Mohammed Omar **Khan**

(+971) 56 286 1951 | mkhan0138@gmail.com | itsmokha.github.io | itsmokha | m-omarkhan

Experience

Midis Group — *UAE*

Big Data Intern — *May 2023 - Present*

- Created data reports on tickets raised and solved by the team for stakeholders.
- Created Oracle Database clusters on Linux machines using Exadata and improved OLTP capabilties for customers by 3.6x times .
- Facilitated data warehousing using Exadata for customers.
- Improved data mining and analysis workflows for customers.

Heriot-Watt University — *Dubai*

Intern — *Sept. 2022 - May 2023*

- Mentored 50+ students in Python, machine learning algorithms, data visualization techniques, Jupyter, networking tools, VMs, Java, and other tools to enhance their practical skills by leading regular labs.
- Assisted professors in various courses taken under the Computer Science program at the university
- Assisted students in gaining hands-on experience in various projects.
- Analysed student performance in the courses and created reports on overall class performance for professors.

Education

Heriot-Watt University — *Dubai*

B.S. in Computer Science with Artificial Intelligence — *Sep. 2020 - June 2023*

- Received the Deputy Principal's Award which is given in recognition of academic excellence.
- GPA: 3.7/4.0

Certificates

July 2023 **Azure Data Fundamentals**, Microsoft
May 2023 **Career Essentials in Data Analysis**, Microsoft
July 2022 **Deputy Principals Award**, Heriot-Watt University
Dec. 2021 **Certified Junior Associate - Java**, Oracle

Projects

Dubai Property Market Analysis

https://github.com/itsmokha/Property-Price-Visualization-Dubai

- Analysis of dataset containing over 50,000 listings to visualise property rates in Dubai.
- Created the dataset by scraping Bayut.com and currently planning to scrape data from PropertyFinder to increase the dataset for potentially better analysis.
- Identified the overall market rates by neighborhood.
- Visualized the price per square foot in Dubai neighborhoods using PowerBI.

Customer Conversion Rate using Data Analysis

https://github.com/itsmokha/Customer-Conversion

- Analysis of a dataset containing 67000+ rows of data to analyze the effect of promotional campaigns on customer conversion rate and graphs to provide visual representation of collected data and gain insights on customer behavior.
- Achieved sentiment classification accuracy of 86% using various Machine Learning algorithms such as neural networks, regression, XGBoost and Random forests, logistic regression.
- Identified the best responding group for promotions using Uplift Modeling and used Qini curves to visualize accuracy.

Advertisement Detection using Machine Learning

https://github.com/itsmokha/Advertisement-detection-using-machine-learning

- Final year project where URLs were scanned to detect advertisement links using libraries such as Scikit-Learn, Tensorflow, Pandas, and Seaborn.
- Created a dataset using various sources totaling more than a million elements and extracted necessary features from the data.
- Worked with various machine learning algorithms and identified the best model for the task.
- Obtained an accuracy of 80% using various Machine Learning algorithms such as neural networks, regression, XGBoost, Support Vector Machines, LightGBM and Random forests and performed hyperparameter tuning to find the optimal parameters for the model.

Amazon Review Sentiment Analysis

Scikit-Learn, Tensorflow, Seaborn

- Analysis of dataset containing over 227,000 Amazon reviews to identify customer satisfaction with the product with the help of natural language processing, using Tensorflow and graphs to provide visual representation of the data.
- Achieved sentiment classification of 90% using various Machine Learning algorithms such as neural networks, regression, XGBoost , Light-GBM, Random forests, SVM (Support Vector Machines) classifiers and hyper parameter tuning.
- Identified overall customer satisfaction with various products and made inferences on best selling products.
- Visualized the most common sentiments regarding various products within the data.

Twitter Sentiment Analysis

Scikit-Learn, Tensorflow, Seaborn

- Analysis of dataset containing over 10,000 tweets to identify the sentiment expressed with the help of natural language processing, using Tensorflow and graphs to provide visual representation of collected data.
- Scraped tweets from Twitter using Python, and extracted, transformed and loaded (ETL) them to gain valuable insights.
- Achieved sentiment classification accuracy of 92% using various Machine Learning algorithms such as neural networks, logistic regression, XGBoost and Random forests.
- Identified common sentiments regarding artificial intelligence, and visualized the common themes within the data using word clouds and other graphs.

Airline Passenger Satisfaction

Jupyter Notebook, Scikit-Learn, Tensorflow, Seaborn

- Analysis of customer reviews to analyze customer satisfaction with airline.
- Thorough data analysis and machine learning models were conducted on airline customer review dataset.
- Worked with a mix of categorical as well as numerical data.
- Provided visual representation of the results using various graphs.
- Applied concepts such as Clustering, Decision Trees, K nearest neighbors and various others.

vscan

https://github.com/HWTechClub/Virus- Scanner

- A Python-based virus scanner that uses different APIs to scan files and websites for malicious content.
- Worked as project lead and identified the best tools to use.
- Improved security by implementing best practices regarding API keys
- Published the project on PyPi, making it pip-installable.

Extracurricular Activity

Heriot-Watt University Tech Club

Dubai

Tech Lead(Teamsize: 8)

Sept 2020 - Apr 2022

- Tech lead for various projects undertaken by the Tech Club.
- Published vscan project on PyPi.

Skills

Programming	Python, C, SQL, C++, Java, DAX, Bash Scripting, JavaScript, HTML, CSS
Data Science	Machine Learning, NLP, pandas, NumPy, Scikit-learn, PyTorch, Tensorflow, Keras, Jupyter, PowerBI, Seaborn, Matplotlib
Tools	Excel, Jupyter, PowerBI, Tableau, Git, VS Code, Bash
Soft Skills	Team Player, Quick Learner, Leadership, Communication, Motivated Worker
Languages	English, Hindi, Urdu