





Tarun S

 GitHub |  LinkedIn |  tarunskumar11@gmail.com |  +91-8217232584

EDUCATION

Don Bosco Institute of Technology

June 2025

B.E. in Computer Science (Artificial Intelligence and Machine Learning)

CGPA: 8.00

EXPERIENCE

Unified Mentors

October 2024 – February 2025

Machine Learning Intern

- Developed predictive ML models for healthcare and performed classification tasks using **Python** and **Scikit-learn**.
- Implemented advanced Machine Learning techniques like **PCA**, **K-Nearest Neighbours**, **Logistic Regression**, **Random Forest**, **Gradient Boost** and **Vector search algo** to improve insights and overall model performance.

Varcons Technologies

October 2023 – December 2023

UI/UX Designer

- Worked on designing an **attractive website** for an e-commerce platform which focused on selling sneakers for Gen-Z.
- Developed a design system, created **wireframes** based on initial concepts, and finalized the web design using **Figma**.

Snap. Inc

June 2023 – January 2024

Snapchat Opinion Leader

- Built and led a campus community by creating immersive **Augmented Reality** experiences as part of the Program.
- Designed and published **46+** vibrant Augmented Reality Lenses, collectively reaching over **20,209,305** (20M+) views.

PROJECTS

Semantic Book Recommender using LLM

- Implemented a multi-model NLP pipeline leveraging **DistilRoBERTa** for emotion classification and **BART-large-MNLI** for **zero-shot text classification**, achieving an average accuracy of **77.8%** across diverse test datasets.
- Built a search-driven system using **LangChain**, **HuggingFace** sentence-transformers, and **Chroma** vector database, enabling similarity-based recommendations with real-time emotional tone filtering for personalized user experience.

Log Analysis for AIOps

- Developed an end-to-end anomaly detection pipeline using **pandas** and **numpy**, by parsing unstructured system logs, with features such as **log severity scores** and message length, and preparing **structured inputs** for **ML models**.
- Applied **Isolation Forest** with a **10%** contamination threshold to **automatically** flag suspicious log patterns, enabling **proactive** detection of anomalies, with lesser manual efforts, thus improving **reliability** in **AIOps workflows**.

DockerGen AI

- Developed an intelligent **containerization automation** tool that dynamically generates optimized **DockerFile** configurations for multiple programming languages, streamlining **DevOps** workflows, reducing manual setup time.
- Designed and implemented a **dual AI** model architecture integrating cloud-based Google **Generative AI** API with local **Ollama** inference engines, enabling flexible deployment options, seamless scalability, and cost-efficient operations.

COURSES AND CERTIFICATIONS

- Supervised Machine Learning: Regression and Classification , **DeepLearning.AI**
- Mastering Generative AI, **GeeksforGeeks**
- The Complete SQL Bootcamp, **Udemy**
- Large Language Models, **Google Cloud**

TECHNICAL SKILLS

Languages: Python, SQL, PostgreSQL, Java, HTML, CSS

Framework & Libraries: NumPy, Pandas, Matplotlib, Seaborn, Sci-kit, NLTK, SpaCy, LangChain, Ollama

Machine Learning & AI: Supervised Learning (Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, KNN), Natural Language Processing, Semantic Search, Recommender Systems, Transformers, Large Language Models

DevOps & Deployment: Linux, Docker, Docker Swarm, YAML, Git, GitHub