VITOR SOUSA - Data Scientist

in vitorhcsousa vitorhcsousa.github.io/

vitorhcsousa

@ vitorsousavhcs@outlook.com

Portugal



EXPERIENCE

Data Scientist

FARFETCH

Jun 2022 - Present

Porto.Portugal

- Development of advanced size recommendation models employing LSTM, GRUs, and attention models. Achieved a substantial increase in top1-accuracy by 22%-45% compared to the previous solution.
- Implemented innovative features, including the incorporation of new signals such as "add to bag" resulting in a remarkable 24.5% increase in user coverage.
- Successfully initiated and spearheaded an onboarding project with the primary objective of creating an automated Size & Fit recommendation identification model. This pioneering model, validated through manual verification, is currently in production, and it has already contributed valuable notes for over 25,000 products, significantly enhancing product recommendations and user experience.
 - The model demonstrated a successful result by reducing return rates by 7% in Men and 3.6% in Women.

Data Scientist

Sonae MC

May 2021 - Jun 2022 Porto, Portugal

- Engineered predictive models (RFM, behaviour segmentation, churn prediction, sentiment analysis) using diverse data sources, notably Continente Card, Portugal's premier loyalty program.
- Enhanced marketing impact predictions and strategies by blending machine learning expertise with business acumen.
- Developed an interactive Streamlit platform for analyzing models, empowering crossfunctional business teams.
- Successfully migrated models from SAS to Python/PySpark on Azure, ensuring robust deployment and ongoing maintenance.

Data Scientist & Engineer

Valuekeep (PRIMAVERA BSS)

苗 Aug 2020 - May 2021 📍 Braga, Portugal

- Employed dual methodologies for component failure prediction: Regression to predict failure timing, and Classification to forecast failure within specific intervals.
- Created micro-services to capture and store data within Azure Data Lake. Developed pipelines for data transformation and processing, primarily utilizing Azure Data Factory and Databricks. Established a software access point to integrate model outcomes.
- Nominated for Fast Shining Award that distinguishes the revelation of the vear.

Software Engineer

Valuekeep (PRIMAVERA BSS)

苗 Jun 2018 - May 2021 📍 Braga, Portugal

- Develop a cloud based software for maintenance with .NET/.NET Core (C#), SQL, EntityFramework, and Angular.
- Managed product life cycle, overseeing Azure environment for best practices and CI/CD implementations. Create an access point for our software to feed in the model results.

Software Developer

Pinto Brasil Group

苗 Oct 2017 – Jun 2018 📍 Guimarães, Portugal

*Summer Internship Jun-Aug 2017 (from 15h to 20h, conciliated with my work at Idioma de Tons(06h-14h))

- Develop web solutions- Idea Management Software; R&D products interfaces.
- R&D project of Prediction and identification of the moments and causes of failures of the department's products.

Textile Operator

Idioma de Tons

• First Professional Stint: Catalytic for financial independence and university education pursuit. This phase, a vital juncture, honed the work-study balance, shaping identity and fostering achievements. Anchors all subsequent successes through profound personal growth.

EDUCATION

Engineering and Management of Information Systems (Integrated Master's)

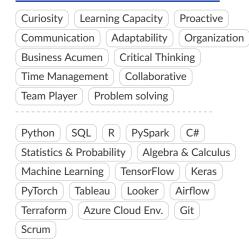
University of Minho

Sept 2015 - Dec 2020

© Excellence Scholarship Student with the best grade of the course

Thesis: Artificial Intelligence for Predictive Maintenance (17/20)

SKILLS



LANGUAGES

Portuguese **English**



VOLUNTEERING

Data Scientist

Data Science for Social Good Portugal

Mar 2021 - Present

PUBLICATIONS

Journal Articles

- Candeias, A., Silva, I., Sousa, V., & Marcelino, J. (2023). Tailor: Size recommendations for high-end fashion marketplaces. 16th ACM Conference on Recommender Systems - Fashion RecSys, 2023.
- Ferreira, L., Pilastri, A., Sousa, V., Romano, F., & Cortez, P. (2021). Prediction of maintenance equipment failures using automated machine learning. Intelligent Data Engineering and Automated Learning - IDEAL 2021, 259-267.