

Sofi Altamsh Shah

Email: sofialtamsh123@gmail.com | Phone: +91-9156704982 | Bengaluru, Karnataka
<https://www.linkedin.com/in/sofi-altamsh-409831193/> | <https://github.com/sofialtamsh>

Professional Summary

NLP Engineer with over 1 year of hands-on experience building and deploying LLM pipelines using LangChain, RAG and Hugging Face. Specialized in semantic search, AI agents and generative text systems with proven production deployments in EdTech. Reduced response time in QnA systems by 92% and drove 23% uplift in hiring decisions using predictive ML models

Technical Skills

- Machine Learning: Regression, Clustering, **PCA**, XGBoost, Preprocessing, Feature Engineering, Hyperparameter Tuning
- Deep Learning: CNN, RNN, GANs, Transformers (RoBERTa, T5, GPT), **LoRA**
- NLP & LLMs: Text2Text, **RAG**, TF-IDF, Word2Vec, **NLTK**, NER, Text Summarization, Hugging Face, **LangChain**
- MLOps: Streamlit, Flask, HDFS, AWS Lambda, Vector Databases (FAISS), **MySQL**, MongoDB
- Libraries & Frameworks: NumPy, Pandas, **TensorFlow**, PyTorch, **Scikit-learn**, SciPy, OpenCV, OCR
- Visualization: **Power BI**, Matplotlib, Seaborn, **Excel**, Metabase

Experience

Natural Language Processing Engineer | Masai, Bengaluru | July 2025 – Present

- Automating ticket resolution processes with LLMs using LangChain-based RAG pipelines and AI agent frameworks.

Machine Learning Engineer | Masai, Bengaluru | Jan 2025 – July 2025

- **Reduced** user query response time by **92% (to less than 3.2 seconds)** by building an Answering System using **RAG** pipeline
- **Increased 23% hiring rate** of Masai students by analysing their performance and developing a **CTC Prediction Model**
- Trained ML and DL models then deployed them in production. Analyzed data trends and created visualizations on Metabase

Data Research Analyst | Masai, Bengaluru | Oct 2024 – Jan 2025

- Worked on Data pipeline, **Top Hiring Companies** Analysis project and delivered insights to increase placement rate by 15%.

Data Analyst Associate | Masai, Bengaluru | Apr 2024 – Oct 2024

- Analyzed job market data across LinkedIn, Naukri, and Instahyre to extract hiring trends.
- Supported the data team in structuring and cleaning large-scale recruitment datasets.

Data Science Intern | IStop.ai, Bengaluru, Karnataka | Jun 2022–Nov 2022

- Trained predictive models for house loan value with more than **90% accuracy**
- Enhanced performance using **PCA**, Visualized clusters & created Dashboards

Projects

- **Fake News Detection Using Multi-Modal Learning**
 - Built a hybrid model combining RoBERTa for text and CNNs for image inputs to detect fake news articles
 - Achieved 88.7% accuracy on a custom multimodal dataset of 10,000+ samples.
 - Tech Stack: Python, Transformers (Hugging Face), RoBERTa, CNN, PyTorch, NumPy
- **Face Mask Detection System**
 - Built a real-time system using TensorFlow and **OpenCV** achieving **94.6% accuracy** trained with 1500+ images
 - Tech Stack: Python, TensorFlow, OpenCV, Keras, NumPy

Education

- **B.Tech - Computer Science Engineering**
Government College of Engineering, Aurangabad | 2019-2023 | CGPA: 8.1