

Tushar Arora

Ghaziabad, Uttar Pradesh
+91-88510-84852
tushar.ai.dev@gmail.com

linkedin.com/in/tushararora-dev
github.com/tushararora-dev
tusharx.vercel.app

SUMMARY

Data Scientist with expertise in Analytics, Machine Learning, Deep Learning, Computer Vision, Natural Language Processing (NLP), Generative AI, and MLOps/LLMOps. Experienced in developing and deploying models and building scalable data pipelines. Adept at solving real-world business problems using AI-driven solutions.

TECHNICAL SKILLS

Languages & Libraries: Python, R, SQL, HTML/CSS, JavaScript, Numpy, Pandas, Matplotlib, Seaborn, Plotly, NLTK, SpaCy, BeautifulSoup, Selenium, PySpark, Pydantic

Machine & Deep Learning: Scikit-learn, TensorFlow, Keras, PyTorch, OpenCV, Optuna, XGBoost, LightGBM

Gen AI: RAG, Lang Chain, Prompt Templating, Fine-tuning, Vector Database(FAISS, Pinecone), Hugging Face

MLOps & LLMOps: Git, GitHub, Streamlit, Gradio, Flask, Django, Fast API, Docker, DVC, Dags Hub, CI/CD Pipeline, MLflow, Postman, Pytest, AWS, GCP, Azure

Tools: Linux, MySQL, Mongo DB, MS Excel, Power BI, Tableau, Databricks, N8n, Octoparse

EXPERIENCE

Banao Technologies

Sept 2024 – Nov 2024

AI/ML Engineer Intern

- Researched AI-driven audio-video course creation for automated learning content generation.
- Designed a WhatsApp chatbot using Gallabox for warehouse booking and UI/UX audit report generation.

Synapse Health Tech

Jun 2024 – Aug 2024

ML Engineer Intern

- Engineered 11 ML models and reports for healthcare issues like arrhythmia detection, opioid crisis, and FWA.
- Researched Agentic chatbot for patient appointment scheduling using fine-tuned LLMs (Large Language Models).

Forge Job Simulation

Apr 2024 – May 2024

Data Analyst & AI Engineer

- Designed AI model to reduce supply chain costs by 15% for Gala Groceries (Cognizant).
- Created a financial insights chatbot by parsing 10-K and 10-Q filings (BCGX).
- Visualized key metrics to address 4 business questions for an online retail store (TATA).
- Analyzed 7 datasets to extract social trends for strategic planning (Accenture).

Project-Based Internship

Jan 2023 – Jan 2024

Data Analyst & ML Engineer Intern

- Conducted medical market segmentation in a biotech startup (Feynn Lab).
- Analyzed gameplay data from Game Rust's server logs (Nitrado).
- Developed thyroid disease prediction model using structured medical data (iNeuron).
- Reviewed ad campaign data to assess performance for discontinuation (Globalshala).

PROJECTS

SEO Optimizer AI Agent

[Repository](#)

Generative AI

Streamlit, LangChain, PyTune, OpenAI

- Built a web app using GPT-4 to generate SEO content (tags, timestamps, titles, descriptions). Integrated DALL-E 3 for thumbnail generation and deployed using Docker on AWS EC2.

Self-Driving Car Mechanism

[Repository](#)

Deep Learning Project

Python, TensorFlow, Keras, OpenCV

- Trained CNN for steering, throttle, and reverse control. Achieved 0.03 validation loss via data augmentation. Deployed on Udacity simulator using Flask and SocketIO.

Uber Demand Prediction System

[Repository](#)

Machine Learning Project

Scikit-learn, Optuna, MLflow, DagsHub, AWS

- Forecasted cab demand by segmenting the city into 30 geo-clusters using MiniBatch KMeans. Engineered 15-min interval features. Achieved 0.079 MAPE with Linear Regression. Orchestrated pipeline with Streamlit.

Spotify Hybrid Recommendation System

[Repository](#)

Recommendation System

Dask, DVC, GitHub, Docker, Pytest

- Built hybrid system (30% content, 70% collaborative) with diversity-aware personalization. Used CI/CD with DVC. Released on AWS with S3, ECR, EC2, CodeDeploy, ASG, and CloudWatch via Blue-Green strategy.

Personal Projects

[Repository](#)

Data Analytics / NLP / CV

Web Scraping, Object Detection, Transfer Learning

- WhatsApp chat/PhonePe dashboard, Discord bot with 6 metrics & 9 graphs, adult content classifier, reverse parental alert system, GTA5 race recommender, YouTube influencer plugin, image diff PyPI library & more.

EDUCATION

MCA in Artificial Intelligence

CGPA: 8.9

Jain (Deemed-to-be University)

Jul 2023 – Jul 2025

BCA in Data Science

CGPA: 8.0

Institute of Technology and Science

Jul 2019 – Jun 2022

EXTRACURRICULAR

Participation: Blogathon, Hackathon, Volunteer, Class Representative, E-Sports Tournaments, Caster, Moderator

Interests: Research & Development, Management, Budget Traveling, Itinerary Planning, Following AI Trends

Certifications: Full Stack Data Science from iNeuron, CampusX, Udemy, YouTube, Applied Roots

Hands-on Experience: Stack Overflow Tag Prediction, Zomato/Swiggy Delivery Time Prediction, Microsoft Malware Detection, Netflix/Amazon/Facebook Recommended Systems, Quora Question Pair Similarity

Ongoing Projects: Child Image Generation using GAN/VAE (Shadi.com), Virtual Dressing Room using Deepfake, Hand Gesture-Based Nurse Calling (Arduino), File-based Cheating Detection in E-Sports Tournaments