

Arnab Deogharia

+91-8240851562 arnab.deo@bhu.ac.in [LinkedIn](#) [Github](#) [Blogs](#)

Education

Banaras Hindu University <i>Masters in Statistics & Computing – CGPA 8.16 (till semester 2)</i>	2024-Present Varanasi, U.P.
Sister Nivedita University <i>Bachelor of Science in Statistics (Hons.) – CGPA 7.98</i>	2021-2024 Kolkata, W.B.

Experience

SpectoVX Assistive Technology Pvt Ltd. click <i>GenAI Trainee Intern</i>	Remote July 2025– Present
<ul style="list-style-type: none">Working in a team on multiple projects involving AI technologies like large language models, image generation, or multimodal AI systems.	
Institute of Data Engineering, Analytics and Science Foundation – ISI Kolkata click <i>Data Science Intern</i>	On-site May 2025 – Present
<ul style="list-style-type: none">Working on a project on topic 'A Smart Material Prediction System for Planetary Exploration' under supervision of professor Dr Supratik Pal, ISI Kolkata.	

Technical Skills

Programming & Tools: Python, R, Jupyter Notebook, Collab, VSCode, Excel, LaTeX.
Data & ML Libraries: Pandas, NumPy, SciPy, XGBoost TensorFlow, Keras.
LLM Frameworks: Transformers, LangChain, LangGraph, HuggingFace.
Vector Stores & AI Orchestration: Chroma, Qdrant, n8n, CrewAI, Smolagents.
Visualisation: PowerBI, Matplotlib, Seaborn, Plotly.
Deployment: Streamlit, Flask, AWS, Render, Replit.
Databases: MySQL, MongoDB.

Projects

- RustgzhHelps - Telegram AI Bot** | Aioqram, Groq API, Replit [click](#)
- RustgzhHelps is a Telegram bot built using Python and the Aioqram framework, integrated with the Groq API to access LLaMA 3 models, and deployed on Replit.
- PDFIQ - Ask your PDF** | RAG, QdrantDB, Streamlit [click](#)
- A PDF-based question-answering app that uses LangChain, Qdrant, and OpenAI to retrieve context-aware responses. Built with Streamlit for interactive querying and powered by LLM embeddings for semantic search.
- Music Recommender (Lyrics Based)** | Recommender system, NLP, Streamlit [click](#)
- Lyrics-based music recommendation app built with Streamlit, leveraging NLP to match songs by meaning.
- Cancer Diagnosis using Medical Literature** | Supervised, Ensemble Learning & Neural Networks. [click](#)
- Classifying the given genetic variations/mutations based on evidence from text-based clinical literature.
 - Built a neural network model to classify genetic mutation types (0–9) using clinical text data, achieving 90.8% accuracy and outperforming traditional ML models by 15%.
- Insights into P -value Behaviour under both the Null & Alternative hypotheses.** | Hypothesis Testing, Statistical Analysis. [click](#)
- In a continuous setup the distribution of P -value under alternatives depend on distributions of test statistic and generally it is skewed distribution unlike P -value under null hypotheses which follows Uniform $[0,1]$.

Training & Certifications

Introduction to Transformer-Based Natural Language Processing | NVIDIA [click](#)
Generative AI - Learn by doing | IIT Kharagpur (Ongoing) [click](#)
Data Analytics with Python | NPTEL (Top 2% among 10,000+ participants) [click](#)

Academic Achievements

GATE STATISTICS '24 | Rank AIR 169 (91.02 percentile) [click](#)
TIMES SPARK SCHOLARSHIP (2019-20) & (2021-22) – Awarded for top performance after clearing competitive assessments.