

Lucia Montero Sanchis

M.SC. COMMUNICATION SYSTEMS AND DATA ANALYTICS

contact@luciams.com

github.com/lumosan

EDUCATION

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

2017 – 2019

M.Sc. Communication Systems (specialized in Data Analytics)

Universidad Pontificia Comillas, Madrid, Spain

2012 – 2016

B.Sc. Telematics Engineering

EXPERIENCE

Machine Learning Engineer Intern

Feb. 2019 – Sep. 2019

Schindler Elevator Ltd., Switzerland

- Matched patents to product specifications using natural language processing and information retrieval.
- Achieved 70% accuracy using FastText word embeddings with Python, NLTK, Gensim and scikit-learn.
- Tested and improved the performance in a real-case scenario working with patent attorneys and engineers.

Research Assistant

Sep. 2018 – Feb. 2019

HCI in Learning and Instruction Lab, EPFL, Switzerland

- Analyzed writing strategies during online collaborative writing tasks, with Python and MySQL.
- Defined a set of features to analyze paragraph quality and consistency.
- Calculated the semantic similarity between authors at different stages of the writing process with sentence embeddings.

Data Scientist Intern

Jul. 2018 – Sep. 2018

Dathena Science Ltd., Singapore

- Deployed new features to a production environment in Scala and Spark for classifying sensitive documents.
- Improved by 30% the accuracy of an optical character recognition pipeline for noisy documents, by correlating and denoising partial matches using reference text corpora with Python, Tesseract and scikit-learn.

Collaboration grant

Jan. 2017 – Jun. 2017

Big Data Lab, Universidad Pontificia Comillas

- Designed and tested a distributed programming environment with Docker and Spark for future student projects.
- Deployed pipeline using Kafka, Flume and Cassandra as streaming data sources for the analysis of tweets with Spark.

PROJECTS

Recommender system with collaborative filtering algorithms and model stacking

- Trained nearest neighbor methods and latent factor models to predict user ratings with Python, NumPy and scikit-learn.

Deep learning for predicting the laterality of movements in EEG recordings

- Trained LSTM and CNN models with PyTorch for predicting the laterality of finger movements from EEG recordings.
- Evaluated the effects of downsampling the signals and applying preprocessing to improve the accuracy.

Bills in Congress visualization

- Created an interactive data visualization for insights on bills introduced in US Congress with D3.js and React.js.

Python Flask app using IBM Cloud services

- Deployed a Python Flask application for language translation and visual recognition with IBM Cloud, using Travis CI.

PROGRAMMING LANGUAGES AND TECHNOLOGIES

Languages Python, Java, Scala, JavaScript, HTML/CSS, C
Analytics SciPy, pandas, NumPy, R, Spark, Matlab

Machine learning PyTorch, TensorFlow, scikit-learn
Miscellaneous Git, Linux, Bash, SQL