ALBERT DOMINGUEZ

Interested in exploring both the **Artificial Intelligence** and the **Physics** worlds, **building bridges** between them as I find my way in, always while **exploring** other worlds too. Highly **motivated** by **difficulties** and non-trivial problems.

EDUCATION

Data Science and Engineering Bachelor's Degree

Universitat Politècnica de Catalunya   *Sept. 2017 – Present*

- Mix between telecommunications, statistics, mathematics and computer science, focusing on data-oriented processes.
- Relevant subjects coursed in the degree:
  - Algorithmics
  - Statistics
  - Mathematical Optimization
  - Machine Learning and Deep Learning
  - Data Analysis and Visualization
  - Signal Processing (Image and audio)
  - Elective courses in Quantum Computing & Reinforcement Learning

Scientific Baccalaureate

IES Sant Feliu de Guíxols   *Sept. 2015 – Jun. 2017*

- Awarded with "Honours" distinction.

EXPERIENCE

Junior Data Scientist

AI Department, Abi Global Health (Barcelona)   *Feb. 2020 – present*

A telemedicine startup which enables users to send health questions to real doctors from inside their favorite chat apps

- Development and maintenance of a tool to easily deploy and monitor models to AWS SageMaker, as well as Deep Learning models themselves mostly oriented towards solving medical NLP tasks.
- Assistance in analytical reporting to several departments.

Data Science Intern

Business Intelligence Department, Softonic   *Jul. 2019 – Feb. 2020*

An established software platform containing one of the biggest software catalogs in the world

- Developed an app which generates new program descriptions based on existing ones using a transformer architecture based on ELMo.
- Carried out maintenance and tuning of an ML-based ad system, using TensorFlow and GCP.

CONTACT INFO

- Barcelona, Spain
- +34 682 81 47 85
- albertadm99@gmail.com
- linkedin.com/in/albert-dominguez-mantes
- https://github.com/AlbertDominguez

SKILLS

- **Machine Learning**
- **Deep Learning**
- **Natural Language Processing**
- **Kernel Methods**
- **Computer Vision**
- Programming Languages:
  - Python: ●●●●●●
  - R: ●●●●●○
  - C++: ●●●●●○
  - SQL: ●●●●●○
  - MATLAB: ●●●○○○

- **C++: OpenMP, MPI,** …
- **Python: PyTorch, Keras, Sci-Kit Learn,** Numpy, Pandas, Plotly, …
- **AWS: SageMaker**
- **Google Cloud: AI Platform**, BigQuery
- **Docker**

LANGUAGES

- Catalan – Native
- Spanish – Native
- English – Written and oral fluency C2 equivalent certificate issued by Cambridge Assessment English in 2016.
- French – Elementary proficiency
- German – Elementary proficiency

PROJECTS

Infinite-width neural networks: combination and inference (2020)

- In this work, we studied certain kernel functions, namely the Neural Network Gaussian Process (NNGP) and the Neural Tangent Kernel (NTK), which have strong links with neural networks with an infinite number of neurons per layer. We derived that linear combinations of these kernel functions can be interpreted as ensembles of infinite-width neural networks. Furthermore, we experimentally tested the power of these kernel functions on a challenging problem of more than a million features and just 520 observations, overperforming finite neural networks. This project was developed under the supervision of professor Lluís A. Belanche.

Unsupervised clustering for ancestry inference with neural networks (2021)

- Developed as my final thesis project of the bachelor's degree in collaboration with the Department of Biomedical Data Science of Stanford University, CA. The work consisted in implementing a neural network-based version of the ADMIXTURE algorithm, which is a well-established but slow algorithm in population genomics. The goal was to perform at least on par with the ADMIXTURE algorithm while achieving a large speedup factor. Other significant work included modifications to this neural network to allow for local inference and several cluster assignments in a single forward pass of the network. The work was supervised by professors Xavier Giró-i-Nieto (UPC), Daniel Mas Montserrat (Stanford) and Alexander Ioannidis (Stanford).

Bioinformatic study of the *Influenzavirus* (2016)

- Developed as a research project during my baccalaureate, this project aimed to create a consensus genome sequence for the hemagglutinin and neuraminidase proteins of the *Influenzavirus A*, as well as a 3D render of the consensus of the mentioned proteins. Several techniques. such as BLAST and sequence alignments, were explored.

Clustering and Visualization of Hyperdimensional Medical Scans of the Eye (2020)

- Developed as a team of 10 as a project in the bachelor's degree, this project aimed to meaningfully cluster patients using limited OCT data. Techniques from both classical image processing, as well as recent Deep Learning techniques (Variational Autoencoders or U-Net variations) were used. Non-linear dimensionality reduction methods techniques embedded in a Dash app were implemented for data visualization.

Spam detection using string kernels (2019)

- Developed as a team of 2 as a project of the Machine Learning II course in the bachelor's degree. Several string kernels were used with support vector machines, as well as a custom implementation of kernel logistic regression, to solve the infamous spam dataset task.

OTHER ACTIVITIES

- **Participated** in:
    - Particle Physics Workshop, organized by Universitat de Barcelona (2017)
    - Google Hash Code (2019, 2020)
    - Ion Switching Kaggle competition (2020)