MOHAMED MAZY – Data Scientist

London | +44 7 845 884 037 | mazymohamed06@gmail.com
French Native & English B2 | LinkedIn | Github

EDUCATION

CENTRALESUPECLEC (PARTNERSHIP W/ MICROSOFT) — Paris, FR
Master's degree in Artificial Intelligence — Expected August 2022

UNIVERSITE NICE COTE D'AZUR — Nice, FR
Bachelor's degree in Computer Science — Sep 2017 - Jun 2020

WORK EXPERIENCE

ALAN MCFETRIDGE PHOTOGRAPHY LTD — London, UK
Junior Data Scientist — Feb 2022 – Present

- Applied the Data Science on artistic ecologic field. No final goal, researcher role.
- Geospatial Data Analysis in order to show the impact of wildfire over the boreal forest.
- Task: Imagery satellite analysis, Air quality analysis, Temperature analysis.
- Stack: Python, Google Earth Engine, NASA and ESA API, QGIS software.

TOP SOUVENIRS & LUGGAGE LTD — London, UK
Full Stack Developer — Sep 2021 – Feb 2022

- I had to build an e-commerce website for a small shop in order to be more visible.
- Stack: Django (Python), React (TypeScript), Tailwind (CSS), REST API, Docker, PostgreSQL, ElasticSearch.

GOOGLE WORKSHOP MACHINE LEARNING — Paris, FR
Student — Nov 2019

- Selected by Google London during my third year university in order to discover their products. It was the first time I heard about Machine Learning. Google London team explained how to use Tensorflow and showed different ML applications.

PROJECTS

RECOMMENDATION SYSTEM — Jun 2022

- Situation: A company want to build a mobile app of books reading and build a recommender system to recommend new books to users.
- Task: Use a serverless architecture for the back end and to link it with the front end. Build a recommender system, to train and save the recommendations in database.
- Action: EDA, Pre-processing, Matrice User-Item, Models testing, Save in ComosDB on Azure, Use Android Studio to test the app.
- Results: Mobile App based on an MVP architecture.

SELF-DRIVING SEGMENTATION — May 2022

- Situation: The company want to build a model in order to make an autonomous car. There are three task to do: image processing, image segmentation and decision making.
- Task: Do the image segmentation task using a segmentation model.
- Action: EDA, Pre-processing, , Data Augmentation, Models testing, Deployment model on Azure, Flask WebApp with REST API
- Results: Unet + MobileNet with 0.63 IoU score.

DETECT BAD BUZZ WITH DEEP LEARNING — Apr 2022

- Situation: An airline company want to detect all negative reviews against their company in order to improve their services.
- Task: Build a Deep Learning model in order to maximize the correct detection of negative sentences.
- Action: EDA, Pre-processing, Vectorization, Models testing, Deployment model on Azure, Streamlit WebApp with REST API
- Results: A Neural Network with 0.78 AUC score.

TOPIC MODELING & IMAGE LABELING Mar 2022

- Situation: A customer reviews company want to know the dissatisfaction topics and to label automatically images uploaded by a customer.
- Task: Build a ML model (unsupervised learning) for NLP in order to detect dissatisfaction topics and another Deep Learning model to label images.
- Action: Scrapping, EDA, Pre-processing, Vectorization, Model testing, Transfer Learning.
- Results: NMF used for NLP and we obtained 15 dissatisfaction topics, and for labelling a CNN using VGG16 by fine tuning we obtained 0.92 accuracy score.

CUSTOMER SEGMENTATION Dec 2022

- Situation: E-commerce company want to segment their customers in order to make a marketing decision based on their habits.
- Task: Build a Machine Learning model (unsupervised learning) to segment customers in clusters, then build a Machine Learning model (supervised learning) to interpret clusters based on features. Finally estimate every how long the model must be retrained.
- Action: EDA, Data Cleaning, Features Engineering, Models testing, Model interpretation, ARI.
- Results: 6 clusters with K-Means (0.49 Silhouette score and 0.76 Davies Bouldin score) and interpreted with XGBoost and Shape. The model must be retrained every month.

LOAN PREDICTION Nov 2022

- Situation: Financial company want to automize the checking process if a customer is eligible to loan and to explain why the customer is eligible or not.
- Task: Build a Machine Learning model (supervised learning) to minimize the risk to give a loan to a customer who will have payment default.
- Action: Exploratory Data Analysis, Data Cleaning, Features Engineering, Models testing, Model interpretation.
- Results: LightGBM was the most performant with 0.74 AUC and 0.4 F2 score and a model well explained.

SKILLS

PYTHON, PANDAS, TENSORFLOW, REACT, DEEP LEARNING, MACHINE LEARNING, NUMPY, KERAS, AZURE, SEABORN, COMPUTER VISION, NLP, DJANGO, DOCKER, PLOTLY, REST API, FLASK, SQL

CERTIFICATIONS