

Akash Chandra Baidya

Resume

Emil Figge strasse 25
44227 Dortmund, Germany
📞 +4915758474346
✉️ akashbaidya2@gmail.com
in akashbaidya
🌐 AkashBaidya
🌐 akashbaidya.github.io



Passionate data scientist with experience in research, development and project management. Currently pursuing a Master's degree in Data Science at TU Dortmund, with a strong focus on data science and machine learning. Seeking opportunities to apply my skills in a challenging work environment

Experience

- 2024–2024 **WerkStudent Data Science, IMSP - Institut für Marktforschung, Statistik und Prognose, Munich, Germany**
- Developing customer data science models in the area of customer and business analytics for well-known national and international companies.
 - Supporting the team on projects and taking sub-projects on initial tasks independently.
 - Providing trainings on topics of data science, data analysis, and data engineering for clients.
- 2023–2024 **Research Assistant (WHF) on Data Visualization, Forschungszentrum Borstel, Leibniz Lungenzentrum, Borstel, Germany**
- Generated visualizations of scientific data using web-based tools and custom scripts with dedicated libraries (Link).
 - Contributed to two projects on proteomics data in exploratory data analysis:
 - Analyzed protein cysteine oxidation data to provide insights and visualizations for research purposes.
 - Visualized "Batch Corrected Proteomics" samples to facilitate analysis and interpretation of experimental results.
- 2021–2024 **Research and Teaching Assistant (WHF), Technische Universität Dortmund, Dortmund, Germany**
- Utilized R and Deseq2 packages to simulate datasets for analysis, aiding in the identification of differential genes.
 - Developed a project focused on simulating datasets tailored for Deseq2 analysis, contributing to advancements in gene expression research.
 - Contributed to teaching various courses, including "Big Data Analysis", "Advanced Statistical Learning", and "Probabilistic Reasoning & Machine Learning". Responsibilities included creating exercises, conducting help desks, and grading assignments.
- 2021–2022 **Research Assistant (WHF), E.ON RWTH Energy Research Center-Aachen, Aachen, Germany**
- Developed web applications utilizing Python and the Django framework, with a focus on Docker and Ansible for streamlined deployment.
 - Employed probability frameworks to construct simulation models, improving the precision and dependability of energy system evaluations.

- 2018–2021 **Assistant Project Manager (Remote)**, *ITM System LLC.*, USA
- Led multiple Django-based project developments, including an Appointment System featuring patient alerts, reminders, and confirmation via voice and text messages, serving clients in the USA, Philippines, and Vietnam.
 - Implemented Django-based appointment systems to support COVID testing and vaccination aid programs in Arizona, USA.
- 2018–2021 **IT Officer, In-Charge (Software)**, *SINGER Bangladesh Ltd.*, Bangladesh
- Utilized SQL for generating crucial business insights and crafted informative dashboards and reports using Jasper BI.
 - Engineered and deployed a Django-based Rent Management System tailored for precise rent calculations and ensured compliance with IFRS 16 standards across organizations.
- 2017–2018 **Software Engineer (Android)**, *Ubitrix BD Ltd.*, Bangladesh
- Significantly contributed to the development of multiple projects utilizing Python, Android, and PHP, such as Speech Therapy Online, Ag Palliative Care, and Natural Family Planning.

Projects

- 2022 – 2023 **Time Series Processing, Anomaly Detection, and Explanation**, *Pytorch, Scikit, Pandas*
- Developed and tested machine learning models to process time series data and detect anomalies in energy consumption data.
 - Utilized various models including Long Short-Term Memory (LSTM), Temporal Convolutional Network (TCN) encoder-decoders, and a Deep Gaussian Hidden Markov Model (DGHL) generative model.
 - Analyzed and compared the performance of the models using data collected from over ten facilities of TU Dortmund.
- 2023 – 2024 **Bundesliga 23-24 Analysis (GitHub Link)**, *Power BI*
- Utilized Power BI Desktop to import, clean, and format .csv data, simplifying analysis.
 - Designed a comprehensive dashboard with sections showcasing key Bundesliga season stats and player metrics.
 - Implemented visually appealing visualizations and interactive features for effective data exploration and insights.
- 2023 – 2024 **Credit Card Fraud Detection (GitHub Link)**, *Python, Machine Learning*
- Developed and implemented a machine learning model to detect fraudulent credit card transactions, utilizing Python and various machine learning algorithms.
 - Preprocessed and cleaned the dataset, handling missing values and outliers to improve model accuracy and reliability.
 - Utilized techniques such as logistic regression, decision trees, and random forests to train and evaluate the model's performance.
 - Performed feature engineering to identify important features and reduce dimensionality, enhancing the model's efficiency and interpretability.

Education

- 2020–2024 **Master of Science in Data Science**, *Technische Universität Dortmund, Dortmund, Germany*, *GPA: 2.4*
<https://www.tu-dortmund.de/>
- Major Courses Advanced Statistical Learning, Statistical Theory, Information Mining, Deep Learning, Big Data Analytics, Bioinformatics, Machine Learning Paradigms for Complex Data & Industrial Data Science

2012–2017 **Bachelor of Science in Computer Science & Engineering**, *Khulna University of Engineering & Technology, Khulna, Bangladesh*, GPA: 2.48
<https://www.kuet.ac.bd/>

Thesis Distributed MAC protocols for VANET (Criticalities for RR-Aloha, MS-Aloha, RR-Aloha+ and proposed a new distributed Aloha MAC protocol for VANET)

Skills

Programming Python (NumPy, Pandas, Scikit-learn, Keras, PyTorch, FastAPI, Requests, Django),
Languages R (dplyr, ggplot2), Java, PHP, SQL
Visualization Power BI, Tableau, RShiny, ggplot2, Matplotlib, Seaborn
Other Git, Docker, Azure, AWS, AWS Redshift, AWS Kinesis, AWS SageMaker, BigQuery, Agile Methodologies, IaaS, SaaS, Databricks
Data Analysis Advanced Excel (formulas, pivot tables, Power Query, VBA excel), Statistical Analysis (SciPy, StatsModels, SAS, hypothesis testing, regression modeling)
Big Data Apache Hadoop, Apache Spark, ETL (Extract, Transform, Load) tools (Apache Airflow)
Database SQL Databases (MySQL, PostgreSQL, Oracle), NoSQL databases (MongoDB)
O.S. Windows, Linux (Redhat, Ubuntu), MacOS
Applications Microsoft Office (365): Word, Excel, PowerPoint

Languages

English Fluent, C1 level
German Beginner, A2 level
Bengali Native Speaker

Online Certification

2024 **Azure Fundamentals**, *Datacamp*
2024 **Data Visualization with Power BI**, *Datacamp*
2023 **Deep Learning with pytorch**, *Datacamp*
2021 **Cluster Analysis in R**, *Datacamp*
2021 **Cluster Analysis in Data Mining**, *Coursera*, Issued by University of Illinois at Urbana-Champaign
2019 **SQL for Data Science**, *Edx*, Issued by University of California, Davis
2019 **Applied Scrum for Project Management**, *Edx*, Issued by University of Maryland, College Park