PRIYANKA ROY

☎ +49 15234700630 /+91 9748893245 ▣ LINKEDIN ⭘ GITHUB 🔗 PORTFOLIO ✉ PROY272001@GMAIL.COM

EDUCATION

Universität Mannheim   Mannheim, Germany
Masters in Data Science   *Sep 2022 - Mar 2025*

Maulana Abul Kalam Azad University of Technology   Kolkata, India
Bachelors in Computer Applications   *July 2019 - June 2022*

EXPERIENCE

Universität Mannheim   Nov 2024 – Jan 2025
Research Assistant   *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   Apr 2023 – Aug 2024
Working Student as Product Data Manager   *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   Nov 2022 – Apr 2023
Research Assistant   *Mannheim, Germany*
- Focused on comprehensive analysis of census data to understand demographic trends and patterns

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.

PROJECTS

Master Thesis - Chair of Artifical Intelligence | *Online Speaker Diarization, Speech Processing Pipelines, VAD, ASR*
- Developed a real-time speaker diarization system optimized for multi-speaker conversations achieving a 23% Diarization Error Rate (DER), representing a 4% absolute improvement over comparable research baselines.
- Integrated 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.
- Applied unsupervised clustering like k-means and Oracle-based clustering to segment speakers effectively.
- Implemented dimensionality reduction techniques like PCA to evaluate embedding quality.
- Profiled system performance to eliminate bottlenecks for scalability.

Foodle - Chair of Artifical Intelligence | *Automatic Speech Recognition, 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.

- Applied topic modeling to cluster and summarize user cooking queries, enhancing LLM-powered retrieval.
- Built interactive speech functionality using LangChain enabling hands-free recipe navigation and multilingual support, enhancing accessibility for diverse users.
- Published on *Foodle: ASR*

SBA Loan Approval - Data Mining | *Python, Pandas, Scikit-learn, Google Colab*
- Built predictive models using logistic regression and gradient boosting to assess loan approval likelihood.
- Applied SHAP for model explainability, highlighting feature importance in approval decisions.
- Focused on large-scale data cleaning and exploratory analysis.

PARTICIPATIONS AND ACHIEVEMENTS

Winner – DMI Challenge at 3rd Jivs Hackathon, hosted by Data Migration International (2025)

Winner – DMI Challenge at 2nd Jivs Hackathon, hosted by Data Migration International (2024)

Runner up – SystAI3r, hosted by J3D.AI (2023)

CERTIFICATIONS

Deloitte Data analytics Job Simulation | Tableau

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

Python for Data Science, AI & Development | Python

LANGUAGES

English, Hindi, Bengali, German