

# Mayank Kumar

LinkedIn: mayankkumar7

Email: mayank132.kumar@gmail.com

Mobile: +91 9582087528

## EXPERIENCE

- Eberhard Karls University of Tuebingen

Research Assistant

Tuebingen, Germany

April 2025 - August 2025

- Dysarthria Speech Correction Pipeline:** Built a modular system using HuBERT, KMeans, and Unit HiFi-GAN for detection and restoration of dysarthric speech, improving intelligibility without relying on transcripts or parallel supervision.
  - Clinical Dataset Integration & Evaluation:** Preprocessed and harmonized UASpeech and TORGO datasets and designed evaluation metrics combining symbolic edit distance, intelligibility scoring, and synthesis quality to benchmark model performance.
- SCOR SE

Analyst

Cologne, Germany

April 2023 - March 2025

**Key Skills:** Palantir Foundry, DB2, LLM, FastAPI, GPT, IBM Cognos Analytics, Microsoft Power BI, Python, Microsoft Excel, IBM OpenPages

  - LLM-Integrated Risk Intelligence:** Built a document processing feature using Azure OpenAI to extract insights from compliance reports; optimized prompts and token usage for scalable multi-departmental deployment.
  - Palantir Foundry for Issues and Risk Tracking:** Developed the data pipeline and assisted in designing a dashboard for issue tracking in Palantir Foundry during its first implementation in the organization.
  - End-to-End Risk Analytics in Power BI:** Independently developed and automated a governance dashboard pipeline; integrated decision-making signals, resulting in a 60% reduction in manual reporting time.
  - License Optimization Reporting in IBM OpenPages:** Led reporting automation for IBM OpenPages license usage; categorized licenses and flagged underutilization, delivering cost-saving insights for IT and Finance teams.
  - IBM Cognos Reporting Optimization:** Reduced report execution time by 40% through SQL tuning and dashboard logic refinement, improving user experience for critical governance workflows.
- Indian Telephone Industries Ltd.

Research Intern

Bengaluru, India

June 2019 - September 2019

- Research Documentation:** Conducted comprehensive documentation for diverse departments, with a focus on manufacturing and research. Gained valuable insights into the manufacturing and research workflows, enhancing overall organizational knowledge.

## EDUCATION

- Eberhard Karls University of Tuebingen

Master in Computational Linguistics

Tuebingen, Germany

April 2022 - July 2025
- Guru Gobind Singh Indraprastha University

Bachelor of Technology

New Delhi, India

August 2017 - August 2021

## PROJECTS & PUBLICATION

- BabyLM: Language Model Pre-Training on Constrained Datasets

Key Skills: Transformers, Knowledge Distillation, Pre-Training, GPT-2, LLaMA, Hugging Face

August 2024
- AfriSenti- SemEval: Sentiment Analysis of Low resource African Languages

Key Skills: Python, SVM, Random Forest, Huggingface, Ensemble modeling, Transformers

July 2023
- Forecasting the Future Power Consumption of Germany

Key Skills: Python3, ETS, SARIMA, LSTM, Alteryx Designer, ENTSO-E Dataset

March 2022

## CERTIFICATIONS

- McKinsey Forward

McKinsey

McKinsey

May 2025 - July 2025
- Microsoft Certified: Azure Data Scientist Associate

Microsoft

Microsoft

February 2025
- AI Programming With Python Nanodegree

Funded by Amazon Web Services (AWS)

Udacity

July 2022 - November 2022
- Python for Data Science Certificate

Indian Institute of Technology, Madras

NPTEL

January 2021 - March 2021

## SKILLS SUMMARY

- Coding:

Python, SQL, R
- Frameworks:

NumPy, Pytorch, Scikit, TensorFlow, LanfChain, Hugging Face, Transformers, RAG, LoRA, PEFT
- Tools used:

Alteryx Designer, BigQuery, Microsoft Excel, IBM Cognos, PowerBI, Palantir Foundry, KNIME, Tableau
- Platforms:

Linux, AWS, Google Cloud, Azure ML, Git
- Languages:

English, Hindi, German, Maithili