em: 06 jan. 2025.
- Python for Data Science and Machine Learning Bootcamp. Disponível em: <https://github.com/barbarafernandesmadera/Python-for-Data-Science-and-Machine-Learning-Bootcamp>. Acesso em: 06 jan. 2025.

6 Atividades Desenvolvidas

Durante o estágio no Instituto de Pesquisa Eldorado, as atividades foram voltadas para o desenvolvimento de soluções em análise de séries temporais e modelagem preditiva utilizando Python. A primeira etapa do estágio envolveu a coleta e pré-processamento de dados meteorológicos fornecidos pelo Instituto Nacional de Meteorologia (INMET). Foram aplicadas técnicas de limpeza de dados e normalização para garantir a qualidade dos dados antes da aplicação dos modelos de previsão. Além disso, foi analisado o uso ou não de outliers durante a implementação dos códigos, com isso, durante o projeto há implementações para ambos os casos.

6.1 Análise e Modelagem de Séries Temporais

A principal atividade técnica desenvolvida foi a aplicação de modelos de previsão de séries temporais, como ARIMA, SARIMA, XgBoost e Nbeats com a colaboração da minha dupla de estágio. Estes modelos foram escolhidos devido à sua eficácia na captura de padrões sazonais e tendências em séries temporais meteorológicas. Após essa etapa inicial, implementei dois modelos mais avançados: XGBoost e N-BEATS, para previsão de variáveis como a temperatura. A feature de tratamento de dados de temperatura, com foco na detecção e tratamento de outliers, foi desenvolvida com o intuito de melhorar a precisão das previsões de 4 semanas à frente. Esse processo impacta no resultado final dos resultados, pois a qualidade dos dados impacta diretamente na acurácia.

6.2 Integração com Ferramentas Corporativas

Uma parte significativa do estágio envolveu o uso de ferramentas corporativas, como GitHub para versionamento de código, Bitbucket para gerenciamento de repositórios, Microsoft Teams para comunicação e colaboração em equipe, além de outras ferramentas relacionadas ao ambiente profissional. Essa integração com as ferramentas corporativas garantiu a organização e rastreabilidade dos resultados, além de facilitar a colaboração com a equipe de pesquisa.

6.3 Relacionamento com a Empresa e Aspectos Administrativos

Além das atividades técnicas, pude observar e aprender com a dinâmica organizacional da empresa. A interação com os membros da equipe de Software colaborou para o aprendizado sobre a aplicação prática de modelos preditivos. A empresa adota uma filosofia

colaborativa e ágil, onde a comunicação entre os diferentes níveis hierárquicos é clara e eficiente. Essa interação contribuiu para o desenvolvimento de habilidades interpessoais importantes, como colaboração em equipe, gestão de tempo e comunicação técnica.

Ademais, participei de treinamentos internos que abordaram boas práticas no uso de ferramentas de versionamento de código, além de técnicas avançadas de modelagem de dados e análise de séries temporais. Essas oportunidades foram importantes para o meu aprendizado e aperfeiçoamento profissional dado que eu não possuía uma experiência prévia formal na área.

6.4 Disciplinas Relacionadas

As atividades desenvolvidas durante o estágio estão diretamente relacionadas com disciplinas cursadas na graduação, como:

- Modelagem Matemática e Estatística, que forneceu os fundamentos teóricos para a aplicação de modelos como ARIMA e SARIMA;
- Ciência de Dados e Inteligência Artificial, que foi essencial para o desenvolvimento e implementação de modelos avançados como XGBoost e N-BEATS;
- Análise de Dados, que foi aplicada para o tratamento e pré-processamento dos dados meteorológicos.
- Engenharia de Software, que foi aplicada para o uso de métodos ágeis e gestão de projetos.

Essas disciplinas me permitiram entender as bases matemáticas e computacionais necessárias para realizar as previsões de séries temporais com precisão e confiabilidade.

6.5 Aspectos Éticos e Profissionais

O estágio também proporcionou uma oportunidade para refletir sobre os aspectos éticos e profissionais envolvidos na análise de dados. A empresa adota uma postura ética rigorosa quanto ao uso de dados sensíveis e garante que as informações utilizadas nas análises sejam tratadas de forma responsável e transparente. Além disso, o estágio me permitiu entender melhor a importância de se trabalhar com dados de forma ética, garantindo que os resultados obtidos possam ser aplicados de maneira justa e equitativa.

