Álvaro Domínguez Calvo

Machine Learning Engineer

+34 634 69 06 11
dominguezcalvoalvaro@icloud.com
Website
LinkedIn Profile

Summary

- Machine Learning Engineer with almost three years of experience and solid knowledge in Natural Language Processing. Currently, I work in a data science team to deliver machine learning solutions to meet customer needs related to different car bidding platforms.

Experience

•Ciklum *2021/02 - Present*

Machine Learning Engineer Remote

- In a Agile environment and Continuous Integration/Delivery workflow, my team develops and maintain machine learning/deep learning models that predict the retail price of cars in bidding platforms. My work consists of gathering car data and textual user data, analyzing, training and evaluate different machine learning algorithms with different vectorized car representations to improve the estimation of the retail price.
- In a team of three data scientists, I lead a deep learning project based on a classifier to predict the brand and models of cars given their images with the objective to obtain vectorized representation of the cars. We scrape car images on a weekly basis and then, we feed them to a transformer (Vision Transformer).
- I collaborated closely with the development team to integrate, in different car bidding platforms, a recommender system based behavioural data of the users.

•Ingenia *2019/09 - 2021/02*

Full Stack Developer Málaga - Spain

- I worked with the development team as a full stack developer. We developed, test and maintain the accounting system for Ingenia.

Education

•Universidad Nacional de Educación a Distancia *2020 - present*

Master's degree in Natural Language processing Spain

•Universidad Nacional de Educación a Distancia *2019*

Computer Science Spain

Side Projects

•Automatic Fake News Detection in Social Media *Present*

- Master's thesis in Natural Language Processing oriented to detect fake news within social data. I exploit contextual and user information under the hypothesis that some users are likely to post non-contrasted information.
- **Technologies**: PyTorch, Optuna, Pretrained Models, Fine-Tuning, Word Embeddings

•Multilingual trend analysis and event detection *2021*

- I scraped, processed and analysed Twitter data of the COVID desease and La Palma volcano eruption to detect latent topics contained within the tweets and to cluster tweets semantically similar. The goal of this project is to find evidence that there is no difference in talking about a particular topic given a set of languages.
- **Technologies**: Python, Pandas, Latent Dirichlet Allocation, clustering algorithms, pretrained word embeddings

Technical Skills and Interests

Languages: Python, C#, SQL
Developer Tools: Git, Docker, Elastic Search, pretrained word embeddings
Frameworks: PyTorch, XGBoost, Scrapy, Gensim, Optuna, Dask
Cloud/Databases: Amazon Web Services, Azure Devops (TFS), Octopus
Soft Skills: SCRUM, Teamwork, communication, problem-solving
Areas of Interest: Natural Language Processing, Social Media Mining, Text Mining, word embeddings