Sameena **Tasneem**

RESEARCHER & DATA SCIENTIST

*M.Sc. (Mathematics) | M.Tech. (M&S, Major Machine Learning)*



# Profile Summary

AI/ML Engineer with (14 years total) and, 7+ years of experience delivering production-grade systems across startups and high-growth environments. I specialize in building real-time, scalable solutions using LLMs, multi- agent systems, and multimodal models. From fine-tuning open-source models (TinyLlama, Mistral, FlanT5) to deploying fast, reliable inference pipelines with FastAPI, Docker, and cloud platforms (AWS/GCP), I take ownership of the full ML lifecycle. I’ve built internal tools, automated support workflows, and designed intel- ligent systems that integrate context-rich search, RAG, and prompt engineering. I move fast, love ambiguous challenges, and thrive in teams that prioritize velocity, impact, and integrity

# Technical Skills

* **ML & Deep Learning Frameworks:** PyTorch, Hugging Face Transformers, scikit-learn, Spacy, NLTK, matplot, LangChain, LlamaIndex, crewai
* **NLP & LLM Fine-Tuning:** LoRA, DPO, SFT, Embedding Extraction, Prompt Engineering, Retrieval-Augmented Generation (RAG), multiagent
* **Experimentation & Evaluation:** ROUGE, PR-AUC, NDCG, MMR, Precision/Recall, TensorBoard,
* **Multimodal & Vision Models:** BLIP-2, BLIP-3, Vision-Language Embeddings, Fusion Architectures
* **Data Tools & Pipelines:** Pandas, Numpy, PySpark, SQL, time-series analysis, matplotlib, Streamlit.Fastapi.
* **Cloud & Infrastructure (Basic):** AWS (S3, SageMaker, Athena), GCP (BigQuery, Cloud Storage)
* **Programming Languages:** Python, R

# Relevant Experience

## Network-based engagements *Remote*

**INDEPENDENt AI/LLM CONSultING PROJECtS** *Feb 2025 - June 2025*

* Collaborated on multiple AI/LLM initiatives with startup teams and personal network :
* Developed a RAG pipeline to parse and query PDFs, Excel, CSVs, and TXT files, including table extraction, QA, summarization, and translation
* Built a multi-agent system using LangChain for text-to-SQL, and tool use. Applied few-shot prompting and chain-of-thought reasoning for improved task completion.
* Built a CrewAI-based multi-agent helpdesk assistant for contextual query handling and internal documentation support.
* Assisted a startup in prompt engineering and LLM integration (Mistral-based) for a mental health consultation platform

**course hero** *Remote*

**data SCIENtISt-ML** *August 2022 - November 2024*

* Designed and evaluated a multimodal document quality classifier using BLIP-3 vision-language embeddings and neural head layers. Conducted 8+ structured ablations across fusion strategies, achieving 79% PR-AUC, outperforming baseline by 24%.
* Fine-tuned TinyLlama using SFT, for educational keyword generation. Achieving 75% ROUGE-1/L and strong alignment with human evaluation.
* Built RAG QA systems for long-form text and video-based conversations, integrating LangChain, vector stores, and OpenAI APIs. Tuned re-ranking modules using NDCG, MMR, and precisionk to maximize relevance.
* Developed a multimodal traffic forecasting model using BLIP-2/3, integrating image, text, tabular, and geospatial features. Achieved a 40% improvement over baseline accuracy through embedding fusion techniques and architectural optimization.
* Trained a FlanT5-based toxic content classifier with frozen decoder for efficient inference. Achieved 83% accuracy across 11 toxicity categories; evaluated effects of label granularity and alternate loss functions.
* Fine-tuned LLaMA and Mistral models on domain-specific data, comparing SFT, LoRA, and DPO. Integrated training insights with Ten- sorBoard, and monitored performance via Grafana dashboards and Slack alerts.
* Developed contextual response generation models using BERT and OpenAI embeddings to support conversational agents. Bench- marked models using semantic similarity, ROUGE, and human evaluation metrics.
* Designed and deployed a real-time anomaly detection system using time series models, with automated alerts via Grafana and Slack. BuiltFastapi-based inference services and deployed on GCP using Docker and Kubernetes.
* Collaborated cross-functionally with product and design to frame ambiguous product questions as ML tasks and interpret outcomes for user-facing features.
* Optimized LLM inference pipelines using quantization and knowledge distillation techniques, significantly reducing latency and mem- ory usage in production deployments without compromising model performance
* Managed full model lifecycle—including training, CI/CD-integrated deployment using FastAPI, Docker, and Kubernetes, and real-time monitoring with TensorBoard, Grafana, and Slack alerts

