# Keval Saud

kevalsaud25@gmail.com Mumbai, MH, IN
in/keval-Saud github.com/Tobaisfire My-PortFolio

## Professional Summary

Experienced in data scienece, Python development, and machine learning. Proficient in extracting insights from complex datasets, building scalable web applications, and implementing machine learning models. Currently pursuing a Master's degree in Computer Science with a focus on advanced data science and machine learning techniques.

## Areas of Expertise

Programming - Data Science - Data Mining - Automation - LangChain - SQL- LLMs - Machine Learning - NLP - Requests - Communication - Problem Solving - Teamwork - Adaptability - Leadership

## Professional Experience

**Research Intern** *Physics Department UoM* **Mumbai, MH, IN** *March 2024 – May 2024*

- Conducted in-depth data analysis on Coronal Mass Ejections (CMEs) using NASA WIND Database of 20 years.
- Developed and maintained Python scripts utilizing Pyspedas for efficient data extraction from the WIND API. Implemented data cleaning algorithms using Pandas, addressing null datasets (CDF) and ensuring data integrity.
- Performed exploratory data analysis (EDA) on multi-year WIND Database datasets to identify key insights and trends.
- Conducted scientific calculations to derive new parameters from WIND data, optimizing logic and time complexity. Created and optimized IP plots and visualizations with Matplotlib to present analytical findings effectively.
- Managed codebase and version control on GitHub, facilitating collaboration and maintaining project structure. Researched and reviewed relevant literature on CMEs and data analysis techniques to support project development.

**Jr. Software Developer,** *SquareYards* **Mumbai, MH, IN** *May 2022 – Jun 2023*

- Automated data extraction and processing (Data/Web-Mining) using Python requests and Selenium, **reducing manual effort by 30%** and **increasing data processing speed by 40%**.
- Led the development and integration of AI chatbots leveraging Langchain, Cohere embeddings, and OpenAI GPT-3.5, **improving customer interaction and satisfaction by 25%.**
- Optimized data storage and retrieval by transferring Cohere embeddings to vector databases (Pinecone and Weviate), **resulting in a 30% increase in efficiency.**

**Freelance Python Developer,** *Freelance* **Self-Employed** *Nov 2021 – May 2022*

- Developed predictive models using Python libraries such as pandas for data cleaning, and utilized Matplotlib for visualization.
- Trained machine learning models for prediction and classification, resulting in an average **accuracy improvement of 15% in predictive models.**
- Worked closely with international clients to understand objectives and deliver customized solutions within agreed timelines.

## Education

**Master of Science in Computer Science** *University of Mumbai* **Mumbai, MH, IN** *Aug 2023 – Aug 2025*
**Bachelor of Science in Computer Science** *University of Mumbai* **Mumbai, MH, IN** *Jan 2018 – Mar 2021*

## Major Projects

**Web Mining** **Python, Web Scrapping, Data pre-processing MongoDB**

- Developed a sophisticated web application leveraging Requests and Selenium methods for automated web scraping and robust database management..
- Included task scheduling and execution, **optimizing automation workflow up to 20% faster**.
- Achieved a significant **50% improvement in work efficiency**, accompanied by a remarkable **70% reduction in manual involvement**, demonstrating advanced automation capabilities.e.

**LLM-CoversationalAI** **Langchain, Cohere, Python, Weviate**

- Integrated Langchain with AI chatbot, using Cohere embeddings.
- Utilized OpenAI GPT-3.5 , Llama by meta, Gemini by Google and open sorce models from HuggingFace for conversational capabilities.
- Transferred Cohere embeddings and implemented a similarity search algorithm on the vector database.

**LoRA Fine Tune - Diffusion Model** **Python, Diffusers, ML**

- Trained a Stable Image GEN Model using LoRA, enhancing training outcomes through improved Image generation.
- Evaluated model training with **30% less consumption of GPU**, guiding future improvements.

**Named Entity Recognition (NLP Model)** **Python, NLP, BI-LSTM**

- Implemented a BI-LSTM model for Named Entity Recognition (NER), achieving 98% accuracy.
- Preprocessed data with NumPy and Pandas. Trained the model on labeled data, optimizing hyperparameters.
- Evaluated performance using precision (97%), recall (96%), and F1-score (97%). Visualized with Matplotlib and Seaborn.

## Skills

- **Technical Skills:** Programming, Data Science, Data Mining, Automation
- **Frameworks:** LangChain, Django, Flask, Streamlit, Gradio
- **Databases:** MongoDB, Cassandra, Redis, Hbase, MYSQL
- **ML-AI:** LLMs, ML, Data Science, NLP, Diffusion
- **Libraries:** Selenium, Requests, OCR, Pandas, Matplotlib, Scikit-Learn, NumPy, Computer Vision
- **Languages:** Python, SQL, JavaScript, C++, Java
- **Typesetting:** LaTeX
- **Soft Skills:** Communication, Problem Solving, Teamwork, Adaptability, Leadership

## Certifications

- **ML Workshop at IITB** : Completed an intensive workshop conducted by esteemed faculty at IIT Bombay.
- Scientific Computing [Python]: Completed Advance python Skills by **FreeCodeCamp.org**
- Data Science 101 [Python]: Completed Basic of DS using Python.

## Tools and Utilities Used

- **LLMs:** Gemini, GPT, Falcon, MPT-30B, Claude, Llama, Phi, Coral, Commad-R, Mistral
- **Dev Tools:** HuggingFace, Kaggle, Visual Studio Code, Git, GitLab, PyCharm, Visual Studio, NetBeans, Anaconda, Cohere

## Declaration

I do hereby declare that all statements given by me as above are true, complete, and correct to the best of my knowledge and belief.