Uvais Saifi

uvaissaifi523@gmail.com | 9354038020 | India, New Delhi | Portfolio | Linkedin | Github

PROFILE

Highly motivated Data Scientist and aspiring AI/ML Engineer specializing in Natural Language Processing (NLP) and Generative AI. Proficient in Python, TensorFlow, and Google Cloud, I am eager to apply my skills to innovative, data-driven projects.

SKILLS

• Programming: Python, SQL, Bash, JavaScript, HTML

NLP: NLTK, TF-IDF, Word2Vec, GloVe, Transformers, Hugging Face, Sentiment Analysis, Text Classification, RAG, Gemini AI, Prompt Engineering

• Tools: Jupyter Notebook, Google Colab, Git, GitHub, Streamlit, Flask, REST APIs, JSON

• ML/DL: Scikit-learn, TensorFlow, Keras, PyTorch, CNN, LSTM, ANN, XGBoost, LightGBM

• Data: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Data Visualization, EDA, Statistics

• Cloud: Google Cloud (Vertex AI), AWS (Basics), Heroku

PROJECTS

Mental Health AI App
- Developed a holistic Mental Health AI App using Python and Kivy that integrates CBT, Law of Attraction, and guided reframing, helping users improve mindset, emotional balance, and mental well-being.
- Implemented a Retrieval-Augmented Generation (RAG) chatbot powered by embeddings and Gemini API, enabling users to ask complex mental health queries and receive contextual, document-based responses.
- Designed visualization and imagery features with Black Forest image generation API, allowing users to create personalized affirmations and goal-oriented visuals for therapeutic and motivational practices.
- Integrated clinical tests (PHQ-9, GAD-7, DASS-21, Stroop, Cognitive Bias) along with journaling, affirmations, and voice interaction, combining ML pipelines with an interactive interface for personalized feedback.

NLP Query Engine for Employee Data
- Built NLP query engine using sentence embeddings (all-MiniLM-L6-v2) and FAISS vector indexing for semantic search across employee documents with sub-2-second response times.
- Developed dynamic schema discovery system with SQLAlchemy to auto-generate SQL from natural language without hard-coded dependencies across multiple database structures.
- Implemented hybrid AI search combining vector similarity, keyword matching, and entity recognition with confidence-based filtering for structured and unstructured data.
- Designed ML pipeline with query caching and batch embedding generation handling concurrent requests while maintaining production-grade performance.

AthleteRise
- Developed AthleteRise, an AI-driven cricket analytics system using Python, OpenCV, and MediaPipe, performing real-time pose estimation, biomechanical analysis, and motion tracking to evaluate batting techniques and optimize player performance.
- Implemented MediaPipe pose detection with custom pipelines to extract key metrics including bat speed, swing angles, joint trajectories, and timing from videos, leveraging computer vision for actionable AI/ML-driven insights.
- Generated annotated videos with pose skeleton overlays for instant visual feedback to players and coaches.
- Created performance reports with scores and charts for tracking player progress and data-driven decisions.

Flexible Emotion Detector
- Developed **Flexible Emotion Detector**, a multi-modal system using **CNN** for facial emotion, LSTM for audio via MFCC, and a pre-trained **NLP model** for text, detecting emotions from video, audio, or text inputs.
- Implemented a Streamlit-based interactive interface allowing users to upload video, audio, or text inputs and receive real-time emotion predictions, improving accessibility and engagement.
- Optimized models with data augmentation and hyperparameter tuning for improved accuracy across all input types.
- Added logging and error handling, ensuring reliability and smooth user experience during emotion detection

Anime Recommendation System End-to-End
- Trained an Anime Recommendation System using cosine similarity and TF-IDF vectorization to suggest relevant anime titles.
- Built a Flask-based web interface for real-time recommendations, allowing users to input anime names and get suggestion.
- Processed data with pandas and trained vector models using scikit-learn, optimizing system performance for recommendations.

CERTIFICATES

- Python with DS - AEDIFICO
- Data Science Bootcamp-Udemy
- Inspect Rich Documents with Gemini Multimodality & RAG - Google Cloud Skills Boost
- Build Real-World AI Applications with Gemini & Imagen - Google Cloud Skills Boost
- Machine Learning A-Z™ (Hands-On Python & R)- Udemy
- Generative AI with Large Language Models-Coursera
- Prompt Design in Vertex AI - Google Cloud Skills Boost
- Develop GenAI App with Gemini & Streamlit - Google Cloud Skills Boost
- Managing Big Data with MySQL-Coursera
- Gen AI Academy-Hack2skill
- Explore Generative AI with the Gemini API - Google Cloud Skills Boost
- Natural Language Processing on Google Cloud - Google Cloud Skill Boost

EDUCATION

Bachelor of Computer Applications — 2022 – 2025
IGNOU — New Delhi, India

LANGUAGES

• Japanese (JLPT N4) – Basic Proficiency
• Hindi – Native
• English – Intermediate