ELMEHDI LAMSYAH

Data Scientist/ Data Engineer

lamsyahmehdi@gmail.com | +212622694448 | Casablanca
elmehdi-lamsyah-a04bba149 | elmehdiLAMSYAH

PROGRAMMING

Python | R | JavaScript | Scala | Julia | C++ | SQL | NoSQL

FRAMES-LIBS

Tensorflow | Pytorch | Keras | bert | Transformers | cuda | NLTK | Gensim | openCV | Scikit-learn | pandas | numpy | Django | flask | Spark | Hadoop | Rasa | D3.js | bokeh | Three.js | BeautifulSoup | Selenium

ADD-SKILLS

business intelligence
Statistical inference
Operations Research
Data Visualization
GCP | AWS
Docker
VCS
Agile Mindset: scrum

LANGUAGES

Arabic: Native
French: B2
English: B2

EXPERIENCE

Data engineer @ GAISIE'SO

March 2022- Now — Casablanca (remote)

- Automatizing building chatbots based on RASA
- Machine Learning models integration in AWS instance (Django backend)

NLP Engineer @ Integrate An Affiliate of Kantar

November 2021- February 2022 — Casablanca

- Improving the accuracy of a Moroccan Arabic automatic speech recognition system
- Language model for predicting next word for Arabic language
- Sentiment analysis for Moroccan Arabic
- word embedding model for North African Arabic
- clients clustering

AI-ML engineer intern @ Vokaris

march-august 2021 — Casablanca

- French Speech Recognition for e-commerce
- Sentence Textual Similarity
- Cloud based train/deploy on **GCP**
- Use and integrate **GCP** solution APIs

WEB| BI Developer intern @ Presidency of UM5

January-February 2021 — Rabat

- Django and SQLlite in back-end
- React in front-end
- Reporting Dashboard
- Employee Segmentation

PROJECTS

Moroccan Arabic analytics

- Description : WEB based platform containing multiple NLP tools :
  - Word cloud
  - Word embedding
  - Sentiment analysis
  - Audio improvement
  - Speech to text
  - Form answer based clustering
- Techniques: Django, Transformers, GCP APIs, LSTM, TensorFlow

French Automatic Speech Recognition (0.11 WER)

- Description : transfer learning of the open-source Speech-to-Text engine **DeepSpeech** model from Mozilla Common Voice based on Google TensorFlow and Baidu paper

- Techniques: Python, CNN, RNN, LSTM, CTC, Transformers, audio processing, TensorFlow, Pytorch

Article Classification

- Description : BERT|Transformers based model for multi-classification text
- Techniques: Python, Transformers, Gensim WV

Twitter Streaming Analysis

- Description : The model returns the sentiment (POS, NEG, NEU) of tweets in Streaming process (real time)
- Techniques: Flask, Scala, Spark-stream, NLTK, Stanford-coreNLP, twitter-API, KAFKA

Soccer Match Result Prediction

- Description : return regression prediction based on scraped data from the Web
- Techniques: Flask, Python, Selenium, BeautifulSoup, LSTM, RNN

Facial Expression Recognition

- Description : multi-classification model returns facial expression (7 emotions) based on LBP texture feature extraction tools
- Techniques: Python, TensorFlow, CNN, LBP. OpenCV

E-Book Store Recommendation System

- Recommendation System: collaborative filtering, item-based filtering
- Back-end Development: NodeJS (expressJS), MySQL ,REST API
- Front-end : reactJS

EDUCATION

Master's Degree | Intelligent Processing Systems

2019 – 2021 — mohammed 5 university, RABAT, Morocco

Bachelor's Degree | Computer Science

2016 – 2019 — Sidi mohammed ben abdellah, TAZA, Morocco

Baccalaureate | MATH -A-

2015 — Ali ben barri, TAZA, Morocco

CERTIFICATES

Stanford University Online

- Machine Learning by Andrew Ng

IBM

- Machine Learning using Python

Deeplearning.ai

- Neural Networks and Deep Learning
- Natural Language Processing with Classification and Vector Spaces
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Convolutional Neural Networks

University of California San Diego

- Introduction to Big Data

University of Minnesota

- Agile Software Development
- Software Development Processes and Methodologies