



Diogo Eugénio Pinto

Nacionalidade: Portuguesa Data de nascimento: 26/10/1999

📞 Número de telemóvel: (+351) 926774879

✉️ Endereço de email: diogoeugenipinto@gmail.com

🌐 Sítio Web: <http://www.diogopinto1.com>

📍 Casa: Sintra (Portugal)

SOBRE MIM

Curiosity → Data → AI → Impact

EXPERIÊNCIA PROFISSIONAL

▣ **Horizon Europe - INESC TEC** – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Data Scientist - Bolsa de Investigação

[01/01/2025 – Atual]

Projeto Horizon-AI: Previsão do nível freático e mitigação da poluição hídrica no Mediterrâneo.

- Criação de pipeline de data mining e modelação preditiva de séries temporais.
- Estruturação de base de dados relacional e integração de dados climáticos e hidrológicos.
- Aplicação de métodos estatísticos (Spearman e Pearson, com time lags) para avaliar correlações entre variáveis e preditores externos para os modelos de previsão.
- Desenvolvimento de modelos Machine Learning (Random Forest, LSTM, ARIMAX) para previsão da profundidade do nível da água.
- Integração e automatização do pipeline com foco em reproduzibilidade e monitorização contínua dos dados num dashboard final.

Tecnologias: Python, Pandas, NumPy, Scikit-learn, TensorFlow (Keras), PostgreSQL, Git, Azure, Docker, Visualização (Matplotlib, Seaborn, Plotly), FullStack(html, css e js) e React.

▣ **HIDROSPH** – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Data Scientist - Dissertação de Mestrado

[12/2023 – 11/2024]

Dissertação Mestrado: Preenchimento inteligente de lacunas de dados sensoriais - Aplicação de *Machine Learning* em Sistemas de Rega.

- Desenvolvimento de um pipeline de *Machine Learning*, metodologia CRISP-DM, para correção de inconsistências em dados provenientes de sensores IoT.
- Garantia de continuidade e qualidade dos dados, assegurando a sua prontidão para utilização futura em algoritmos internos da empresa e integrações operacionais.
- Seleção e implementação de serviços de cloud computing para ingestão, processamento e serving dos modelos.

Tecnologias: Sensores IoT, SIG (NDVI/SAVI), serviços de Cloud Computing (AWS, s3), Pandas, NumPy, Scikit-learn, TensorFlow (Keras), MySQL, Git, Docker e REST APIs.

▣ **The Summer Berry** – Odemira, Portugal

Cidade: Odemira | País: Portugal

Data Scientist/ Data Engineer - Projeto Hackathon

[10/2023 – 12/2023]

Projeto Hackathon: Modelo preditivo (*Forecasting*) de produtividade para a produção de frutos vermelhos.

- Fluxo de Ciência de Dados (*ML-driven Automation*) em equipa de 5 pessoas para previsão de produtividade de framboesas.
- Recolha e preparação de dados de campo (padronização, limpeza e extração de *features agronómicas*).
- Desenvolvimento em base de dados local, com esquema relacional para históricos de colheita, clima e fenologia.
- Criação de modelo de Machine Learning para previsão de produtividade (*Random Forest*), com validação temporal e afinação de hiperparâmetros.
- Implementação de dashboard interativo (aplicação final) para monitorização de previsões e suporte à decisão.

Tecnologias: MariaDB (SQL), DBeaver, OpenRefine, Python, Pandas, NumPy, Scikit-learn (RandomForest), Plotly e Matplotlib.

Syngenta – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Assistente Regulatório - Estágio Extracurricular

[08/2023 – 09/2023]

Regulamentação e Sustentabilidade - Consultoria no Departamento Regulatório com foco em conformidade e práticas sustentáveis.

- Gestão e armazenamento de informação (organização, normalização e controlo de versões de documentação técnica e legal).
- Colaboração no projeto **Operation Pollinator**, apoiando a implementação, monitorização e reporte de iniciativas de biodiversidade e sustentabilidade.

Tecnologias: Ferramentas de gestão documental (SharePoint/Google Drive) e Office 365 (Excel).

Águas do Tejo Atlântico - Grupo Águas de Portugal – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Data Scientist - Estágio Extracurricular

[06/2023 – 08/2023]

Projeto Aquasafe: Análise e tratamento de dados, e desenvolvimento do IDE.

- Centralização da informação (Base de dados relacional PostgreSQL) com esquema operacional e histórico.
- Monitorização do curso de água com base em diferentes sensores (nível, velocidade, caudal, volume) incluindo validação e controlo de qualidade dos dados.
- Atualização e desenvolvimento do dashboard interativo no Aquasafe, com agregações semanais e mensais por localização para as variáveis Nível, Volume, Velocidade e Caudal.
- Automatização da ingestão e agregação temporal, documentação e versionamento.

