Hassan Ali Khan

Helsinki, Finland | P: +358 40 4487880 | hassankhan8818@gmail.com | linkedin.com/in/hassan8818/

SUMMARY

As a Data Science professional with over 1.5 years of experience, I am driven by a passion for building robust data infrastructures to enhance organizational decision-making and performance. With expertise in designing and managing data pipelines, ETL processes, and big data technologies, I ensure seamless data flow and system stability. My experience spans developing interactive business intelligence dashboards, implementing machine learning models, and creating automated reporting solutions that drive significant business value. Proficient in Python, SQL, and Power BI, I efficiently query, optimize, and manage data storage systems while translating complex data into actionable insights.

SKILLS

Skills: Azure (Certified), Airflow, Databricks, Power BI, ETL Processes, Big Data Processing, Azure Functions, Azure Data Factory, Data Pipeline Development, Data Analysis, Data Processing, Database Management, CI/CD, ER Modeling, PostgreSQL, MySQL, Python, SQL, NLP, Machine Learning, TensorFlow, PyTorch, LLMs.

EXPERIENCE

Systems Limited — Nov 2023 - Aug 2024

Data Engineer

- Proficiently managed databases like PostgreSQL, MySQL, and Teradata, overseeing data modeling (ER modeling), daily ETL operations, and CI/CD pipelines to ensure seamless data flow, system stability, and high performance data processing.
- Collaborated with cross-functional teams to translate business requirements into scalable, end-to-end ETL processes, leveraging Airflow for orchestration and ensuring reliable data infrastructure.
- Developed and optimized data pipelines to support predictive churn models, enabling efficient customer retention strategies through big data processing and analysis.
- Utilized Power BI for data visualization, along with Python and SQL for data manipulation and NLP tasks.

Codistan Ventures — Aug 2023 - Nov 2023

AI / ML Engineer

- Built and managed data pipelines for efficient preprocessing and feature engineering, integrating data from diverse sources to support churn analysis, model training, and machine learning workflows, leveraging Azure AI services for deployment.
- Collaborated with teams to design scalable data pipelines and implement AI workflows, utilizing machine learning, computer vision models, neural networks, deep learning frameworks (TensorFlow, PyTorch), and reinforcement learning to solve complex business problems.
- Engineered and deployed advanced LLMs, including Llama, BERT, and GPT-3, by developing data pipelines and applying transfer learning to address complex NLP challenges. Utilized Azure Machine Learning and Azure OpenAI for efficient deployment and scaling of these models in production environments.

EDUCATION

University of Eastern Finland – Joensuu, Finland
International Master's in Information Technology

NUCES, FAST – Islamabad
Bachelor of Computer Science

PROJECTS

BLINKIT – Retail Analytics Dashboard

- **Tools Used:** Power BI, DAX, Data Modeling
- **Project Impact:** Developed an interactive retail analytics dashboard for Blinkit, processing data from 5,500+ items across multiple store locations. The dashboard helped management track $776K in sales performance and make data-driven decisions about inventory and store operations.

AITRAX – Pulmonary Fibrosis Progression Using Deep Learning (Kaggle Competition)

- **Tools Used:** TensorFlow, Keras, Python, Deep Learning, CNN, Medical Imaging Analysis
- **Project Impact:** Developed and implemented a deep learning model for predicting pulmonary fibrosis progression using medical imaging data. The solution leverages neural networks to analyze CT scans, contributing to improved treatment strategies and patient care outcomes. Successfully deployed a model that aids healthcare professionals in tracking disease progression and making informed treatment decisions.

CERTIFICATIONS

Microsoft: Microsoft Azure Fundamentals (AZ-900)
IBM: Machine Learning with Python
IBM: Data Analysis With Python
IBM: Tools for Data Science
IBM: Databases and SQL for Data Science (Python)