Junaid H
Machine learning Developer

About Me

I'm an ML engineer with 6+ years of experience building intelligent systems and scalable data pipelines. With a strong academic background in Machine Learning & Software Engineering, I've led projects involving AWS, time-series, and customer analytics. My recent work includes synthesizing time series, LLMs agents, and time series forecasting.

Professional Experience

eBrands | Machine Learning Developer
Sep 2023 -

Key responsibilities:
- My responsbilites include; communicating with stakeholders to define the problem. Doing R&D for the ideas the C-suite has, developing those ideas into projects, deploying them in AWS, and then integrating them into the legacy system.
- The key projects I have contributed to include developing Sales Forecasting models, and implementing multiple RAG-based (Retrieval-Augmented Generation) Large Language Model (LLM) for deriving actionable customer insights and understanding the data (Ads, Marketing, Finance etc).

Nokia | Machine Learning
May 2022 -Sep 2023

Key responsibilities:
- Worked in the domain of Machine Learning, particularly on Synthetic Time Series Generation using different generative techniques.

Xisys | Machine Learning Engineer
July 2020 - April 2022

Key responsibilities:
- Developed AI solutions for enterprise security, including intelligent reporting, anomaly detection, & face recognition etc

Northbay Solutions | Software Engineer
July 2019 - July 2020

Key responsibilities:
- Worked on an enterprise Data-lake orchestration on AWS cloud to process data at different stages.

Data Science Journey | Academic Instructor
Part-time

Key responsibilities:
- Taught students from python programming for data science.

Awards and Honours

2019	University Gold Medal
2020	ICPC Regional - Pakistan: Rank: 3, Bronze Medal, + Asia West Continent Finalist, Tehran, Iran
2021	Erasmus+ Scholarship Holder
2016 -	Scored an upright position in over 10 national programming competitions

Tools & Technologies

- Python
- NLP
- Generative AI
- RAG
- AWS
- Forecasting
- Vector Databases
- Embeddings
- Prompt Engineering
- NodeJs
- C++
- Elasticsearch (ELK stack)
- Flask
- SQL
- Backend (Python)
- TypeScript
- Git
- Jira
- MLFLow

Education Background

- Dual Master's in Machine Learning & Software Engineering | Erasmus+ Joint Master's degree

  Mälardalen University and Åbo Akademi University

  GPA 5.0/5.0 & 4.94/5.0

- Bacehlor's in Software Engineering
  Universtiy of the Punjab – 2019

  GPA 3.92/4.0 (University Gold Medal)

Publication(s)

2023	Identifying gender bias in blockbuster movies through the lens of machine learning

Junaid H
Machine Learning Developer

Soft Skills

- Communication Skills
- Management
- Problem Solving
- Adaptability

Hobbies

- Chess
- Dancing
- Cooking
- Trying Out new things

Languages

- Urdu – Native
- Punjabi – Native
- English – Full working Proficiency (C2)
- Finnish – Basic (B1)

Volunteer Work

- ESN ÅA – Treasurer
- Exchange Tutor
- Volunteer – PoBS
- Publicity Head – Event Management Society (PUCIT)

Relevant Projects

- **Time Series Forecasting**

Tools and technologies: *Python, Python, AWS (S3, SageMaker), SQL*
1. I developed a time series forecasting pipeline to predict sales, inventory levels, and other key business metrics.
2. This involved building both statistical (ARIMA, ETS) and ML-based models, as well as integrating Amazon Forecast for scalable deployment. I implemented anomaly detection to flag outliers in historical data, improving model accuracy and reliability.
3. The system supported automated data ingestion, preprocessing, and retraining workflows on AWS, providing forecasts for business intelligence dashboards and operational planning.

- **LLM-Based Retrieval-Augmented Generation Platform**

Tools and technologies: Python, AWS Bedrock, SageMaker, Lambda, Vector DataBase, Docker
1. Designed and implemented a chat-based application based on a RAG pipeline for internal users.
2. Integrated internal (financial, ads, forecast data from ML models, etc) and external data sources(customer demographics engineered, weather data, etc).
3. Enabled the LLM to access this data either using knowledge bases or making them available on run time for retrieval and analysis.
4. Empowered stakeholders to gain insights into products and business by leveraging the LLM's ability to analyse and compare data effectively.
5. This is an ongoing project, where I continue to experiment with new models and techniques to improve performance, usability, and the overall user experience.

- **Synthesizing time series data using GANs**

Tools and technologies: *Python, Pytorch, tensorboard, GANs, Embeddings*
1. Developed a software for time series generation using GANs such as DoppelGANger, VQ-VAE, etc.
2. Proposed a hybrid using VQ-VAE improving the performance.
3. Worked on different quantitative and qualitative evaluation metrics such as downstream tasks and visualizations to measure the performance.
4. Optimized parameters and performed comparative analysis.
5. **Thesis can be found [here](#)**

- **Anomaly Detection Engine**

Tools and technologies: *Python, TensorFlow, Deep Learning, Unsupervised Machine Learning, ELK stack.*
1. Created a full ML pipeline that retrieves data from elastic search logs and segment users.
2. Then using different models including proximity-based, statistical, deep-learning and htm-based models trained user profiles, and kept track of the experiments using MLflow
3. The entire software was developed as a microservice and integrated into the legacy system.