

# Gulshan Kumar

Machine learning engineer

📞 9103091562    ✉️ gulshankumar100g@gmail.com    🔗 linkedin.com/in/gulshan100    📍 Jammu, India  
🐙 github.com/gulshan-100    🔗 https://gulshan100.netlify.app/

## PROJECTS

### RAG based Chatbot - Medicare 🔗

Developed an AI driven chatbot leveraging Retrieval Augmented Generation (RAG) and the Gemini API to handle inquiries efficiently. Enhanced user experience with intuitive, automated responses and real time assistance.

### AI Powered Calorie Burn Prediction and Management Tool 🔗

Developed a Flask web application with XGBoost for predicting exercise calorie burn from a Kaggle dataset. Integrated a Bootstrap-based UI, containerized the application using Docker, and implemented CI/CD pipelines with CircleCI.

### Synthetic Data Generation using NLP 🔗

- **Data Preprocessing:** Cleaned and enhanced a dataset of 16,669 product reviews using standard scaling, text lemmatization, and word count analysis.
- **Model Development:** Created a custom deep learning model with attention mechanisms and Bidirectional LSTM layers to generate coherent product reviews.
- **Review Generation:** Implemented the model to generate high-quality product reviews.

### Fine tuned T5 Model for Text Summarization 🔗

Implemented and trained a T5 small model on the CNN/DailyMail dataset for text summarization tasks, utilizing the Hugging Face Transformers library and PyTorch, with custom data preprocessing.

### Plant Disease Prediction using Convolution Neural Networks

Designed a CNN model to classify plant diseases from leaf images and built a web app for instant diagnosis via image uploads.

### LLM Voicebot 🔗

Developed an LLM Voicebot using Gemini, AssemblyAI, and TTS to transcribe user voice, generate responses with LLMs, and deliver synthesized speech. Built the backend with Flask and the frontend with HTML, CSS, and JavaScript.

### Video Transcription and Sentiment Analysis using Assembly AI API 🔗

Built a web application with AssemblyAI API for video transcription, timestamping, and sentiment analysis. Used HTML, CSS, and JavaScript for the frontend and Flask for the backend.

## SKILLS

### Data science, Machine learning and Deep learning

Python, Supervised and unsupervised algorithms, Neural Networks, ANN, CNN, RNN, LSTM, Feature engineering, NLP, Transformers, Fine Tuning etc.

### Python Packages and Frameworks

Scikit-learn, Tensorflow, Keras, Numpy, Pandas, Matplotlib, NLTK, OpenCV, Langchain, Streamlit, etc.

### Languages and Databases

HTML, CSS, Javascript, PostgreSQL

### Generative AI

GPT, RAG, Gemini, Pinecone, Hugging Face, Langchain, Fine Tuning, Agents

### MLOPS

GCP, Flask, Git, Docker, CI CD Github Actions

## EXPERIENCE

### Data Science Intern

*Zidio Development*

02/2024 – 03/2024

- Worked on data science projects focusing on regression and classification problems.
- Applied feature engineering and data cleaning techniques to prepare data for model development.

## ORGANIZATIONS

### Chegg India

*Subject Matter Expert*

05/2023 – present

Part time Subject Matter Expert at Chegg, solved over 350 probability and statistics questions.

### GeeksforGeeks

*Technical writer*

01/2023 – 05/2023

Authored 23 clear technical articles for GFG.

## CERTIFICATIONS

Machine learning specialization

Deep learning specialization

Langchain for LLM Application development

SQL for Data Analysis - Internshala

NLP Specialization