

Anubhav Verma

Data Scientist • Applied ML Engineer • NLP & GenAI Specialist

anubhavverma9369@gmail.com | +91-9451954022 | GitHub | LinkedIn | Tableau Public

PROFESSIONAL SUMMARY

M.Sc. Data Science candidate (VIT Chennai, CGPA: 8.56) with **5+ production-deployed ML/NLP systems** spanning healthcare diagnostics, misinformation detection, personal finance, and research grant analytics. Achieved **96% text classification accuracy** (BERT, 50k+ articles) and **88.3% brain tumor classification accuracy** (Hybrid Quantum-Classical, +6% over DenseNet121 baseline). Currently interning as **ML Engineer at CDAC India** and **GIS & Data Pipeline Engineer at ShieldNex** (IN-SPACe NGP 2024). DST INSPIRE Scholar (Top 1% nationwide). Seeking early-career **Data Scientist / ML Engineer** roles to turn data into measurable business decisions.

EXPERIENCE

GIS & Data Pipeline Engineer

Jun 2026 – Present

ShieldNex · Part-time · Remote · Andhra Pradesh, India

- Building satellite imagery preprocessing and geospatial data pipelines for a defence-linked startup operating under **IN-SPACe NGP 2024** compliance frameworks.
- Handling sensitive geospatial datasets to support real-time spatial analytics at scale.

Machine Learning Engineer Intern

May 2026 – Present

CDAC India (C-DACINDIA) · Internship · Remote · Chennai, Tamil Nadu

- Contributing to an ML-based project automating and simplifying the **end-to-end machine learning workflow**, reducing manual pipeline overhead for production deployments.

Data Science Researcher

Sep 2025 – May 2026

Google Developer Group (GDG VIT Chennai) · Part-time · On-site · Chennai, Tamil Nadu

- Active member of the GDG Data Science community, participating in workshops, technical discussions, and collaborative AI/ML projects focused on real-world problem solving.
- Contributed to group-led research initiatives and knowledge-sharing sessions on machine learning, deep learning, and emerging data science trends.

PROJECTS

Real-Time Misinformation Detection Platform | BERT, Flask, JavaScript, Deep Learning

Sep 2025

- Fine-tuned BERT on 50k+ multi-source news articles to achieve **96% classification accuracy** across politically diverse datasets, with robust generalisation beyond training distribution.
- Deployed the model via a **Python/Flask REST API** consumed by a JavaScript browser extension, enabling **sub-second inference** at the point of reading with zero user friction.

AI-Powered Research Grant Evaluation & Monitoring System | SciBERT, XGBoost, SHAP, Streamlit

Nov 2025

- Achieved **76.8% funding-decision accuracy** across **9,200+ grant records** via a hybrid NLP + gradient boosting pipeline (SciBERT embeddings → XGBoost) with **SHAP decomposition** for explainable, stakeholder-defensible predictions.
- Deployed a real-time Streamlit dashboard flagging **fund misuse anomalies** and visualising allocation metrics — providing program administrators a single portfolio-level view.

PaperMind: LLM-Powered Research Paper Q&A with RAG | LLaMA 3.1, Groq, Pinecone, Sentence-Transformers, PyMuPDF

Mar 2026

- Built a **retrieval-augmented generation pipeline** (semantic chunking → 384-dim embeddings → Pinecone vector search → top-5 grounded context → LLaMA 3.1), reducing hallucinations versus a prompt-only baseline.
- Deployed a live Streamlit app with **chunk-score transparency** confirming retrieval quality in production and giving non-technical users a zero-setup research assistant.

Hybrid Quantum-Classical Brain Tumor Classifier | PennyLane, DenseNet121, FastAPI, React, Grad-CAM

May 2026

- Engineered a **hybrid quantum-classical architecture** combining DenseNet121 feature extraction with a **4-qubit variational PennyLane circuit**, achieving **88.3% test accuracy** on **6,000 MRI scans** (BRISC 2025) — a **+6% improvement** over the classical baseline — across Glioma, Meningioma, Pituitary, and No-Tumor classes.
- Deployed a **full-stack diagnostic web application** (React/Vite → Vercel; FastAPI → Render) with **Grad-CAM heatmap visualisation**, 4-qubit circuit diagrams, and auto-generated clinical reports — enabling interpretable, research-grade inference for non-technical users.

Splendidly — AI-Powered Personal Finance Mobile App | React Native, Expo, Firebase, TypeScript, Zustand

Apr 2026

- Built a **cross-platform mobile finance application** with Firebase Auth, Zustand state management, and **ML-based auto-category detection** (e.g., “Uber” → Transport) across 13 spending categories — deployed via Expo Go with QR-accessible live demo.
- Delivered a full analytics suite: 7-day bar charts, month-over-month comparisons, per-category budget progress with colour-coded overspend alerts, and **CSV/PDF export** — translating raw transaction data into actionable financial insights.

TECHNICAL SKILLS

- **Programming & Databases:** Python, JAVA, C, SQL, R, MongoDB, Hadoop
- **ML & Statistical Modeling:** Scikit-learn, XGBoost, TensorFlow, PyTorch, Regression, Classification, Clustering, Time-Series, A/B Testing, Hypothesis Testing, Cross-Validation
- **Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Power BI, Tableau, Excel, EDA, Feature Engineering, Statistical Inference
- **NLP & GenAI:** BERT, SciBERT, BART, Sentence-BERT, Hugging Face Transformers, RAG Pipelines, LLMs, FAISS, Pinecone, Fine-Tuning, Prompt Engineering, OCR
- **Quantum ML:** PennyLane, Variational Quantum Circuits, Hybrid Classical-Quantum Architectures
- **MLOps & Deployment:** Flask, FastAPI, Streamlit, Gradio, REST APIs, Git/GitHub, Vercel, Render, CI/CD Pipelines
- **Explainability & Analytics:** SHAP (XAI), Grad-CAM, Model Interpretability, Bias Detection, Vertex AI, Groq API

EDUCATION

M.Sc. Data Science Vellore Institute of Technology, Chennai, Tamil Nadu <i>Coursework: Machine Learning, Deep Learning, NLP, Statistical Modelling, Big Data Analytics</i>	<i>2024 – 2026</i> <i>CGPA: 8.56</i>
B.Sc. Physics, Chemistry & Mathematics (PCM) Dr. Rammanohar Lohia Avadh University, Ayodhya, Uttar Pradesh <i>Coursework: Linear Algebra, Probability & Statistics, Calculus, Classical Mechanics, Quantum Physics, Analytical Chemistry</i>	<i>2021 – 2024</i> <i>CGPA: 8.1</i>

ACHIEVEMENTS & CERTIFICATIONS

- **DST INSPIRE Scholarship** — Government of India *2021 – 2026*
Full merit scholarship for Top 1% nationwide STEM performance; sustained 3+ consecutive years.
- **NPTEL Elite Certification: Introduction to LLMs** — IIT Madras *Oct 2025*
12-week MoE-funded advanced course on large language models; **Elite distinction (72%)**.
- **Generative AI Certification** — Intel Unnati Programme *Aug 2025*
Industry-sponsored GenAI certification: foundation models, prompt engineering, and deployment pipelines.
- **Basics of Quantum Information** — IBM Quantum *Mar 2026*
IBM-issued badge covering quantum states, circuits, and information-theoretic foundations.

SOFT SKILLS & LANGUAGES

Soft Skills: Strong Analytical Thinking, Clear Communication of Complex Ideas, Fast Learner, Comfortable with Ambiguous Problems, Collaborative Team Player, Attention to Detail

Languages: Hindi (Native), English (Professional)