

Mohammad Farhat

mohammadfarhat1333@gmail.com

71 912 713

Data Science graduate Lebanese University, Faculty of Information

Education

Bachelor's in data science

Lebanese University, Faculty of Information 2020 – 2024

42school student

42Beirut,

2024 – present

Professional Experience

Intern, Salem Group *HR Company*

September 2023 – November 2023

- Utilized BeautifulSoup for web scraping to gather data from websites.
- Automated processes using Selenium Python for tasks like login, data extraction, etc.
- Applied Pandas and NumPy for data cleaning and analysis.
- Developed data analysis algorithms to predict the best suitable major and skills for a job posting.
- Cleaned and preprocessed data from LinkedIn profiles.

Work, Norna Analytics *Data Scraping Company*

February 2024 – June 2024

- Leveraged Data Scraping techniques to gather data from websites.
- Automated processes using Selenium Python, Scrapy and Playwright for tasks like login, data extraction, etc.
- Utilized proxies and cookies to bypass any website blockage.
- Worked on fixing and maintaining the scraping config files on a daily basis.
- Ensured every website was scraped correctly and stored in MongoDB with all details needed.

Projects

Mental Health Chatbot Development

- Developed a mental health chatbot using an already collected dataset about emotion recognition.
- Preprocessed and cleaned the data, performing exploratory data analysis (EDA) to extract valuable insights.
- Tokenized the data and transformed extracted word similarity using Word2Vec.
- Conducted A/B testing on various models, including regression model, Naive Bayes model, and SVM model.
- Explored advanced transformer models, including Electra and Roberta, to improve results.
- Identified overfitting issues with Roberta transformer and selected Electra transformer for integration.
- Integrated Electra transformer with Dialogpt pre-trained model API for enhanced performance.

Customer Segmentation Project

- Leveraged machine learning techniques to divide data into clusters/groups (5 clusters obtained).
- Utilized Python and scikit-learn for data preprocessing, feature engineering, and model training with Kmeans Clustering.

Technical Skills	Soft Skills
<ul style="list-style-type: none"> • Languages: Python, R, Java • Data: Pandas, NumPy • Web Scraping: BeautifulSoup • Automation: Selenium • Machine Learning: Scikit-Learn • Visualization: Seaborn, Matplotlib, Tableau • Database: SQL, MongoDB 	<ul style="list-style-type: none"> • Motivated: Enthusiastic and results-driven. • Problem Solver: Innovative and effective at simplifying complex challenges. • Creative: Visionary and inventive problem-solving. • Perseverant: Tenacious and resilient in overcoming obstacles.