Sumit Shahu

Aspiring Data Scientist | B.E. Data Science | Python • ML • Deep Learning | Azure & Oracle Certified

✉ sumitshahu25@gmail.com   📞 9321150515   in sumit-shahu-558271266   ⌗ sumitshahu

SKILLS

Programming & Scripting: Python, Java, C, Keras, TensorFlow
Machine Learning & AI: Regression, Classification, Clustering, Ensemble Learning, LSTM, GRU, Model Evaluation Metrics, Gradient Descent, Regularization, Deep Learning Concepts, ML Pipeline
Data Analysis & Visualization: Data Cleaning, Feature Engineering, Exploratory Data Analysis (EDA), Pandas, Scikit-learn, Seaborn, NLP, Statistics
Generative AI & LLM Tools: LangChain, Prompt Engineering, Vector Databases (conceptual), RAG
Cloud & Tools: Azure Cloud Fundamentals, Azure AI Services, Azure Data Factory, SQL, Git

INTERNSHIPS

- Completed AICTE Virtual Internship in Data Analytics through AWS Academy
- Completed AICTE Virtual Internship in Cybersecurity through Zscaler

EDUCATION

B.E, Data Science | AP Shah Institute of Technology, Thane(Mumbai University)

CGPA: 8.4 | (July 2021-May 2025)   July 2021 - May 2025
XII (HSC) | K)V Pendharkar College, Dombivli (Mumbai University)

86.83%| (June 2020 -May 2021) - 86.83%   June 2020 - May 2021

PROJECTS

Sentiment-Infused Stock Price Prediction | LSTM + GRU Ensemble ↗

- Built a hybrid deep learning model combining LSTM & GRU for stock price forecasting, improving accuracy through ensemble learning.
- Incorporated financial news sentiment analysis using BERT to capture market behavior and sentiment influence on stock movement.
- Predicted the next-day closing price range and generated sentiment-driven insights by integrating financial data and news headlines.
- Designed an automated pipeline using Python, TensorFlow, Pandas, yfinance, NewsAPI, with interactive prediction visualizations in Django.
- Key Skills: Deep Learning, Sentiment Analysis, Time-Series Forecasting, Ensemble Models, Django, API Integration, Data Engineering

Movie Recommendation System | Content + Collaborative Filtering ↗

- Developed a recommendation system using content-based filtering and KNN-based collaborative filtering models.
- Analyzed user behavior and movie metadata to generate personalized recommendations.
- Implemented UI + backend using Python, Django, optimized using vector-based similarity and matrix techniques.
- Key Skills: Recommendation Systems,Python, Django, KNN, Data Processing

Olympics Data Analytics Dashboard (1896–2016) | Streamlit

- Performed EDA on 120 years of Olympic data to identify medal trends, athlete performance, and country-wise insights.
- Built an interactive analytics dashboard using Streamlit with dynamic charts, filters, and search features.
- Used Python, Pandas, Matplotlib/Seaborn, and hosted dashboard for real-time exploration.
- Key Skills: Data Analysis, Data Visualization, Streamlit, Pandas, Plotting Libraries, Dashboard Development

CERTIFICATES / ACHIEVEMENTS

Microsoft Certified: Azure AI Engineer Associate (AI 102) ↗

Oracle Cloud Infrastructure 2025 Certified Data Science Professional ↗

Oracle Cloud Infrastructure 2025 Certified Generative AI Professional ↗

Published Research Paper: Sentiment-Infused Stock Price Prediction: A Hybrid Model for Capturing Market Sentiment and Financial Trends — IEEE