+447465951120
Nottingham, United Kingdom
danielabreu1738@gmail.com

Daniel Pereira De Abreu
Data Scientist - MSci Computer Science & AI

Portfolio: danielabreu.co.uk
github.com/daniel2tio
linkedin.com/in/daniel2tio

TECHNICAL SKILLS

Languages	**Proficient:** Python, Typescript and SQL **Experience:** MATLAB and R
Tech Stack	Tensorflow, AWS, Grafana, MySQL & PostgreSQL, Apache Spark/Airflow, MLFlow, Docker, DVC, Git, Anaconda and Visual Studio Code

EXPERIENCE

Data Science Intern   June 2023 — September 2023
Vis4Sense   *Nottingham, UK*

- Developed a custom JupyterLab extension to automatically capture and archive the history of generated images from a PyTorch-based diffusion model, enhancing reproducibility and traceability for artists.
- Leveraged TypeScript and Python within a Conda environment to deliver a robust, modular solution for a scalable development.
- Integrated real-time monitoring using Grafana dashboards to visualize key performance metrics and hyperparameter trends, facilitating data-driven model optimization.
- Established the foundation and pipeline that has since been adopted and expanded by the team, driving ongoing project innovation and scalability.
- Followed an agile workflow with version control via Git and weekly feedback sessions to ensure iterative development and alignment with project goals.

Artificial Intelligence Engineer Intern   June 2022 — September 2022
BluSkeye AI   *Nottingham, UK*

- Enhanced an advanced emotional intelligence facial recognition model by improving its capacity to analyze mental health through valence/arousal metrics, facial action units, and driver behaviors, now supporting applications in health and well-being, automotive, and social robotics.
- Automated the conversion of human-annotated datasets into structured CSV files for seamless comparison with software-generated annotations such as Nova.
- Developed and applied DBScan clustering algorithms to evaluate and benchmark the alignment between human-generated and synthetic data, ensuring data quality and consistency.
- Organized a data collection program, gathering and coordinating 100 participants to enrich the training datasets.

Machine Learning Software Intern   September 2021 — May 2022
Ideagen   *Nottingham, UK*

- Designed and developed an advanced meeting summarizer that uses state-of-the-art transformer-based models like BERT and T5, utilizing attention mechanisms and positional embeddings to produce accurate and efficient natural language summaries.
- Applied the sentence-transformer model with cosine similarity algorithms to map sentences into vector space and generate personalized summaries tailored to user input keywords, significantly enhancing summary relevance.
- Implemented Named Entity Recognition (NER) to classify entities such as dates, people, and locations, enabling streamlined content categorization and information retrieval.
- Collaborated with the team to deploy the model using AWS, integrating S3 Bucket for storage, Lambda functions for serverless processing, and SQL for data categorization, ensuring scalability and cloud efficiency.

EDUCATION

MSci Computer Science & Artificial Intelligence, *The University of Nottingham*	September 2020 - June 2024
A Levels - Computer Science (A), Mathematics(A) & Geography(A), *Luton Sixth Form College*	September 2018 - June 2020
D*D* (A*A*) IT National Diploma Level 3 , *Luton Sixth Form College*	September 2018 - June 2020
9 GCSEs Grades 9-7 (A*- A), *The Reach Free School*	September 2013 - May 2018

EXTRAS

Certifications
- Deep Reinforcement Learning (Hugging Face)
- Machine Learning A-Z Python & R (Udemy)
- The Ultimate MySQL Bootcamp (Udemy)
- Deep Learning with PyTorch for Medical Imaging (Udemy)
- Natural Language Processing (Udemy)
- Complete MLOps Bootcamp (Udemy)

Projects
- Gesture-controlled 3D Hologram System
- Telco Customer Churn Prediction
- Synthetic Data Generation for Cancer Prognosis Detection
- Ant Colony Optimization
- Scaling of time-series classification in Big Data
- "PromptAll" - JupyterLab Extension