Sergio Marrero

Data Scientist - Machine Learning Engineer

Data Scientist-Machine Learning Engineer with professional experience on A-Z machine learning projects and used to work on agile-minded data teams. I spent the last year developing **Deep Learning Architectures** applied to **Natural Language Processing (NLP)**. I define myself as a passionate worker, which is why I like to work with passionate coworkers. **I am looking for a position closer to backend engineering.**

Personal Info

Address
Barcelona, Spain

Phone
(+34) 667 563 883

E-mail
sergiomarreromarrero@gmail.com

Date of birth
1987-12-03

WWW
https://rpubs.com/SergioMarrero

GitHub
https://github.com/SergioMarreroMarrero

LinkedIn
www.linkedin.com/in/sergio-marrero

Software

Python ■■■■■

R, SQL ■■■■■

Python ML stack ■■■■■

Languages

English ■■■■☐

Spanish ■■■■■

Experience

2020-07 - present

Data Scientist Specialist

CornerJob | Eurofirms Group

Principal Tasks:

- **A-Z Deep Learning Model for candidate's photography acceptability prediction**
  - *Description:* Development and deployment of a candidate's photography acceptability prediction
  - *Technologies/tools used* : Deep Learning (Transfer learning) architecture, Python, Pytorch, AWS

- **A-Z Natural Language Processing (NLP) model development to detect fraudulent and scam offers intentionality.**
  - *Description* : Natural language processing model to detect fraudulent intent and spam in user's offers
  - *Technologies/tools* : Python, Hugging Face, Deep Learning: Transformers (BERT), CNNs, word embeddings. Pytorch, Spacy, AWS

- **Natural Language Process POCs (proof of concept)**
  - *Description:* Other NLP models were developed but not implemented in the production environment. For example:
    - Name Entity Recognition (NER) to detect jobs position inside the offers
    - NER + *Linguistic* Desambiguation to detect context spam words
  - Technologies/tools: Python, Hugging Face, Transformers (BERT), Pytorch, Spacy, AWS, Word2vec representation, AWS

- **Design and create dataset annotation processes**
  - *Description:* Data infrastructure development and team coordination to label new samples with the goal of building an NLP model in top of that.
  - *Technologies/tools:* prodigy, PostgreSQL, SQLite

- **Maintenance and improvement of the backend of the web application used by the moderation team to detect fraudulent content and spam.**
  - *Description* : Acquisition of new requirements, debugging and improving the processes, planning and proposing new functionalities to be implemented, coordination with stakeholders
  - Technologies/tools: Python, Django, ElasticSearch, MongoDB, AWS

- **Statistics Analysis - AB testing:**
  - *Description:* Statistical tests and data visualization in order to understand and validate hypothesis or expand data knowledge

- *Technologies/tools:* Python, Pandas, scipy, pymc3, altair, plotly, matplotlib, seaborn

2019-03 -
2020-07

Data Scientist

Vueling Airline

Principal Tasks:

- **A-Z unsupervised machine learning model**
  - *Description:* Development and implementation of a customer segmentation model to segment customers based on their purchasing patterns and behavior
  - Technologies/tools: Python, scikit-learn, kmeans, pandas, PCA, AWS

- **A-Z supervised machine learning model for Business-Leisure probability prediction**
  - *Description*: A-Z ML project to predict Business-Leisure probability
  - Technologies/tools: Python, R, sklearn, Spark, Hive, xgboost, AWS

2017-07 -
2019-01

Data Scientist

Quantified Statistics & Data Analysis

Principal Tasks developed:

- Developing statistical and probabilistic models for betting markets
- Technologies/tools: R, Python, tidyverse, environment, shinyapps, beatifulsoup, sklearn, JAGS

2017-03 -
2017-06

Data Scientist Junior

Domingo Alonso Group

Principal Tasks developed:

- Cleaning and preparing data sets for analysis with R
- A/B testing machine learning models

2016-01 -
2017-02

Collaboration grant for research at the University Institute of Intelligent Systems SIANI

Las Palmas de Gran Canaria University

Supporting to the University Institute of Intelligent Systems SIANI in ULPGC

Education

2017 -
present

PhD Student in Bayesian Deep Learning

Thesis Title: *Contributions to the Uncertainty on Deep Learning Models (Bayesian Deep Learning)*

University*: Las Palmas de Gran Canaria University*

University *Institute*: *University Institute of Intelligent Systems and Numeric Applications in Engineering SIANI*

(PhD Thesis was not submitted)

2016

M.Sc Intelligent System and Numeric Application on Engineering

(Master Thesis was not submitted)

2007 -
2014

Industrial Engineering with mention on Optimal Systems and Digital Electronic

These studies are equivalent to: B.Sc + M.Sc

Thesis delivered with Cum Laudem

Publications and Congresses

2017-07   **A comparison study of the Morgenstern-Price methods coupled with a heuristic optimizer forslope stability analysis.** *ISBN: 978-84-697-4340-9, Edited by: ScienceKnowconferences, Canary Islands,Spain (2nd Applied Computing in Science and Engineering (ACSE2))*

2017-06   **Integration of renewable sources of energy in low voltage distribution network. Case smallwind turbine.** *ISBN: 978-84-697-4150-4, Edited by: ScienceKnowconferences, Canary Islands, Spain (7th International Congresson Energy and Environmentand Management (CIIEM7))*

2017-06   **Sustainable energy application in a multifunctional industrial warehouse.** *ISBN: 978-84-697-4150-4, Edited by: ScienceKnowconferences, Canary Islands,Spain (7th International Congresson Energy and Environmentand Management (CIIEM7))*

Additional Activities

2017-05 -   **Dr.Stetter ITQ**
2017-07
Organizer and instructor of a Deep Learning Workshop

Organizer and instructor of the Workshop: *Deep Learning (first edition in ULPGC-SIANI)* refunded by german tech company named Dr.Stetter ITQ