



# Arpit Baranwal *Data Scientist*

✉ arpitbaranwalaec@gmail.com    ☎ +491789391284    📍 Freiburg

🔗 [medium.com/@arpitbaranwal](https://medium.com/@arpitbaranwal)    🌐 [github.com/Arpit-Baranwal](https://github.com/Arpit-Baranwal)

📄 Work Authorization: Germany (no sponsorship needed)

## PROFILE

---

Data Scientist with 5+ years building ML pipelines, statistical models, and applied AI across healthcare and energy research. Expert in Python, SQL, and AWS with hands-on experience in LLM-based RAG, uncertainty quantification, and end-to-end model deployment. Published researcher skilled at translating complex analytics into data-driven product decisions

## PROFESSIONAL EXPERIENCE

---

**Independent Data Engineer & ML Developer** 04/2026 – Present

- Built an end-to-end Databricks pipeline (Bronze → Silver → Gold) ingesting **3 data sources** (POS, weather, operational) processing **~6K+ monthly transactions**.
- Engineered a KPI dashboard tracking **15+ metrics** and a sentiment analysis pipeline across **200+ customer reviews** to surface actionable insights for menu optimisation.
- Developing a dish demand forecasting model across **30+ menu items** using **12 months** of historical order data and weather signals.

**Student Research Assistant, Fraunhofer ISE** 06/2024 – 04/2026

- Developed ML models for risk prediction and anomaly detection across **3+ physical system configurations**, integrating **30+ real-world sensor signals** per setup.
- Built a probabilistic surrogate model (TabPFN) trained on **~500 simulation samples**, delivering uncertainty-aware predictions **~100x faster** than full physical simulations.
- Designed a Kriging-Monte Carlo prediction framework across **10,000+ Monte Carlo iterations**, achieving **R<sup>2</sup>=0.81** through rigorous cross-validation and uncertainty benchmarking.
- Deployed a RAG knowledge system indexing **1000+ internal technical documents** across **3 research teams**, reducing manual information retrieval time by **~30%**.

**Associate Cloud Intern, RealWorldOne** 09/2023 – 02/2024

- Implemented secure, scalable ML infrastructure on AWS, contributing to model deployment workflows across development and production environments.
- Designed IAM-based access control policies ensuring compliant, role-based data handling across the engineering team.
- Contributed to CI/CD pipelines improving deployment reliability and release consistency for ML services.

**AWS Python Developer, Legato Health Technology** 06/2019 – 05/2022

- Developed **5+ backend services** and data processing APIs in C# .Net, building the foundational healthcare data layer that later fed into ML pipelines.
- Engineered **3+ ML-driven ETL pipelines** processing **2M+ healthcare records monthly**, maintaining data integrity with built-in anomaly detection.
- Built an automated OCR identity-verification system processing **thousands of documents monthly**, eliminating manual review across **2+ business units**.
- Designed **4+ serverless ingestion architectures** (AWS Glue, Lambda, AppSync) provisioned via Terraform across **3 environments** (dev, staging, production).

## EDUCATION

---

**Master's in Embedded Systems Engineering, University of Freiburg** 04/2022 – 03/2026

Specialization In AI with coursework in deep learning, reinforcement learning, computer vision, robotics, microsystems engineering and optimization, building strong skills in applied ML and system modeling. Freiburg

**Bachelor's in Electronics and Communication Engineering, Asansol Engineering College** 2016 – 2020  
Asansol, India

## SKILLS

---

- **Programming** — Python, C++, C#, SQL, NoSQL, Pyspark, .Net
- **Machine Learning** — Predictive modeling, Anomaly detection, Time Series, Transformers, TabPFN, Classical ML, RAG
- **Statistics** — Probabilistic modeling, Monte Carlo methods, uncertainty quantification, A/B testing
- **Frameworks** — PyTorch, TensorFlow, Yocto
- **ML Ops / Deployment** — AWS (Lambda, Glue, Fargate, AppSync), CI/CD
- **Data Engineering** — ETL pipelines, data quality assurance
- **Tools** — Git, JIRA, Confluence, GraphQL, MATLAB, n8n, Databricks, PowerBI, Excel
- **Soft Skills** — Stakeholder communication, analytical storytelling, cross-team collaboration

## PUBLICATIONS

---

- Data-Driven Spatial Risk Assessment of Indoor Gas Dispersion using Kriging Regression and Monte Carlo Uncertainty Analysis** <sup>Ⓞ</sup>, 2026  
IEA Heat Pumping Technologies TCP – Project 64 Annual Report 2026  
In Press (Master Thesis)
- A Diagnostic Framework for CFD Validation of R-290 Dispersion in an Indoor Heat Pump Installation Room,** 08/2026  
IIR Gustav Lorentzen Conference on Natural Refrigerants (GL2026), Hamilton, New Zealand.  
(Submitted)
- Ensemble Method Of Feature Selection using Filter and Wrapper Techniques with Evolutionary Learning** <sup>Ⓞ</sup>, Springer Nature 16/04/2022

## Conference proceeding

---

- Chewing sound interpretation by Deep Learning** <sup>Ⓞ</sup>, Cosima 2023  
Developed an end-to-end, AI-based signal processing pipeline for healthcare diagnostics, Dresden  
creating a novel system using sensor-equipped eyewear and deep learning for real-time audio classification.

## PROJECTS

---

- EEG signal Analysis using Vision Transformers** 2024
- Applied **Vision Transformer** models to analyze **EEG signals**, enhancing feature extraction through attention mechanisms.
  - Developed techniques to capture complex temporal and spatial patterns in neural data for improved classification accuracy.
- Optimization-Based Rocket Landing Control (MPC)** 2024
- Implemented a **nonlinear MPC** pipeline using **CasADi** to stabilize and land rocket in simulation.
  - Modeled **system dynamics from first principles**, performed **constrained optimization** of control inputs, and validated performance through **PyMunk physics simulation**.

## COMMUNITY & LEADERSHIP ACTIVITIES

---

- Head Chef, Vietnamese Restaurant (Part-time)** 2023 – Present  
Freiburg
- Managed kitchen operations and workflow during peak hours, ensuring consistent quality, teamwork and time critical decision making in a high pressure environment.
- Cashier, Student Dormitory** 2023 – 03/2026  
Händelwohnheim, Freiburg
- Handled cash transactions, reconciled accounts and supported community operations with accuracy and responsibility.
- Event Coordinator, Freiburg Student Festival** 2023  
Freiburg
- Helped organize one of Freiburg's largest student festivals, coordinating logistics, budgeting and volunteer teams for a 300 plus attendee event.