7 Implementação de Algoritmos

As atividades foram realizadas com base em Python, utilizando bibliotecas como **pandas**, **numpy**, **matplotlib**, **scikit-learn**, **xgboost** e **pytorch-forecasting**. Foram criados notebooks específicos para a implementação de modelos:

- **Notebook XGBoost:** `serie_temporal4_6.ComOut.XgBoots.ipynb`.
Implementou-se um modelo de séries temporais com XGBoost, incluindo tratamento de hiperparâmetros, divisão treino/teste e análise de métricas (RMSE, MAE).
- **Notebook N-BEATS:** `serie_temporal4_6.ComOut.NBeats.ipynb`.
Implementou-se a arquitetura de redes neurais N-BEATS para previsão de séries temporais. Incluiu preparação dos dados, normalização e visualização dos resultados. Sendo este um algoritmo mais lento para ser rodado com dados muito complexos de análise.

Todo o código implementado durante o projeto encontra-se no seguinte link:

Repositório de Inteligência Artificial Preditiva Disponível em: <https://github.com/barbarafernandesmadera/predictive-artificial-intelligence>. Acesso em: 07 mar. 2025.

Além disso, participei de reuniões semanais com os supervisores Leonardo Trigueiro e Renata de Paris, para alinhamento de tarefas e análise de progresso.

8 Análise Crítica

Pontos Fortes:

- Aprendizado e aplicação prática de técnicas de séries temporais.
- Uso de ferramentas corporativas para controle de versões e dashboards.
- Experiência de integração de dados reais e modelagem preditiva.

Pontos a Melhorar:

- Ajustes de hiperparâmetros exigiram mais tempo de experimentação.
- Automação completa dos pipelines pode ser aprimorada.
- Modularização de códigos implementados durante o estágio.

Contribuições do Curso: As disciplinas de estatística, álgebra, programação da graduação e inteligência artificial forneceram a base necessária, mas a aplicação em séries temporais foi aprimorada de forma prática no estágio tanto por implementação de códigos quando de cursos teóricos online.

8.1 Comprovante de Estágio

O comprovante de estágio está em anexo a este relatório, como mostrado na figura 1. Abaixo, segue a imagem do documento que valida a realização do estágio no Instituto de Pesquisa Eldorado.



Figura 1 – Comprovante de Estágio no Instituto de Pesquisa Eldorado.

9 Conclusão

O estágio obrigatório permitiu aplicar e consolidar conhecimentos de Engenharia de Computação, com foco em Ciência de Dados aplicada a Séries Temporais meteorológicas. O aprendizado prático no Instituto de Pesquisa Eldorado proporcionou experiência técnica e profissional, alinhada às necessidades do mercado de trabalho.

Durante o estágio, enfrentei alguns desafios relacionados à complexidade dos modelos de previsão de séries temporais, como ARIMA e XGBoost, além da adaptação ao ambiente corporativo e à colaboração com equipes multidisciplinares. No entanto, essas dificuldades foram superadas com o suporte da equipe de pesquisa e a aplicação dos conhecimentos adquiridos durante minha formação acadêmica, como os conceitos de modelagem matemática e análise de dados.

Acredito que a experiência foi de grande importância para o meu desenvolvimento profissional, pois além de aplicar os conceitos aprendidos, tive a oportunidade de aprimorar habilidades técnicas, como o uso de ferramentas de versionamento de código e análise de dados em grande escala. Além disso, a interação com outros profissionais me permitiu entender melhor as dinâmicas de trabalho em equipe e as exigências do mercado de tecnologia.

Como sugestão para futuras edições do estágio, acredito que uma maior ênfase em visita técnica, como o laboratório da Apple em parceria com a Eldorado de Brasília o qual se encontra na Universidade Católica de Brasília , a qual pode proporcionar uma experiência ainda mais completa e enriquecedora. No geral, o estágio contribuiu significativamente para a minha formação, ampliando meu conhecimento e preparando-me melhor para os desafios profissionais que enfrentarei no futuro.

10 Referências

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 10719:2015 — Relatório técnico e/ou científico – Apresentação. Rio de Janeiro: ABNT, 2015.
- INSTITUTO NACIONAL DE METEOROLOGIA. Dados meteorológicos. Disponível em: <https://portal.inmet.gov.br/>. Acesso em: 06 jan. 2025.
- UDEMY. Master Time Series Analysis and Forecasting with Python 2025. Disponível em: <https://github.com/barbarafernandesmadera/Master-Time-Series-Analysis-and-Forecasting-with-Python-2025>. Acesso em: 06 jan. 2025.
- UDEMY. Python for Data Science and Machine Learning Bootcamp. Disponível em: <https://github.com/barbarafernandesmadera/Python-for-Data-Science-and-Machine-Learning-Bootcamp>. Acesso em: 06 jan. 2025.