

# Muhammad Umar Salman

DATA ENGINEER · DATA SCIENTIST · NLP SCIENTIST

4.08, Hydro Building, Masdar City, Abu Dhabi, United Arab Emirates

☎ (+971) 561-732912 | ✉ umar.salman@mbzuai.ac.ae | 🏠 umar1997.github.io | 📱 umar1997 | 🌐 umar-salman

## Education

### Mohamed Bin Zayed University Of Artificial Intelligence

Abu Dhabi, UAE

MASTERS OF SCIENCE

Aug 2021 - Present

- **Major:** Natural Language Processing      **CGPA:** 3.95
- **Relevant Courses:** Advanced Natural Language Processing, Big Data Processing, Speech Recognition, Artificial Intelligence

### Lahore University of Management Sciences

Lahore, Pakistan

BACHELOR OF SCIENCE (HONORS)

Aug 2016 - June 2020

- **Major:** Computer Science      **CGPA:** 3.50
- **Relevant Courses:** Data Science, Data Mining, Machine Learning, Deep Learning, Data Structures, Algorithms, Databases

## Experience

### Mohamed Bin Zayed University Of Artificial Intelligence

Abu Dhabi, UAE

GRADUATE ASSISTANT

Jan 2020 - Present

- Worked for multiple courses: **Data Mining** and **Machine Learning**
- Helped design and analyze the course structure, lab sessions and various graded instruments for the aforementioned courses
- Facilitated the students with learning and building holistic understanding of various concepts in the courses
- Evaluated student projects, labs, reports and other graded components.

### G42 Healthcare

Abu Dhabi, UAE

DATA SCIENTIST (INTERNSHIP)

May 2022 - Aug 2022

- Designed and worked on the development of a natural language interface that can parse a users statement, transform it into a structured criteria representation and produce a **SQL** query conforming to an EHR (Electronic Health Records) common data model.
- Built a **NLP** pipeline which parses and preprocesses a user's statement and then passes it sequentially through to a Named Entity Recognition model, Entity Linking model and lastly a SQL Generation model which generates an executable query.
- For the project I primarily used **Python's** Deep Learning frameworks and libraries such as **PyTorch**, **Hugging Face** and **Scikit-Learn**.

### Decklaration

Lahore, Pakistan

PYTHON DEVELOPER | DATA SCIENTIST

Dec 2020 - Aug 2021

- Developed the backend of a Crypto + Finance News Aggregator in **Django** (Python) using **Postgres** (SQL) as a database in a team using its rest framework and deployed it on **AWS**
- Designed and implemented a **Recommendation System** based on Collaborative Filtering and using **Natural Language Processing** techniques for a Financial News Aggregator in **Python**.
- Created a complete end-to-end data pipeline solution for clients working with modern cloud-based tools **MYSQL - Spark - PowerBI** to create a seamless flow of data for the BI Team

### United Arab Emirates University

(Remote) Lahore, Pakistan

RESEARCH ASSISTANT

Oct 2020 - Feb 2021

- Developed simulation codes of a selection algorithm for On-Demand Mass Transport in **Python** that models the dynamic routing of public services to minimize the waiting time along with increasing bus capacity utilization

### Afiniti

Lahore, Pakistan

DATA ANALYST

June 2020 - Oct 2020

- Designed models to optimize pairing between call-center agents and customer call groups to maximize the profit of the company in the Artificial Intelligence department where I also analyzed the effects of the model on the data using **SQL, Python and R Language**

### Lahore University of Management Sciences

Lahore, Pakistan

RESEARCH ASSISTANT

Jan 2020 - June 2020

- Working alongside two instructors to construct different machine learning models which predicts possible diabetes after looking at prediabetic vitals and their food consumption behavior of individuals.
- Compared results of different using **Scikit-Learn's machine learning models** such as Naive Bayes, Support Vector Machine, Logistic Regression, Decision Trees, Random Forrest and lastly Multi-Linear Perceptron (MLP) to see which models performed better.

## Projects

---

### Digital Email Assistant

NLP ENGINEER

March 2022 - May 2022

- Developed a model which takes a speech source input from a user and converts it to a structured email and then sends it to the recipient.
- For the project I primarily used **GCP's Speech-to-Text** model, **PyTorch**, **Hugging Face** and **Scikit-Learn**.

### OLX Business Analytics (Capstone Project)

DATA SCIENTIST

Sept 2019 - May 2020

- Provided insights on the behavioral pattern of the customers on the OLX website platform through **clickstream** and **cluster analysis**.
- Constructed a conversion metric to identify serious buyer using various **Machine Learning** and **Association Mining Algorithms**.

### Maedaan App

DJANGO DEVELOPER

Feb 2021 - May 2021

- Developed the backend for an application for booking playing fields in **Django** using its rest framework to provide endpoint API's.
- Deployed the backend application on **AWS' EC2 instance** and connected it to **AWS S3** bucket and **AWS RDS** as a storage and database solution.

### Forecasting Prices for Cryptocurrencies

DATA SCIENTIST

Oct 2019 - Dec 2019

- Built a **time-series model** in **Python** which forecasted future prices for Bitcoin using an **Auto-Regressive Model (ARIMA)**

## Skills

---

**Languages** Python, SQL, Java, Javascript, Golang, R, VBA, LaTeX

**Big Data Tools** Map Reduce, Apache Spark, Apache Airflow, Apache Kafka, AWS EMR

**Databases** MySQL, MongoDB, Postgres, Apache Cassandra, AWS RDS

**Development** Django, Node, Flask, AWS EC2

**Data Modelling** Pandas, Numpy, PyTorch, Tensorflow, Scikit-Learn, Selenium, Scrapy, NLTK

**Visualization** Matplotlib, Dash + Plotly, Charts.js, Seaborn, Tableau, Powerpoint

**Tools** Docker, Git, Slack, Postman, Jupyter Notebook

## Honors & Awards

---

|      |   |                         |
|------|---|-------------------------|
| 2022 | <b>Golden Visa</b> , Received Golden Visa under the category of talented individuals  | <i>Abu Dhabi, UAE</i>   |
| 2021 | <b>3rd Place</b> , 3rd Place at the G42 and Coders HQ #HackForSpace event   | <i>Dubai, UAE</i>       |
| 2020 | <b>High Merit</b> , Graduated with high merit from Lahore University of Management Sciences (LUMS)  | <i>Lahore, Pakistan</i> |
| 2019 | <b>Dean's Honors List</b> , Placed on Dean's Honors List for 3 semesters at Lahore University of Management Sciences (LUMS)                               | <i>Lahore, Pakistan</i> |
| 2020 | <b>Data Science Professional Certificate</b> , Earned an IBM Data Science Professional Certificate after completing a 9-course specialization on Coursera | <i>Lahore, Pakistan</i> |