**IML** *Pune*

**data SCIENtISt-INDIvIDual CONtRIButOR** *July 2018 - July 2022*

* Developed topic modeling using LDA + KMeans clustering, improving clustering accuracy by 30%.
* Created a Python library for ML tasks (classification, regression, visualization) using scikit-learn for non-ML experts.
* Built an ensemble-based classification system (Random Forest, XGBoost, AdaBoost).
* Designed text de-duplication pipelines using Jaccard similarity.
* Developed an automated email analytics system to deliver weekly insights on user behavior and product performance. Leveraged Python, Pandas, and SQL to extract, analyze, and visualize data, generating dynamic plots and summaries. The system delivered actionable reports to sales, product managers, and other stakeholders, enhancing data-driven decision-making across teams.
* Collaborated with research scientists, ML engineers, and domain experts to develop, experiment, and optimize AI models for real- world applications.

**Exela-Rule14** *Pune*

**data SCIENtISt** *August 2017 - June 2018*

* Worked on text extraction and NLP problems. Developed models using regex, NLTK, and spaCy to extract structured features from raw text and generate actionable insights
* Led internal ML knowledge-sharing sessions, providing clear insights and practical guidance on translating research ideas into production-level code.

## CSIR-National Chemical Lab, Pune *NCL-Pune*

**RESEaRCH INtERN** *May 2016 - August 2017*

* Successfully built a predictive model to predict onion price, which can also be used to predict other commodity prices.
* Modified the model ARGO from one research paper and our modified model performed better than the suggested one giving more accurate prediction for Pune, Nagpur and Mumbai region. Applied different feature scaling and feature engineering techniques.
* Mentored junior engineers in ML best practices and effectively communicated concepts and results to non-ML audiences through presentations.

## Centre for Modeling and Simulation, SPPU-Pune *CMS-Pune*

**RESEaRCH StuDENt** *Sep 2015 - May 2016*

* successfully built a predictive model for epidemic out break.
* Developed custom codes for evolutionary algorithms (genetic, ant colony optimization, firefly)
* Participated in kaggle solved different problems like loan prediction and titanic survivor classification using R and python.

# Experience

**Rajarshi Shahu College of Engineering Buldana** *RSCE-Buldhana*

ASSISTANT PROFESSOR, ADJUNCT RESEARCHER *July 2009 - August 2015*

* Collaborated in a college innovative lab, working on different papers focused on differential equation and numerical analysis. Men- tored 5-7 engineering teams.
* Instructed Engineering graduate students in subject like Engineering mathematics, Calculus, Numerical Analysis and operation re- search.

# Education

## PUNE UNIVERSITY(Centre For Modeling and Simulation) *Pune, India*

M.TECH. IN MATHEMATICAL MODELING AND SIMULATION

With major in Machine Learning *june 2015- July 2017*

* 3.8/5.00 CGPA

**SGBA** *Amravati,MH*

M.SC. MATHEMATICS *june 2006-july 2008*

* Percentage: 60.90 %

**SGBA** *Amravati,MH*

B.ED. MATHEMATICS *june 2005 - march 2006*

* Percentage: 74 %

**SGBA** *Amravati,MH*

B.SC. MATHEMATICS *june 2002 - june 2004*

* Percentage: 69.11 final year %

# Talks and Presentations

* ”Onion price prediction” at CMS-Pune.
* ”Epidemic outbreak prediction using google trend data” at CMS-Pune.
* ”Statistical Approach to Clustering in Pattern Recognition.” at CMS-Pune.
* ”Advance in the Perihelion of Mercury” at RLT-College Akola.
* ”Applications of differential equation in real life” SSGMCE-Shegaon.
* ”Laplace Transform” at CMS-Pune.