Ali Asghar Marvi

✉ aliasghar_marvi52@hotmail.com 📞 +4917689027927 📍 Rheingutstrasse 32, Konstanz, Germany ⚲ aamarvi.com

in linkedin.com/in/aamarvi ●● medium.com/@aliasgharmarvi ○ github.com/dsaam94

PROFILE

Results-oriented Data Scientist/ Data Analyst with 5+ years of experience translating complex datasets into actionable insights. Proven expertise in machine learning, statistical analysis, and data-driven strategy. Skilled in Python, R, SQL, and data visualization tools. Adept at enhancing decision-making processes and optimizing operations through impactful data solutions.

PROFESSIONAL EXPERIENCE

Data Scientist (Part-Time)
KNIME AG

03/2021 – present
Konstanz, Germany

- As a part of the Educator Alliance team, I contributed to the development and implementation of strategies that resulted in a 30% growth in downloads within the academic community.
- Developed a Time Series Extension in KNIME using Python, serving as a wrapper for Statsmodels and leveraging Pandas Date & Time functionality.
- Created and contributed to specialized spaces on KNIME Hub dedicated to Marketing Analytics, Supply Chain Analytics, and Digital Healthcare to promote KNIME applications within academia and the community.
- Named contributor in "Machine learning for marketing on the KNIME Hub: The development of a live repository for marketing applications ⧉ " published in the Journal of Business Research, December 2021.
- Engineered an Arduino weather station interfacing with the KNIME Server API for sensor data collection. Devised and implemented an ETL pipeline for efficient data processing and time series analysis.
- Designed and instructed L4 courses in Data Engineering (DE) and Data Visualization (DV) within KNIME.
- Developed a Text-to-SQL chatbot utilizing Large Language Models (LLM) as an interactive dashboard, showcasing proficiency in innovative solutions.

Process Automation: Continuous Integration/Continuous Deployment (CI/CD)
BaseX GmbH

04/2021 – 06/2021
Konstanz, Germany

- Polished dataset of research papers from basic design to DBLP suitable paper using CI/CD Process Automation in GitLab.
- Created Docker images for handling Node.js and Python-based dependencies for multiple integration tests.
- Configured GitLab to handle Docker files and monitor the CI/CD pipeline.

Data Engineer/Data Analyst
JS Bank Pvt. Ltd.

11/2019 – 10/2020
Karachi, Pakistan

- Developed data cleaning pipelines to generate refined data for analysis using PowerBI and .NET charts for data visualization.
- Applied data modeling techniques to design an optimized database schema for web applications, centralizing cash information for streamlined branch operations.
- Applied time series analysis to forecast cash demand for a specific branch, enabling proactive cash resource allocation. Successfully achieved a 30% reduction in costs incurred by the bank to insurance companies.

Data Analyst
K-Electric Pvt. Ltd.

02/2018 – 10/2019
Karachi, Pakistan

- Implemented DenseNet-121 to automate meter readings, optimizing the process and significantly reducing data capture time for meter readers.
- Produced data reports for Human Resources to track Employee Self-Service (ESS) requests received from employees and monitor the turnaround time of processing each request.
- Aided new connection business in report creation and development of an interactive dashboard in QlikView. Merged data from SQL Server and SAP sources for comprehensive analysis.

EDUCATION

Master of Science in Computer and Information Science
University of Konstanz

10/2020 – 11/2023
Konstanz, Gemany

Bachelor of Science in Computer Science
Institute of Business Administration

02/2014 – 12/2017
Karachi, Pakistan

PROJECTS

Ensemble Methods for Image Quality Assessment using Individual or Group Ratings   11/2022 – 06/2023
Master's Thesis

Applied ensemble methods, including bagging, boosting, and stacking, in Objective Image Quality Assessment (IQA) using the KonIQ-10k dataset, coupled with data processing to identify user clusters through Ward's minimum variance algorithm for Hierarchical Clustering (Agglomerative), enhancing evaluation metrics such as Spearman-Rank and Pearson-Linear Correlation Coefficients.

Spark Retail - Big Data Solution for Retail   01/2017 – 01/2018
Final Year Project (FYP) - Bachelor's

Created a comprehensive final-year project leveraging the Hortonworks sandbox environment to configure a robust four-computer cluster with Apache Hadoop, Spark, Oozie, Hive, Sqoop, and Ambari. Developed and implemented bash scripts in the Linux environment for sequential service management and integrated external RDBMS and BI tools for enhanced data visualization, strengthening analytical capabilities.

SKILLS

Python, **SQL** *(PostgreSQL, SQL Server)*, **Pandas**, **Statsmodels**, **Tensorflow**, **scikit-learn**, **Keras**, **PyTorch**, **R**, **KNIME**, **Matplotlib** *(Seaborn)*, **Tableau**, **D3.js**, **PowerBI**, **QlikView**, **Airflow**, **Big Query**, **Git**, **AWS**, **DBT**, **Problem Solving**, **Business Intelligence**, **Geospatial**