Shivam Dabral

shivamd04@gmail.com | +49 (0)176 8701 1837 | Dortmund, Germany

Summary

I completed my **M.Sc. in Data Science (Statistics)** at TU Dortmund and recenlty worked at **RWE**, focusing on automation, data pipeline optimization, and predictive modeling. My expertise spans **statistical learning, machine learning for high-dimensional data, NLP, bioinformatics, and big data analytics**, applied across both research and industry environments. I have practical experience with **Python, R, SQL, Azure (CI/CD, Databricks), and AWS Cloud**, as well as data visualization and monitoring tools such as **Tableau, Power BI, Elasticsearch, and Kibana**. Driven by the goal of transforming complex datasets into actionable insights, I strive to design innovative, data-driven solutions that support strategic and operational decision-making in dynamic settings.

Professional Experience

Working Student (Data Analyst) Dec 2023 – Mar 2025
RWE Supply & Trading GmbH, Germany

- Developed and optimized data pipelines and built predictive models for **energy trading, supply chain, and risk analysis** using statistical learning techniques.
- Constructed forecasting models for trading and supply chains, leveraging **Azure and AWS cloud** to scale processing, reduce prediction error, and enhance efficiency.
- Integrated real-time data ingestion via **RESTful APIs** with automated reporting and risk monitoring to enhance data quality and system reliability, enabling seamless analysis through interactive dashboards.

Intern (Data Analyst) Jul 2022 – Aug 2022
KPMG, Germany

- Performed data integrity diagnostics and **exploratory multivariate analysis** on demographic and transactional datasets to identify patterns and anomalies.
- Developed statistical models in **Python** and designed interactive dashboards in **Power BI**, enabling evidence-based business decision-making.

Student Research Assistant Jul 2021 – Jul 2022
FLW, TU Dortmund

- Technical and conceptual support in research and industrial projects.
- Data modelling, pipelining and evaluation of new scenarios in simulation environments using **deep learning and neural network techniques**.
- Configured **simulation frameworks** for temporal data analysis and designed test scenarios for **human activity environments**.

Software Developer Dec 2017 – Mar 2020
Accenture (BMW), India

- Designed and deployed an **ML-powered dashboard** to optimize manufacturing processes and predict production capacity using supervised learning and statistical analysis.
- Built data pipelines with **MongoDB, Docker, Microsoft Azure, and AWS** to enable scalable and reliable analytics.
- Collaborated closely with stakeholders within the **Agile framework** to resolve system-level issues, refine requirements, and ensure seamless integration of predictive systems into production environments.

Intern (Data Engineer) May 2016 – Jun 2016
Bharat Electronics Ltd. (Ministry of Defence, India)

- Built a **sensor data monitoring and visualization system** using machine learning (Random Forest, K-Means) to detect anomalies and predict equipment states.
- Applied **PCA for dimensionality reduction** in time-series data, enhancing real-time diagnostics and improving signal clarity in operational diagnostics.

Projects

Interactive AI-Driven Analytics Dashboard (Jun 2025 – Present)
Building a React and Dash-based dashboard with multi-format file ingestion, dynamic visualizations, and an AI-enabled chatbot using LangChain for contextual search and insights.

Master Thesis – Statistical Comparison of Dataset Similarity Measures (Dec 2023 – Jun 2024)
Developed a simulation-based research framework to benchmark and analyse similarity measures, deriving statistical power curve analysis to assess sensitivity, robustness, and limitations in high-dimensional data environments.

BTA Time Series Deep Hole Drilling Process (Jun 2023 – Sep 2023)
Developed FFT and periodogram-based predictive models for real-time noise diagnostics and early-warning systems for real-time chatter detection in stochastic drilling processes.

Styrian Health Map – Geospatial Logistic Analysis (Apr 2023 – Jun 2023)
Applied multinomial logistic regression and geospatial modeling to analyze regional hypertension risk patterns based on ESC/ESH clinical guidelines, integrating demographic and health datasets for predictive mapping.

Optimal Designs for Emax, Log-linear, and Exponential Models (Nov 2022 – Mar 2023)
Simulated D-optimal and ED-optimal experimental designs to improve robustness of dose-response modeling under parameter uncertainty.

Windmills Energy Production Analysis (Dec 2021 – Jun 2022)
Implemented time-series regression with anomaly detection and data augmentation to forecast long-term windmill efficiency degradation.

Data Generator and Analysis of High-Dimensional Data (Apr 2021 – Jul 2021)
Developed an R-based synthetic data generator and analyzed clustering behavior in high-dimensional subspaces.

Education

M.Sc. in Data Science (Statistics) Nov 2020 – Mar 2025
Technical University of Dortmund, Germany
B.Tech. in Computer Science and Engineering Aug 2013 – Jul 2017
Maharaja Surajmal Institute of Technology, New Delhi (GGSIPU)

Skills

Programming Languages: Python, R, SQL, Java, C++, JavaScript
Statistical Research Methods: Hypothesis Testing, Experimental Design, Multivariate Analysis
Machine Learning: Regression, Classification, Clustering, Time Series, Anomaly Detection
Deep Learning: Neural Networks, CNN, TensorFlow, PyTorch, Keras
Data Engineering: ETL Pipelines, PySpark, Databricks, Docker, Kubernetes
Data Visualization BI Tools: Power BI, Tableau, Dash, Matplotlib, Seaborn, Plotly, ggplot2
Cloud Platforms: Microsoft Azure (CI/CD, Databricks), AWS (S3, Lambda), Terraform
Databases: PostgreSQL, MySQL, Oracle, MongoDB

Achievements

- Awarded “Ace Team” (Q4 2019) – Accenture
- Star of the Batch – HP Training Program, 2016
- Event Head – Technical & Cultural Fest, 2016

Affiliations

Global AI Hub – Community Lead (2022 – Present)
CSR Group, Accenture – Member (2018 – 2019)
Computer Society of India (CSI) – Member (2014 – 2015)