

Utkarsh Mishra

utkarshmishrabada@gmail.com
+91 93215 51093

linkedin.com/in/utkarsh-mishra-1aa340227
github.com/Utkarshmishra2k2

Results driven AI professional with hands-on experience building production grade systems at government and tech organizations. Specialized in LLM architecture, RAG pipelines, statistical modeling, and full stack development. Proven ability to translate complex ML concepts into scalable, user facing products with measurable impact on performance, accuracy, and business outcomes.

Skills

- **Language:** Python, R, SQL, Linux OS
 - **Machine Learning:** Supervised Learning (Regression, Classification, SVM, k-NN), Unsupervised Learning (Clustering, Factor Analysis), Tree based Methods (Decision Trees, Random Forests), Ensemble Methods, Feature Engineering, Regularization (L1/L2), Hyperparameter Tuning
 - **Generative AI & LLMs:** Prompt Engineering, RAG, Finetuning, LangChain, LangGraph, Transformer Models, Ollama, Vector Databases, Embedding Models, Hallucination Mitigation
 - **Statistics:** Hypothesis Testing, Sampling, Distribution Theory, Regression, Classification, Factor Analysis
 - **Data & ML Tools:** Pandas, NumPy, Scikit learn, TensorFlow, Seaborn, Power BI, Matplotlib
 - **Backend & Full Stack:** FastAPI, Python web services, JWT authentication, REST APIs
-

Professional Experience

International Financial Services Centres Authority (IFSCA), Govt of India:

- AI - Large Language Models (LLM) | Jul 2025 – Jan 2026 | GIFT City, Gandhinagar
- Built IFSCA IntelliChat, a production grade Retrieval Augmented Generation (RAG) system using dense embeddings, cross encoder reranking, fine tuned local LLMs, citation enforcement, and confidence scoring.
 - Developed IFSCA IntelliDesk, an AI assisted service desk leveraging LLM based intent classification, deterministic workflow orchestration, RBAC, JWT authentication, and audit logging.
 - Architected IFSCA DeskFlow, a secure internal workflow and ticketing platform with role based access control, concurrency safe persistence, real time notifications, and fault tolerant design.
 - Implemented the IFSCA Mass Communication Engine, a rate limit aware email delivery system using OAuth 2.0, exponential backoff, API integrations, and observability ready logging.
 - Work on Agentic AI architectures, regulatory ML governance, and trained a custom Small Language Model (SLM) from scratch using tokenization, transformer architectures, and modern optimization techniques.

Dynamite Technology: Data Analyst | Mar 2024 – Apr 2024 | Navi Mumbai

- o Built recommendation system using TF-IDF + cosine similarity, achieving 0.78 precision@10 on holdout test set.
 - o Optimized Pandas data pipeline, cutting preprocessing time 45% via vectorized ops and feature selection for production.
-

Projects

Focus on Vision: A Statistical Study of Eye:

Tech Stack: Python, TensorFlow, Power BI, Streamlit

Impact: 1,400+ survey participants, 86.9% accuracy

- o Designed and deployed a 60 question survey capturing demographics, screen time, and genetic factors; built Power BI dashboards visualizing vision risk correlations.
- o Engineered features through EDA, handled missing value imputation, and applied SMOTE to address class imbalance before model training.
- o Built a dual model architecture combining an interpretable logistic regression model (Recall: 0.84) with a high performance TensorFlow neural network (Accuracy: 86.9%, ROC-AUC: 0.81).
- o Deployed Streamlit app for real time personalized risk assessment with confidence scores and feature importance.

AI Powered Data Analytics Platform:

Tech Stack: Gemini LLM, Python, Streamlit, DataStax Vector DB

Impact: Dynamic code generation

- o Architected platform enabling natural language queries to automatically generate executable Python code for data analysis and visualization without manual coding.
 - o Fine tuned domain specific transformer models to improve contextual precision and reduce hallucination on financial and healthcare datasets.
 - o Built stateful Streamlit UI with DataStax backend for session management, embedding storage, and conversational analytics—supporting multi turn dialogue and chat history.
-

Education

- **M.Sc in Statistics** | CGPA: 8.90 / 10 | University of Mumbai | 2023 : 2025
 - **B.Sc in Statistics** | CGPA: 9.40 / 10 | University of Mumbai | 2020 : 2023
-

Certifications

- **Generative AI with Large Language Models:** Coursera
- **Data Analytics using R Software & Excel:** Ramniranjan Jhunjhunwala College [Value Added Courses]
- **TCS iON National Qualifier Test:** IT Programming