Maria Afara

ML/NLP Engineer
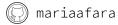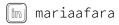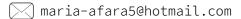
portfolio mariaafara mariaafara maria-afara5@hotmail.com 650082715

EXPERIENCE

ACOLAD | NLP Engineer
March 2021 – Present | Paris, France

➜ Researched, prototyped, built features and optimized machine learning and deep learning models using scikit-Learn, keras, and pytorch on CPU/GPU environments.
➜ Containerized(Docker) and deployed re-producible machine learning models/APIs into production through GitHub actions CI/CD onto Google Cloud Platform.
➜ Automated tools by building and managing data pipelines using Airflow.
➜ Developed/Tested/Maintained internal text-processing libraries.
➜ Solved business issues include classifying domains of translation projects using classification model, identifying personal information based on GDPR using NER models, and automated the creation of neural machine translation models.

SAP LABS | NLP Data Scientist
March 2020 – April 2020 | Sophia Antipolis, France

➜ Researched & developed SOTA techniques for information extraction & sentence level classification
➜ Performed weak supervision approach to label unlabelled data
➜ Designed, developed, and tested several topic classifiers using transfer learning for confidential French dataset.
➜ Created custom Named Entity Recognition models using spaCy.
➜ Built a large English dataset about real people biographies which highly contains PII using DBpedia KB and Wikipedia dumps.
➜ Built a Relation Extraction pipeline that combines Named Entity Recognizer & Relation Classifier to identify the sensitive information according to GDPR.
➜ Mostly worked with Keras, Tenserflow, Spacy, Pandas, Numpy, Matplotlib, Sklearn, mBert, CamemBert, Snorkel, PySpark, BiLSTM, CNN, FastText, SPARQL, Git, Docker

PROJECTS

QUESTION ANSWERING SYSTEM
➜ Implemented question answering system through semantic similarity matching
➜ Bert, MongoDB, Docker, Milvus, Pandas, FastApi...

NEURAL MACHINE TRANSLATION
➜ Worked with Encoder-Decoder architecture with and without attention mechanism to build a seq2seq models to translate English sentences into French sentences.
➜ Tensorflow, LSTM, GRU, Sklearn, Pandas, Numpy

BREAKING NEWS DETECTION
➜ Built a summarization model that summarizes a text and a classifier that predicts if the summarized news is breaking or not, prior to that the training data was labeled by unsupervised learning
➜ Tensorflow, Keras, Pytorch Word2vec, CNN, Sklearn, Pandas, Numpy, LDA

SKILLS

Python • PySpark • SpaCy
NLTK • SparkNLP • Keras
Scikit-learn • TensorFlow
PyTorch • HuggingFace

Airflow • FastAPI • Docker
SQL • GitHub actions CI/CD
Git • GCP • Gradio • Streamlit
Jupyter • WandB • PyCharm

EDUCATION

MASTER 2 IN NLP
UNIVERSITÈ DE LORRAINE, FRANCE
Coursework: Deep Learning, Machine Learning, Data & Text Mining, Recommendation System, Speech & Text Processing, etc.

MASTER 1 IN AI AND BIG DATA
LEBANESE UNIVERSITY, LEBANON
Coursework: AI, ML, Big Data, Cloud Computing, Design Patterns, Distributed Systems, Advanced Databases, etc.

BACHELOR'S IN COMPUTER SCIENCE
LEBANESE UNIVERSITY, LEBANON
Coursework: Data Structures, Graph Theory, OOP, Databases, Project Management, Software Engineering, etc.

LANGUAGES

Arabic	Native
English	Fluent
French	Beginner