Tomás Premoli Computer Science Student – Machine Learning & AI

tomas@premoli.org LinkedIn Github
https://www.tomaspremoli.com/

Education

University of Exeter *MSci Computer Science* 2020 – 2024
Thesis: *Convolutional Neural Network Architectures for the Classification of Alzheimer's Disease MRI Scans*, Grade: 73
Grade Avgs: Year 1: 78, Year 2: 77, Year 3: 74
Computer Science Subject Chair 2021/2022, leading Staff student Liaison Committee meetings on behalf of the student's guild to represent computer science students for the university

Employment History

Amazon *Reliability & Maintenance Engineering* June – Sept 2023
Software Development Engineer Intern. *Barcelona, Spain*

- Migrated and upgraded a large portion of **data engineering** cloud infrastructure to the **AWS CDK**.
- Achieved **CI/CD** and enhanced development experience for essential big data management workflows.
- Furthered **ML** skills by taking Machine Learning University (MLU) trainings in **NLP** (RNNs, LTSMs, Transformers) and **CV** (CNNs, GANs).
- Developed skills in data-oriented cloud infrastructure and setting up DevOps workflows.
- Furthered **data engineering** skillset, and gained experience in **AWS Lambda**, **Glue**, **EMR**, and **IAM**.

Amazon *Search Science & AI* June – Sept 2022
Software Development Engineer Intern. *Barcelona, Spain*

- Formed a part of the Search M5 team, focused on training and deploying **large scale foundation models** to hundreds of teams across Amazon.
- Created an **NLP** model training pipeline, used to train, finetune, distill, and evaluate Transformer-based **ML** models.
- Automated the training of multi-modal models containing over 200 million parameters each, with options for hot-swapping configuration files and creating novel workflows.
- Gained experience using **Apache Airflow**, **AWS EC2**, and **AWS Batch**.

Projects

Latent Diffusion Augmentation 2023 – 2024
Manipulating Diffusion Model Latent Spaces to Improve Image Dataset Augmentation Quality & Diversity.

- Ongoing Master's thesis, experimenting with **Latent Diffusion Models** for the task of dataset augmentation.
- Aiming to manipulate Diffusion Models for the use of class-conditional image dataset augmentation, hoping to improve the quality and diversity of generated images, and evaluating the performance of classifiers trained on the artificial data.
- Completed literature review and training of **LDMs**. Currently analyzing results and designing further experiments.

Technologies used: Python, PyTorch, Hugging Face Diffusers, Google Colab

ADMIRE-DL 2022 – 2023
A Comparison of Convolutional Neural Network Architectures for the Classification of Alzheimer's Disease Patients Using MRI Scans.

- Developed a deep learning model to accurately classify Alzheimer's disease in patients using MRI scans.
- Trained various **CNNs** to classify collections of MRI image slices, achieving an accuracy of **86.36%** with DenseNet201.
- Quantified the impact of model depth and trainable parameter count on performance, along with the efficacy of training from scratch compared to transfer learning.
- Created tools for data preprocessing, model training and evaluation, and wrote a paper documenting experiments.

Technologies used: Python, TensorFlow, AWS EC2

Skills

ML Skills	TensorFlow, Pandas, NumPy. Neural Networks (CNNs, RNNs, GANs) & good foundation in statistics.
Languages	Bilingual fluency in both English and Spanish.
Coding	Python, Java, C/C++, HTML, JavaScript, React, Django …
AWS	s3, Lambda, IAM, EC2, Elastic Beanstalk, MWAA (Apache Airflow), AWS CDK, VPC, EMR, Glue, …
Misc.	Strong presentation and public speaking skills. Excellent communicator and collaborator. Good writer.