

Sahana Khanai

Generative AI Engineer

[LinkedIn](#) | 9353048023 | sahana.khanai.icer.cs061@gmail.com | Mumbai, India 400601 | [GitHub](#)

Objective:

Innovative and technically versatile Engineer with 2 years of experience in full-stack development, now building intelligent applications using Generative AI tools such as LLMs, LangChain, Hugging Face, and RAG pipelines. Skilled in integrating vector databases like FAISS and ChromaDB, and deploying Transformer-based models on cloud platforms including AWS SageMaker and Bedrock. Combines strong backend expertise in Spring Boot, API development, and CI/CD with deep learning, prompt engineering, and modern AI frameworks to deliver scalable, production-ready AI solutions in fast-paced environments.

Education:

- Bachelor of Engineering | Computer Science and Engineering** August 2019 – May 2023
Visvesvaraya Technological University – Belagavi, **CGPA: 9.14**
 - Recipient of **Gold Medal** 2023

Skills:

- Generative AI & NLP** - Large Language Models (LLMs) – OpenAI GPT, Llama2, Mistral, Gemini Pro; LangChain, Hugging Face, Transformers, RAG Pipelines (Retrieval-Augmented Generation), PEFT, Prompt Engineering, Instruction Tuning, Fine-tuning with custom datasets, LangSmith, LangServe, LlamaIndex, Ollama, NLP Pipelines, NLTK, word2vec, Text Embedding, Conversational AI, Foundation Models, NVIDIA NIM, CREW AI
 - Machine Learning & Deep Learning** - TensorFlow, PyTorch, Keras, Scikit-learn, ANN, RNN, LSTM, Deep Learning Pipelines, Model Inference, Model Evaluation, Neural Network Optimization, Transfer Learning, OpenAI API
 - Cloud Platforms, MLOps & DevOps** - Amazon SageMaker, AWS Bedrock, Databricks, Docker, Jenkins, Git, CI/CD Pipelines, SonarQube, Unix, Git.
 - Full-Stack Development & API Engineering** - Spring Boot, Spring Framework, Hibernate ORM, RESTful APIs, Servlets, JSP, JDBC, Streamlit, OpenAPI (Swagger), Postman, Backend-Frontend Integration
 - Frontend** - React.js, JavaScript, HTML, CSS, Bootstrap
 - Databases** - PostgreSQL, MySQL, Apache Cassandra, Astra DB, Chroma DB, FAISS, Query Optimization, Vector Stores
 - Programming** - Python, Java, JavaScript, SQL
 - Soft Skills** - Analytical Thinking, Creative Problem-Solving, Effective Communication, Team Collaboration, Adaptability, Fast Learning, Initiative, Time Management, Curiosity, Attention to Detail
-

Experience:

Tata Consultancy Services (TCS)

Software Engineer – Full Stack Development & Generative AI

Mumbai, India | Feb 2024 – Present

- Developed secure, scalable APIs using Java, Spring Boot, Hibernate, and JWT, improving system response time and maintainability.
- Built dynamic, responsive UIs using JSP, HTML, CSS, and JavaScript, enhancing user experience and engagement.
- Automated CI/CD workflows with Docker, Jenkins, and SonarQube, reducing deployment time by 35%.
- Provided L2/L3 production support and resolved incidents, improving system uptime and performance.
- As part of internal initiatives and upskilling, explored Generative AI tools such as LangChain, OpenAI GPT, FAISS, and Hugging Face Transformers.
- Built a working prototype of a RAG-based chatbot using Python, integrating vector search for contextual retrieval.

- Studied and applied Prompt Engineering, LLM integration, and vector database concepts in self-driven projects using Python and LangChain.

NSREEM

AI Developer – Full Stack & ML Prototyping

Remote | Feb 2023 – May 2023

- Created interactive, data-driven dashboards using Streamlit and Python, improving user engagement and system insights.
- Developed K-Means clustering models for customer segmentation and integrated them into Flask-based APIs
- Implemented basic sentiment analysis for customer feedback, using NLTK and Scikit-learn.
- Prototyped a content-based recommendation engine and explored GenAI integrations using OpenAI APIs and vector similarity.

Knowx Innovations Pvt. Ltd.

