# TANU TIWARI

+91 7999244527 ◇ tanutiwari4527@gmail.com ◇ [Linkedin](#) ◇ [Github](#)

## EDUCATION

**Bachelor of Engineering** in Electronics and Telecommunication Engineering.                    2015-2019
Jabalpur Engineering College, Jabalpur

## SKILLS

| | |
|---|---|
| **Technical** | Python, Pandas, Numpy, Xarray, pytorch, sklearn, SQL, Matplotlib, Seaborn, Nltk, Spacy, DAX, Reports Random Forest , Decision Trees, Logistic Regression, Linear Regression, Regularisation, Neural Networks CNN, RNN, LSTMs, NLP,Transformers, LLM, VLM, Time Series Forcasting, Spatio-Temporal Modeling. |
| **Tools** | VSCode, Jupyter Notebooks, HuggingFace, Mysql, Oracle, MSSQL server, Power BI, Github, Google colab. |

## WORK EXPERIENCE

**Junior Research Fellow (Data Science)**, IIT Gandhinagar, Gujrat                    Oct 2024 - Present

- Processed and transformed large-scale air quality data for years 1998–2025 from CPCB (CSV format) and satellite sources (WUSTL, LandScan, ERA5) using **Python**, **Xarrays**, and **Pandas** for downstream forecasting tasks.
- Investigated the influence of meteorological factors on multi-day PM2.5 forecasting accuracy by applying data fusion techniques to integrate satellite imagery, traffic data, weather patterns, and geophysical variables for improved forecasting.
- Developed an AQI prediction model using **LSTM**, **Transformer**, and **Neural Process**, achieving the best RMSE of 8.1.
- Created **visualizations** and GIFs using **Python** to communicate air pollution trends via **Matplotlib** and **Seaborn**.

**Systems Engineer** - Tata Consultancy Services, Pune, India                    Jan 2020 - Oct 2024

**Data Scientist - Tata Cliq (E-Commerce)**

- Forecasted inventory demand using **LSTM** and **Transformers** on historical sales data to prevent stockouts, leading to a 15% increase in sales during Diwali and Christmas holidays.
- Performed sentiment analysis on reviews using **BERT** to identify underperforming vendors based on customer feedback.
- Built a **RAG**-based POC using **LangChain** to answer customer queries by leveraging relevant product reviews as context.
- Analyzed Customer's past purchases, browsing behavior, demographics to predict the likelihood of a customer purchasing a specific product category using techniques like **Random Forest** based classification methods.

**Data Scientist – Tata Consumer Products (Supply Chain Analytics)**

- Automated ETL pipeline using **Python** to fetch, clean, and transform SQL data into structured CSV files used for inventory forecasting, supplier segmentation, and Power BI dashboarding.
- Clustered suppliers using **K-Means** based on delivery delays, defect rates, and order volumes to identify high-risk vendors.

**Data Analyst - MARS Chocolates and Pet Care (Power BI)**

- Developed interactive dashboards using **Power BI** to visualise key performance indicators (KPIs) for sales and marketing.
- Reduced report development time by 34% by implementing **reusable DAX** measures and calculated columns for reporting.
- Designed custom visuals and triggered alerts tailored to business needs, streamlining executive decision-making processes.
- Wrote **SQL stored procedures** to extract and transform data, ensuring accurate and timely insights.

## PERSONAL PROJECTS

**RAG using Phi-2**  [Github Repo]
Implemented RAG from scratch utilizing MS Phi-2 for a PDF with 1200 pages using PymuPDF, Sapcy, NLTK, and HuggingFace

**FineTuning GPT-2** [Github Repo]
Fine-tune GPT-2 for Twitter sentiment analysis and achieved an accuracy of 70.88% on test set.

**Zero Shot analysis of Florence-2 VLM** [Github Repo]  [HuggingFace App]
Zero shot object detection and image caption generation on MS Coco dataset.