

# PRIYANKA ROY

+91 15234700630 /+91 9748893245    proy272001@gmail.com    LinkedIn    GitHub

## Education

### Universität Mannheim

Master in Data Science

Mannheim, Germany

Sep 2022 - Mar 2025

### Maulana Abul Kalam Azad University of Technology

Bachelor in Computer Applications |Grade: 9.0/10.0

Kolkata, India

July 2019 - June 2022

## Experience

### Universität Mannheim

Research Assistant

Nov 2024 – Jan 2025

Mannheim, Germany

- Conducted large-scale data reconciliation across multiple formats and sources, addressing inconsistencies to ensure accurate and unified datasets for research and analytical use
- Applied R programming for advanced anomaly detection, leveraging statistical techniques and pattern recognition to identify and resolve complex data irregularities
- Enabled seamless data integration across heterogeneous datasets

### ABB

Working Student as Product Data Manager

Apr 2023 – Aug 2024

Heidelberg, Germany

- Managed and optimized master data across multiple product categories, ensuring accuracy, completeness, and compliance with ABB's global data governance standards
- Played a key role in enhancing data consistency and accessibility by identifying and correcting discrepancies across databases
- Supported the development and execution of streamlined product data integration strategies, facilitating smooth onboarding of new products into internal systems and external platforms (PIM)
- Contributed to data migration and harmonization projects, using Excel and data visualization tools to clean, transform, and validate large datasets

### Universität Mannheim

Research Assistant

Nov 2022 – Apr 2023

Mannheim, Germany

- Focused on comprehensive analysis of census data to understand demographic trends and patterns

## Projects

### Master Thesis on Speaker Diarization - Chair of Artificial Intelligence | Python, Pyannote, Silero, Whisper, ECAPA-TDNN

- Designed real-time pipelines for dynamic, multi-speaker conversations.
- Integrated and benchmarked state-of-the-art models to balance latency and accuracy.
- Fine-tuned segmentation and clustering thresholds to reduce error rate while maintaining low-latency inference.
- Profiled system performance to eliminate bottlenecks for scalability.

### Foodle - Chair of Artificial Intelligence | Python, Streamlit, Whisper, Large Language Models

- Developed an interactive real-time cooking assistant for amateur chefs, integrating Whisper for high-accuracy transcription and conversion of culinary audio recordings into detailed textual recipes.
- Published on : Foodle: ASR

### SBA Loan Approval - Data Mining | Python, Pandas, Scikit-learn, Google Colab

- Applied gradient boosting and data mining techniques to predict loan approvals.
- Focused on large-scale data cleaning and exploratory analysis.

## Participations and Achievements

---

### **Hackathon 2024: Jivs** | *Streamlit, Sentence-Transformers, RAG* | Kreuzlingen, Switzerland

- Developed a tool for automated schema mapping across databases using sentence-transformers (all-MiniLM-L6-v2) and RAG.
- **Secured 1st place for the particular theme**

### **Start Hack 2025** | *DeepFilterNet, speechlib, Pyannote, Whisper, GPT-4-o-mini, Qdrant, AWS* | St. Gallen, Switzerland

- Developed a voice-based assistant that identifies speakers, transcribes conversations, and personalizes interactions by remembering user preferences.
- Enhanced user experience by enabling seamless, memory-driven conversations for both new and returning customers.

### **CheffTreff 2024: Otto Dorner** | *Python* | Hamburg, Germany

- Engaged in an analysis of time series data, focusing on demand forecasting. Worked with ARIMA and Prophet Model to forecast future trends based on historical data.

### **Bosch Connected Experience 2024** | *Python* | Berlin, Germany

- Developed a Python-based proof of concept for an automated parking negotiation system, utilizing Pydantic for data validation.

### **HerHackathon 2023** | *Python* | Mannheim, Germany

- Developed a mobile app for Freudenberg Group employees using Flutter, promoting ride-sharing to reduce CO2 emissions, enhanced with gamified incentives

### **SystAI3r** | *Python, PyAudio, SpeechBrain, Whisper, HuggingFace, NLTK, IBM Watson, Streamlit, Plotly* | Munich, Germany

- **Secured 3rd place** by building a tool enhancing meeting dynamics through advanced metrics analysis.

## Certifications

---

### **Deloitte Data analytics Job Simulation** | Tableau

### **Data Visualization and Dashboards with Excel and Cognos** | IBM Cognos Analytics, Microsoft Excel

## Skills

---

**Languages:** Python, R, Java, C/C++

**Concepts:** Artificial Intelligence, Machine Learning, Neural Networks, Natural Language Processing (NLP), Automatic Speech Recognition (ASR), Speaker Diarization, Retrieval-Augmented Generation (RAG)

**Technical:** Text Analytics, Data Analysis, Data Integration and Visualization, Sentence-Transformers, Scikit-learn, Pandas, NumPy, TensorFlow, PyTorch, Whisper, SpeechBrain, Streamlit, Git

**Interpersonal:** Responsible, ability to work in a team and independently in a time sensitive environment.

Flexible and would like to work on different technologies and aspects.