

Praveen B M



 praveenbm1997@gmail.com

 +49 15228266425

 LinkedIn  GitHub

 Koblenz, Germany

PROFILE

Passionate Data Scientist with a Master's in Data Science and 3+ years of experience in data engineering, machine learning, and business intelligence. Skilled in developing forecasting models, data pipelines, and interactive dashboards to support strategic growth and business development. Expertise in data integration, analytics, and predictive modeling to drive decision-making and expansion in dynamic sectors.

EDUCATION

SRH Hochschule Heidelberg,
Master Of Science In Applied
Data Science & Analytics

Oct 2022 – Sep 2024
Heidelberg, Germany

Master Thesis

Specialization: Machine learning & Gen-AI, Advance NLP, Image processing, Data Analysis, Data Engineering, Data Management, Statistical Analysis.

GPA : 1.8

Vidya Vikas Institute of Engineering & Technology,
Bachelor's of Engineering in
Electrical & Electronics

Aug 2015 – Jul 2019 | Mysuru, India
GPA : 7.6/10

LANGUAGES

English ● ● ● ● ●

German ● ● ● ● ●

Kannada ● ● ● ● ●

PROFESSIONAL EXPERIENCE

Data Analyst, Merck KGaA

Jul 2023 – Sep 2024 | Darmstadt, Germany

- Built **automated data pipeline** systems to integrate data from various sources using **Python**.
- Improved **data accuracy** by **10%** through effective preprocessing techniques.
- Performed data analysis to detect **Vulnerabilities, Risk analysis & Incident management**.
- Developed high-performance **data pipelines** using **PySpark & SQL** to optimize data processing efficiency.
- Automated data workflows, **creating 15+ dashboards** and **streamlining ETL processes**.

Software Engineer I, Stats Perform

Sep 2019 – Sep 2022 | Bengaluru, India

- Designed and implemented **ETL pipelines** to integrate data from diverse sources.
- Developed advanced **predictive models** for performance and trend analysis using **machine learning** techniques.
- Created **interactive dashboards** to **visualize KPIs**, enabling **strategic decision-making** for stakeholders.
- Optimized data storage** and **server processes** to enhance performance and reliability.
- Conducted **root cause analysis** to resolve **production issues**, **minimizing downtime**.
- Documented workflows** and **pipeline architecture** to improve collaboration and knowledge sharing.

SKILLS

Programming:

Python, PySpark, SQL, HTML, CSS

● ● ● ● ●

Databases & Data Warehousing:

Oracle, MySQL, PostgreSQL, SQL Server, MongoDB, Azure SQL, Amazon Redshift, S3, DynamoDB, Snowflake, BigQuery

● ● ● ● ●

Data Science:

Data Analysis, Visualization, Statistical Modeling, Machine Learning, NLP, LLMs, Gen-AI, MLFlow, Critical & Analytical Thinking

● ● ● ● ●

Data Engineering:

Data Integration, Modeling, Quality Assurance, Governance, ETL/ELT, Spark, Kafka, Airflow, Databricks, Qlik

● ● ● ● ●

Data Visualization:

Power BI, Tableau, Qlik Sense

● ● ● ● ●

Cloud Technologies:

AWS (SageMaker), Azure (ML), GCP (Vertex AI, Google Analytics), Palantir Foundry

● ● ● ● ●

Version Control:

Git (GitHub, GitLab), SVN, SourceTree

● ● ● ● ●

Other Tools:

Docker, Excel, Power Query, DAX, Power Automate, Jira, Confluence, Bitbucket

● ● ● ● ●