Anup Ramesh Bhagat

Data Analyst · Data Scientist · AI Engineer

Berlin, Germany · **+49 176 29979288** · anupbhagat5080@gmail.com · linkedin.com/in/-anup-bhagat- · anupbhagat77.github.io · github.com/Anupbhagat77

Profile

Data professional with 3+ years in data analytics: three years of BI and KPI reporting at TCS, followed by hands-on work as a Data Scientist / Data Engineer building **RAG solutions**, **LLM workflows**, validated **data pipelines** and **Django** backends. Strong in **SQL**, **Python**, Power BI and statistics, with a focus on data quality and measurable results. Completing an M.Sc. in Project Management & Data Science at HTW Berlin (October 2026); available full-time from November 2026.

Experience

Masem Research Institute GmbH | Working Student, Data Scientist / Data Engineer — 08/2025 – present
Wiesbaden, Germany

- Built **RAG solutions** integrating domain-specific knowledge with LLMs, evaluating models on quality, latency and cost trade-offs.
- Designed structured **information-extraction pipelines** with validation checks, turning unstructured business and financial content into analysis-ready datasets.
- Developed backend applications and REST APIs in **Django**, optimizing data models for scalable storage, retrieval and workflow automation.
- Built AI-driven workflow automations in **n8n** with LLMs and prompt engineering; tested and validated LLM outputs end-to-end before rollout.
- Worked with Docker, Git, CI/CD and Python unit tests in an agile team, including reviewing and being reviewed on pull requests.

Tata Consultancy Services (TCS) | Data Analyst — 08/2021 – 08/2024
Pune, India

- Built **10+ Power BI and Qlik Sense dashboards** from the data model up on sales, revenue and operational KPIs for **20+ stakeholders**; independently owned 2–3 end-to-end.
- Built CXO-level executive dashboards with actual-vs-target and prior-period **variance analysis**; wrote complex DAX, published to the Power BI Service and implemented row-level security.
- Cut reporting preparation time by **~20%** by automating consolidation of SAP HANA, SQL and Excel data with SQL and Python.
- Kept reporting accuracy **above 95%** through data-quality checks and reconciliation, and improved SQL query performance by **up to 50%**.

Selected Projects

LLM Model Cascade for Cost-Efficient Inference — M.Sc. thesis · Python, LiteLLM, GPT-4.1, Gemini
Benchmarked cascade routing on 100 MMLU/GSM8K questions: an **LLM-as-judge router kept 98.8% of GPT-4.1 quality at 25.7% lower cost**; oracle routing set the ceiling at 59.8% savings.

Privacy-Preserving Vision with Generative Inpainting — Technical lead · PyTorch, Stable Diffusion, YOLO, Streamlit
Detected faces (InsightFace) and licence plates (YOLO) and regenerated them with **Stable Diffusion 2 inpainting**; video support and PSNR/SSIM checks.

Sarcasm Detection: Classic ML vs. BERT — Technical lead, team of six · scikit-learn, Hugging Face
Compared five classifiers on TF-IDF features; engineered linguistic features raised sarcastic-class **F1 from 0.74 to 0.83** (ROC-AUC ≈ 0.97); fine-tuned BERT and validated on a hand-labelled gold set.

Music Streaming Analytics & A/B Test Design — Solo case study, UMG Germany dataset · Python, Power BI
Merged plays, tracks and users data in Pandas, built two interactive Power BI dashboards, and designed a 14-day A/B test on personalized recommendations with listening-duration metrics.

Skills

Machine Learning	scikit-learn, XGBoost, Random Forest, Logistic Regression, SVM, Hugging Face Transformers (BERT), YOLO, InsightFace, OpenCV, PyTorch, SHAP, imbalanced data
AI Engineering	RAG, LLM evaluation, model routing, prompt engineering, LangChain, Stable Diffusion (diffusers), OpenAI / Gemini APIs, LiteLLM, n8n, Django, Streamlit, REST APIs
Data & BI	SQL, Python, Pandas, Power BI / DAX, Qlik Sense, Tableau, Looker, Excel / Power Query, hypothesis testing, A/B test design
Engineering	ETL, data modeling, SAP HANA, Pydantic, Git, Docker / Compose, CI/CD, unit testing, Google Cloud; dbt and Snowflake (basic)

Education

M.Sc. Project Management & Data Science 10/2024 – 10/2026
HTW Berlin

B.Sc. Mathematics & Statistics 08/2017 – 08/2020
Savitribai Phule Pune University

Certifications & Languages

Microsoft Power BI Data Analyst (PL-300) · Google Cloud Associate Cloud Engineer · IBM Data Science · IBM Data Analyst · Google Project Management

Languages: English C1 · German B1