KUMAR AWANISH

@ awanish00@gmail.com +4917647167989
Germany kumar.awanish

Profile Summary

Machine Learning Engineer/Data Scientist with 9+ years of experience in end-to-end data solutions, MLOps, predictive analysis and cloud deployments. Proven track record in cross-functional leadership, building ML pipelines, and transforming business objectives into actionable insights.

WORK EXPERIENCE

Senior Machine Learning Engineer

Yara International

May 2021 – Present Berlin, Germany

- Developed and deployed an agentic RAG-based recommendation system using Large Language Models on AWS Bedrock that intelligently matched user queries with personalized responses. This system balanced user experience with business goals, simultaneously increasing engagement by 35% while reducing manual effort by 40% through advanced classification and retrieval models.
- Managed and scaled cloud-based ML infrastructure using AWS CloudFormation, Amazon ECS, and CI/CD pipelines, cutting cloud infrastructure costs by 20% while maintaining robust system uptime and performance.
- Lead a cross-functional team for the development of a real-time product recommendation tool based on variational autoencoders (VAE) that generated 6 % additional leads(sales)
- Designed and developed a data pipeline in AWS using ETL Glue, PySpark, Redshift Spectrum & CICD principles, resulting in a 30% decrease in processing time for Data Lake & Data Modeling.
- Demonstrated strong mentorship capabilities by managing and coaching junior data scientists/engineers. Additionally, took an active role in conducting interviews for potential candidates, contributing to the team's growth.
- **Technologies & Concepts:** Python, NLP, OpenCV, Git, Docker, Airflow, Kinesis, LLM, RAG, Langchain, Spark, AWS (OpenSearch, Bedrock, Sagemaker),ETL/ELT, MLOps, A/B Test, MlFlow, Azure.

Data Scientist

Bundesanstalt für Materialforschung und -prüfung (BAM)

Sep 2020 – Mar 2021 Berlin, Germany

- Worked on an EU project for the digitalization of materials by developing an ETL framework and standard to extract and transform data from different materials experiments and load it into the data warehouse.
- Developed and implemented a machine learning-based model in close cooperation with material experts for the classification of liquids and liquid-cooled glasses based on CNN architecture and Python.
- **Technologies & Concepts:** Python, Scikit-learn, Docker, Flask, PostgresSQL, Git, Apache Airflow, CNN, Keras.

Data Scientist

Dreamlines GmbH

November 2019 – March 2020 Hamburg, Germany

- Designed a XGBoost based predictive recommendation tool for the Marketing Team to predict the next best offers and booking dates, which helped to reduce the cost of sending the newsletter by 20 %.

EDUCATION

M.Sc. in Computer Science(Major in Data & Software Engineering). GPA: 1.6 (Highest:1)

Technische Universität Berlin

2016 – 2019

B.Tech. in Computer Science & Engineering. 73.1%

Bangalore Institute of Technology

2009 –2013

STRENGTHS

Hard-working | Team Player | Self-motivated | Quick Learner | Good communication skills

TECHNICAL SKILLS

- Machine Learning (Regression, Classification, Clustering, Ensembles)
- Deep Learning (Autoencoders, CNNs, RNNs)
- Natural Language Processing (NLP) , Large Language Models (LLMs)
- Keras, PyTorch, Tensorflow, Transformers
- Scikit-Learn, Pandas, Numpy, Spacy
- Matplotlib, Plotly, Seaborn, Tableau, PowerBI
- Retrieval-augmented generation (RAG), Langchain
- Python, Pyspark, Flask, FastApi, Streamlit.
- ETL, Airflow, SQL, NoSQL(Redis, DynamoDB)
- Java, Unit Test, A/B tests, JIRA
- AWS, Azure, Docker, GIT, CI/CD, MLOps.

LANGUAGES

Fluent: English,
Beginner (A2): German

- Designed ETL workflows and contributed to data modeling during the data warehouse development phase.
- Setting up A/B tests together with marketing team.
- **Technologies & Concepts:** Python, Pytest, Scikit-learn, Docker, XGBoost, VAE, Flask, Keras, Redis, Git, AWS, Apache Airflow, Tableau, ETL.

Data Scientist

Fraunhofer FOKUS

April 2017 – November 2019 Berlin, Germany

- Machine learning based model development to predict user behavior for energy trading platform with an accuracy of 97 %.
- Smart Contract development for Blockchain based Energy Transaction, development & deployment of web service to production.
- **Technologies & Concepts:** Java, Python, Scikit-learn, Pandas, Docker.

Senior Software Engineer

Samsung R&D Institute India

July 2013 – September 2016 Bangalore, India

- Responsible for the development of web services for registering smartphone devices through Samsung Internet Of Things(S-IoT) platform.
- Spearheaded 6 people IT automation team that developed Tizen Automation Tool(TAT) to automate 90% of test cases for the Tizen SDK development.
- **Technologies used:** Python, Java, Selenium, Javascript, HTML, Unit Test.

ACADEMIC PROJECT

Master Thesis in "Spatial & Temporal feature extraction for gaming Video Quality Assessment(VQA)"

Technische Universität Berlin

April 2019 – November 2019 Berlin, Germany

- Goal: To extract the temporal & spatial features of Gaming videos and predict their video quality & complexity.
- Developed machine learning-based lightweight No-Reference(NR) metric, named BEG-VQ for the quality assessment of gaming videos which outperforms the state-of-the-art NR metrics.
- **Technologies & Concepts:** Python, Scikit-video, Pandas, Scikit-learn, Jupyter Notebooks, OpenCV, Keras, Neural Network, SVM, XGBoost.

Internet of Services Lab

Technische Universität Berlin

April 2018 – September 2018 Berlin, Germany

- The project's objective was to implement and evaluate supervised & unsupervised machine learning methods for intrusion detection.
- **Technologies & Concepts:** Python, Scikit-learn, Jupyter Notebook, Pandas, PySpark,Keras, t-SNE, Machine Learning(Random forest, Neural Network), Spark MLlib.

HOBBIES

Travelling
Visiting different countries to see & learn about new culture, people and Food.

Sports
Badminton, Cricket.

MISCELLANEOUS

Spot Award
Received SPOT AWARD in Samsung R&D Institute for outstanding performance in a project.

CERTIFICATIONS

Coursera
Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization

Coursera
Neural Networks and Deep Learning

Coursera
Structuring Machine Learning Projects

Coursera
Structuring Machine Learning Projects

Coursera
Google Cloud Platform Big Data and Machine Learning Fundamentals