AI/ML Engineer – Applied Machine Learning Projects

Bengaluru, India | Aug 2022 – Sep 2022

- Built a secure Password Manager application using Python, Tkinter, and MySQL, applying encryption and GUI principles.
- Developed a fraud detection system using a Random Forest classifier, achieving 95% accuracy on financial datasets.
- Preprocessed datasets, conducted feature selection, and trained models using Scikit-learn, improving model performance metrics.
- Explored NLP techniques and began implementing text classification workflows for internal automation use cases.

Projects:

1. ThreatLens – GenAI-Powered Cyber-Threat Assistant

Streamlit | OpenRouter API | LLMs | Python

Built a multi-tab web assistant that harnesses Large Language Models to streamline security analysts' workflows.

- Phishing Analyst: flags suspicious emails, extracts IOCs, and suggests remediation steps.
- CVE Explainer: translates CVE details into plain English, outlining impact and mitigation.
- Log Summarizer: condenses lengthy security logs into actionable insights for faster triage.
- Integrates OpenRouter LLM endpoints within an intuitive Streamlit UI to deliver contextual, on-demand threat intelligence.
- Accelerates incident response and reduces analyst workload by providing cohesive AI support in one place.

2. SOP-Genius – RAG-Based SOP Knowledge Assistant

LLMs | FAISS | RAG | Streamlit | Python

Developed a smart Q&A assistant that transforms static cybersecurity SOPs into an interactive, searchable knowledge system using Retrieval-Augmented Generation (RAG).

- Enables users to upload or query SOP documents and get precise, context-aware answers powered by Large Language Models.
- Utilizes FAISS for efficient vector search and document retrieval.
- Built with a clean Streamlit interface for seamless user interaction.
- Helps cybersecurity teams quickly access relevant procedures, enhancing operational efficiency and decision-making.

3. Customer Churn Classifier

Streamlit | ANN | TensorFlow | Keras | Python

Developed an interactive web application to predict customer churn using an Artificial Neural Network (ANN).

Enabled businesses to proactively retain customers by accurately identifying churn risk, with a clean UI built using Streamlit.

- Achieved >85% prediction accuracy on test data.
- Enabled proactive customer retention strategies.
- Showcased end-to-end ML lifecycle: preprocessing, training, evaluation, and deployment.

4. Movie Review Sentiment Detector

Streamlit | NLP | RNN | Python | TensorFlow

Built a sentiment analysis tool using a Simple RNN to classify movie reviews as positive or negative. Demonstrated the use of deep learning for natural language understanding with interactive web UI.

- Achieved 90% sentiment classification accuracy.
- Visualized predictions and model confidence in real-time.
- Applied tokenization, embedding layers, and sequential modeling.

5. Next Word Predictor

Streamlit | LSTM | NLP | TensorFlow | Keras

Engineered an intelligent word prediction tool using LSTM networks to suggest the next probable word based on previous input, simulating natural typing behavior.

- Demonstrated sequence modeling and text generation capabilities.
- Improved language understanding using pretrained embeddings.
- Built a performant backend optimized for sequential data.

6. AI Picture Translator

TensorFlow | Computer Vision | OCR | Multilingual AI

Created a real-time AI-driven translator app that detects objects and text from images using TensorFlow and translates them into selected languages.

- Achieved 90% object/text recognition accuracy.
 - Boosted user engagement by 25% through interactive visual AI.
 - Combined **vision + NLP** for multilingual application.
-

Certifications:

- Complete Generative AI Course with Langchain and Huggingface
 - Career Essentials in Generative AI – Microsoft & LinkedIn
 - Spring Security 6 & JWT, OAuth2 – Udemy
 - Master Java Web Services & REST API – Udemy
 - Introduction to OpenAI & ChatGPT API – Udemy
 - Generative AI Foundation Curriculum – TCS
 - SQL for Data Science – University of California, Coursera
 - Machine Learning for All – University of London, Coursera
 - Python Data Structures – University of Michigan, Coursera
 - Web Development Bootcamp – Udemy
 - Cybersecurity Awareness & Foundations – TCS & Forage
-

Achievements & Awards:

- Achieved **Gold Medal in the B.E. Computer Science department**.
- Recognized as the **Best Student** in high school and PUC.
- Secured **1st rank** in Computer Competition and SSLC exam, earning accolades from the high school.
- **5 Star** in Python programming in Hackrank, **Top performer** in ILP TCS training program.
- In TCS received **Beyond Performance Awards, Xcelerate Warrior Certificate**.
- **Special Initiative Award** from TCS | June 2025
- Recognized for outstanding contributions and proactive initiatives that positively impacted the team and organization. Appreciated as a role model for dedication and commitment.