Tecnologias: Microsoft SQL, Sensores/Telemetria, Python, Pandas, Docker e Power BI.

CONSULAI – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Data Scientist - Estágio Extracurricular

[06/2022 – 08/2022]

Departamento I&D: Consultoria/Digitalização nos setores Agrícola, Alimentar, Florestal, Mar e Desenvolvimento Rural.

- Research e web scraping de fontes oficiais (INE, GPP, EUROSTAT) com controlo de qualidade e catalogação de metadados.
- Centralização e modelação de dados em Excel, com Power Query e DAX (Power Pivot) para criação de modelo analítico.

- Desenvolvimento de uma API (script orientado ao INE) para automatizar a extração e atualização de dados.
- Análise de mercado do preço do abacate em Portugal e Espanha e criação de dashboard em Power BI para monitorização e reporting.

Tecnologias: Python (BeautifulSoup/Scrapy), REST APIs, Excel (Power Query, Power Pivot, DAX), Office365, SharePoint, Power BI, Git e FastAPI/Flask.

Marinha Portuguesa – Lisboa, Portugal

Cidade: Lisboa | País: Portugal

Oficial da Marinha

[04/2017 – 11/2017]

Prestação de serviço em regime de contrato como Oficial, na classe Fuzileiro.

EDUCAÇÃO E FORMAÇÃO

Mestrado em Ciência de Dados em Agricultura, Alimentação, Floresta e Ambiente

ULisboa - Instituto Superior de Agronomia [01/09/2022 – 30/11/2024]

Cidade: Lisboa | País: Portugal | Sítio Web: <https://www.isa.ulisboa.pt> | Classificação final: 18 | Nível no QEQ: Nível 7 QRQ | Tese: Intelligent Sensory Data Gap Filling: Application of Machine Learning in Irrigation Systems

Licenciatura em Engenharia Agronómica

ULisboa - Instituto Superior de Agronomia [01/09/2018 – 09/06/2022]

Cidade: Lisboa | País: Portugal | Sítio Web: <https://www.isa.ulisboa.pt> | Classificação final: 13 | Nível no QEQ: Nível 6 QRQ

Curso Científico-Humanístico de Ciências e Tecnologias

Escola Secundária de Santa Maria em Sintra [2015 – 2017]

Cidade: Lisboa | País: Portugal | Nível no QEQ: Nível 4 QRQ

COMPETÊNCIAS LINGUÍSTICAS

Língua(s) materna(s): português

Outra(s) língua(s):

Inglês

COMPREENSÃO ORAL C2 LEITURA C2 ESCRITA C1

PRODUÇÃO ORAL C1 INTERAÇÃO ORAL C1

Espanhol

COMPREENSÃO ORAL B2 LEITURA B2 ESCRITA B1

PRODUÇÃO ORAL B2 INTERAÇÃO ORAL B2

Níveis: A1 e A2: Utilizador de base; B1 e B2: Utilizador independente; C1 e C2: Utilizador avançado

COMPETÊNCIAS

Base de Dados (SQL Server e MySQL) / Microsoft Office (Word, Excel, PowerPoint, PowerBI, MSProject) / JavaScript, Shell Script, Power Shell, Python, Java, HTML5, CSS3, Bootstrap e React / Remote Sensing / SIG (ArcGIS, QGIS) / AZURE Cloud / AWS cloud / PowerBI PowerQuery DAX (advanced) / Programação em Python, C e C++ / MLOps / DevOps / Frontend, Backend e Fullstack / LangChain / análise de dados

CARTA DE CONDUÇÃO

Carta de condução: A1

Carta de condução: A2

Carta de condução: B1

Carta de condução: B

PROJETOS

Portfólio Pessoal - www.diogopinto1.com Dei forma a um portfólio pessoal, em *React*, para centralizar e apresentar os projetos que fui construindo ao longo do percurso académico e profissional. Enquanto Data Scientist, este portfólio é essencial para demonstrar resultados, metodologias e impacto dos mesmos de forma mais clara e acessível.

Ligaçāo: <http://www.diogopinto1.com>

QUALIFICAÇÕES

[12/2025]

Advanced Learning Algorithms

Entidade: Standford University

Competências:

- Construir e treinar redes neurais artificiais em TensorFlow para tarefas de classificação multiclasse.
- Desenvolver e utilizar modelos de árvores de decisão e métodos de *ensemble*, incluindo random forests e *boosted trees*, para problemas de classificação e regressão.
- Fazer afinação de hiperparâmetros (*performance tuning*) e avaliação de modelos para melhorar o desempenho e a capacidade de generalização.
- Aplicar boas práticas de desenvolvimento de sistemas de *machine learning* (divisão treino/validação/teste, prevenção de *overfitting* e *data leakage*, etc.) em cenários do mundo real.
- Integrar princípios de IA responsável e ética dos dados no ciclo de vida de modelos de *machine learning*.

