Masihuddin Khan

Phone | Mail | LinkedIn | GitHub | Portfolio | New Delhi, India

Profile Summary

- Data Scientist with expertise in building machine learning models, analyzing large datasets, and delivering actionable insights. Proven ability to implement deep learning solutions for image segmentation, predictive analytics, and recommendation systems. Skilled in Python, SQL, and cloud-based data pipelines.

Technical Skills

- **Programming Languages**: DSA, Python, SQL.
- **Libraries/Frameworks:**: TensorFlow, PyTorch, Scikit-learn, Pandas, Django, Spark.
- **Machine Learning**: Supervised and Unsupervised Learning, NLP, Computer Vision.
- **Tools**: Jupyter, VS Code, Git, Docker, CI/CD Pipeline .
- **Data Visualization**: Matplotlib, Seaborn, Tableau.

Work Experience

Data Analyst Jun 2023 - November 2023
Navatar Consulting Pvt Ltd.w *Noida, Uttar Pradesh*

- Developed predictive models to optimize business processes.
- Worked on creating an end-to-end data warehouse solution using Hive.
- Analyzed large datasets to extract actionable insights and drive decision-making.

Training in Data Science Feb 2022 - Jan 2023
Croma Campus *Noida, Uttar Pradesh*

- Completed comprehensive training in data science methodologies and tools.
- Developed multiple machine learning models using Scikit-Learn,TensorFlow, PyTorch and evaluated their performance.
- Worked on real-world projects, including predictive modeling and exploratory data analysis.

Education

Bikaner Technical University July 2022- June 2025
B.tech Electrical Engineering *Rajasthan , India*

Diploma in Electrical Engineering July 2018- June 2021
Jamia Millia Islamia *New Delhi , India*

12th CBSE July 2016 - June 2018
Boys Seninor Secondary School *New Delhi, India*

Projects

Movie Recommendation System | *Python,ML,NLP* | *https://movierecommendationsystem00.streamlit.app/*

- Python Libraries: **SKLearn, Numpy, Pandas, Matplotlib, Django**.
- Deployed in Microsoft Azure.
- The movie recommendation system successfully provided personalized movie recommendations using a combination of collaborative and content-based filtering techniques. The hybrid approach showed the best performance. The system, deployed on Microsoft Azure, operates in real-time with 90 percentage accuracy.

AI Pipeline for Image Segmentation and Object Analysis | *https://ai-pipeline-for-image-segmentation.streamlit.app/*

- Developed an end-to-end AI pipeline using transformer-based models for image segmentation, object identification, and analysis.
- Implemented U-Net and Vision Transformers to enhance accuracy, generating a summary table with key data for each identified object.
- Optimized model performance using data augmentation and pre-trained models.

Netflix Inside (EDA) | *Python* | *https://github.com/Masihuddinkhan/Netflix-Exploratory-Data-Analysis-s*

- Libraries Used: **Python, Pandas,Numpy Matplotlib.**
- Analysis of user ratings revealed that content with higher production budgets, particularly in genres like Action and Thriller, tended to have higher viewership and better ratings. This insight suggests a correlation between production value and audience reception.