

# TUMBALI VENKATA VISWAS

+91 9121874235 — tumbalivenkataviswas@gmail.com — VenkataViswas — VenkataViswas — Leetcode : VenkataViswas

## Experience

### PowerBI Intern - Dosoft Technologies

May 2024 - June 2024

- Developed and designed interactive dashboards and visualizations using tools like Power BI.
- Analyzed large datasets to extract actionable insights and presented findings to stakeholders.

## Projects

### Retrieval-Augmented Generation (RAG) System

- Designed and implemented an **end-to-end RAG pipeline** to ingest multi-format documents (PDF, TXT, CSV, Excel, Word, JSON) and answer user queries intelligently.
- Utilized **Sentence Transformers (all-MiniLM-L6-v2)** to generate dense vector embeddings and stored them efficiently using **FAISS** for fast similarity search.
- Integrated **Groq LLMs (gemma2-9b-it)** for high-speed inference, contextual question answering, and summarization.
- Built a **modular and maintainable architecture** using LangChain community loaders, enabling easy extension and scalable document ingestion. .

### Sentiment Analysis using Custom Avg Word2Vec and Logistic Regression

- Developed a **sentiment analysis pipeline** using a **custom-trained Word2Vec model** built with **Gensim** and a **Logistic Regression** classifier.
- Implemented comprehensive **text preprocessing** including tokenization, lowercasing, and stopword removal to enhance data quality.
- Trained **custom Word2Vec embeddings** from the corpus and generated sentence vectors using **average word2vec model**.
- Achieved an overall **accuracy of 81.75%** with balanced precision and recall across sentiment classes.
- Utilized **Gensim**, **NumPy**, and **Scikit-learn** to build a modular, interpretable, and reproducible NLP workflow.

### From Raw Data to Deployment: An End-to-End ML Project

- Built a complete ML pipeline from data ingestion, cleaning, and feature engineering to model training and evaluation.
- Performing exploratory data analysis to identify patterns and insights that drive the selection of characteristics.
- Utilized **Scikit-learn** and **Pandas** for pipeline implementation, achieving reproducibility and modularity.
- Designed a clean project structure with **logging**, configuration management, and CI/CD readiness.
- CI / CD implementation using GitHub Actions with **custom AWS EC2 runners** for efficient pipeline execution alongside standard runners.

### Real Estate Price Prediction Web Application

- Developed a **machine learning-based real estate price prediction system** to estimate property values using historical housing data and regression techniques.
- Designed a **Flask-based backend API** to serve predictions, acting as middleware between the trained ML model and a responsive HTML frontend.
- Implemented end-to-end deployment on **AWS**, enabling scalable access to the prediction service in a production-like environment.

### Restaurant Website - MERN Stack

- Built a dynamic **food ordering website** featuring an intuitive menu with smooth scrolling functionality.
- Integrated **secure user login** functionality, a section to view past orders, a food rating system, and an admin page for comprehensive site management.
- Leveraged HTML, CSS, JavaScript, and Node.js to develop a **user-friendly platform** for efficient food ordering and management.

## Skills

- Machine Learning, Deep Learning, Natural Language Processing (NLP), Generative AI
- Retrieval-Augmented Generation (RAG), Information Retrieval, Semantic Search
- Model Training, Evaluation, and Optimization
- Model Deployment, API Development, MLOps
- Data Analysis, Exploratory Data Analysis (EDA), Feature Engineering
- Strong analytical thinking, problem-solving, and adaptability

## Tools and Technologies

---

- **Programming Languages:** Python, Java, C++ , JavaScript
- **Machine Learning & NLP:** Scikit-learn, TensorFlow / Keras, NLTK, Gensim, NumPy, Pandas
- **LLM & RAG Stack:** LangChain (Community Loaders, Retrievers, Chains), Sentence Transformers , HuggingFace, FAISS, Groq LLMs
- **Data Visualization & Analysis:** Matplotlib, Seaborn, Microsoft Power BI
- **Web Development & APIs:** Flask (ML & LLM APIs), HTML, CSS, JavaScript, Node.js, Express.js, MERN Stack
- **Cloud & Deployment:** AWS EC2, Model & Web Application Hosting, Linux-based Cloud Servers
- **DevOps & CI/CD:** Docker, Kubernetes, Jenkins, GitHub Actions, ArgoCD
- **Version Control & Collaboration:** Git, GitHub
- **Operating Systems:** Linux, Ubuntu, Windows

## Education

---

**Master of Technology in CSE DS.** 2025 - 2027

- Gandhi Institute of Technology and Management
- CGPA : Not Yet Completed Course

**Bachelor of Technology in CSE DS.** 2021 - 2025

- Anil Neerukonda Institute of Technology and Sciences
- Affiliated to Andhra University, Visakhapatnam
- CGPA : 8.95 / 10

**Intermediate - MPC** 2019 - 2021

- Tirumala Junior College, R Tallavalasa, Visakhapatnam, Andhra Pradesh
- Marks : 958 / 1000

**SSC Class - X** 2018 - 2019

- SFS High School, Visakhapatnam
- CGPA : 8.8 / 10

## Certifications and Achievements

---

- Participated in the Seminar on Secure Software Development Life Cycle practices
- Achieved Top 5 position in a Web Design Competition Utilizing Figma .
- Competed in the Internal Smart India Hackathon (SIH) and achieved 15th place among participating teams.
- Introduction to AI - 365 Data Science
- Introduction to ChatGPT and Generative AI - 365 Data Science
- Introduction to IOT - NPTEL 86%
- Organized and Conducted Project Expo event 2023-24
- 2-Day workshop on Problem Solving with Design Thinking