Tecnologias: TensorFlow, Python, scikit-learn, Keras, RNN, Jupyter Notebook.

Ligaçao: <https://www.coursera.org/account/accomplishments/verify/WR3I2R3D8NRR>

[11/2025]

Curso de Ética e Deontologia Profissional (107ª edição)

Entidade: Ordem dos Engenheiros - Região Sul

Competências:

- Compreensão dos princípios de ética e deontologia profissional aplicáveis à engenharia.
- Conhecimento do Código Deontológico da Ordem dos Engenheiros e do respetivo enquadramento legal e disciplinar.
- Análise e resolução de casos práticos de dilemas éticos na atividade de engenharia.

Ligaçao: https://diogopinto1.github.io/portfolio/Certificado_CED_107_Diogo_Pinto.pdf

[01/11/2025 – 07/11/2025]

Supervised Machine Learning: Regression and Classification

Entidade: Standford University

Competências:

- Construir modelos de *machine learning* em Python utilizando bibliotecas populares como NumPy e scikit-learn.
- Construir, treinar e validar modelos de aprendizagem supervisionada para tarefas de previsão e classificação binária, incluindo regressão linear e regressão logística.
- Desenvolver modelos de árvores de decisão (CART) para problemas de regressão e classificação.
- Preparar, limpar e transformar dados (engenharia de *features*) para melhorar o desempenho dos modelos preditivos.
- Avaliar modelos de regressão e classificação com métricas adequadas e optimizar hiperparâmetros.
- Comparar diferentes algoritmos de *machine learning* supervisionado e seleccionar o modelo mais adequado ao problema.

Tecnologias: Python, NumPy, scikit-learn, Jupyter Notebook.

Ligaçao: <https://www.coursera.org/account/accomplishments/verify/1AOAEEXANO3F0>

[12/02/2025 – 14/02/2025]

Formação para Docentes - Ciência de Dados

Entidade: Faculdade de Ciências da Universidade de Lisboa

Docente: Luís Correia

Competências:

1. Introdução

- Os principais passos e componentes da ciência de dados

2. Análise e Manipulação dos dados

- Qualidade dos dados
- Visualização e métricas simples
- Extração de características

3. Aprendizagem automática e Prospeção de dados

- Modelos
- Aferição de resultados – métricas
- Seleção de atributos
- Outros tipos de dados – temporais, espaciais, não-estruturados (imagem, vídeo, grafos)

4. Interdisciplinaridade

[12/02/2025 – 14/02/2025]

Formação para Docentes - Control, Intelligent Systems and Robotics

Entidade: Instituto Superior Técnico, Department of Mechanical Engineering

Center of Intelligent Systems - idMEC

Docente: Jorge Martins

Competências:

1. Control

- A brief History
- Foundations: Concept of Signals and Systems
- Representation: Input-Output and State-Space
- Concepts of Feedback and Feedforward
- Stability vs Performance: a trade-off
- Control Techniques: PID, Optimal, Nonlinear

2. Intelligent Systems

- Fuzzy Systems
- Neural Networks
- Learning

3. Robotics

- Kinematics, Dynamics and Control
- Machine Intelligence

[22/08/2023 – 22/08/2023]

Joining Data in SQL

Entidade: DataCamp **Duração:** 4h

Competências:

- Trabalhar com mais de uma tabela em SQL, usar inner joins (outer joins e cross joins);
- Incluir cláusulas (union, intersect e except), criar queries adaptativas

[09/08/2023 – 09/08/2023]

Intermediate SQL

Entidade: DataCamp **Duração:** 4h

Competências:

- Como filtrar e comparar dados; como utilizar funções de agregação para resumir dados;
- Como ordenar e agrupar dados; como apresentar os dados de forma limpa e rápida, utilizando ferramentas como aliasing

[09/02/2023 – 10/02/2023]

Programação Python do zero ao Avançado

Entidade: Udemy **Duração:** 15.5h

Competências:

- Linguagem Python do básico ao avançado.

[01/2023 – 02/2023]

Curso Completo de Ciências da Computação

Entidade: Harvard CS50 **Duração:** 25h

Competências:

- C, Arrays, Algoritmos, Memória, Estruturas de Dados, Python, SQL, HTML, CSS, JS, Cibersegurança

[19/07/2022 – 19/07/2022]

Curso de Geoprocessamento

Entidade: Udemy **Duração:** 3.5h

Competências:

- Geoprocessamento e QGIS aplicados em agricultura de precisão

[09/2021 – 09/2021]

Curso de SIG

Entidade: Instituto Superior de Agronomia **Duração:** -/-

Competências:

- Unleash QGIS with Python