Bhavesh Mahender Jain *Wissenschaftlicher Mitarbeiter*

Neckarsulm, Germany

+4917640476887

bhaveshjain1212@gmail.com

LinkedIn

Summary

Data Scientist with 4+ years of experience developing AI models from classical statistical methods (OLS, time series) to modern Foundation Models and Multi-Agent Systems. Proven expertise in building data pipelines (SAP, APIs), implementing GenAI solutions (LLMs, RAG), and advising cross-functional teams on data-driven strategies. Experienced in agile collaboration (SCRUM).

Work Experience

04/2024 – present
Heilbronn, Germany

Research Associate
Fraunhofer-Institut für System- und Innovationsforschung ISI

- Building an automated trend radar system using forecasting models and Generative AI to detect emerging market trends and evaluate technology readiness. https://jih-trendradar.eu/
- Developing novel competitor analysis tools that automatically parse web data to generate strategic insights on technology adoption and user engagement.
- Translated complex data science research and AI model findings into actionable business strategies for interdisciplinary teams and non-technical stakeholders.
- Tech Stack: **Project Management, Python, Machine Learning, Gen AI**.

05/2022 – 03/2024
Frankfurt, Germany

Data Science Research Assistant
Fraunhofer IML

- Boosted air cargo forecast accuracy by 70% and developed an AI-powered Python app using OpenAI, Cohere, and Llama 2 for real-time cargo insights and sentiment analysis.
- Implemented Natural Language Processing (NLP) and Time Series Forecasting models enhancing accuracy.
- Developed a comprehensive PowerBI dashboard, facilitating informed decision-making and resulting in an improvement in data processing efficiency.
- Utilized Kubeflow for streamlined project workflows, showcasing expertise in Python, machine/deep learning.
- Contributed to the academic field with a Master's Thesis on Time Series Forecasting, demonstrating a profound understanding of the subject.
- Tech Stack: **Python, Kubeflow, Machine Learning, Deep Learning, and PowerBI**.

07/2023 – 12/2023
Dusseldorf, Germany

Data Scientist
KIOTERA GmbH

- Implemented anomaly detection models (Isolation Forest, One-Class SVM, Autoencoders) to optimize predictive maintenance.
- Developed internal company dashboard to record and monitor internal KPIs.
- Enhanced equipment reliability and significant cost savings for the customers.
- Technolgies used - **Python, MLFlow, Machine Learning, Python libraries for ML.**

07/2022 – 06/2023
Berlin, Germany

Data Analyst
SAP

- Successfully transitioned from an internship to a working student role, showcasing adaptability and unwavering commitment.
- Designed and implemented impactful SAP Analytics Cloud Digital Boardroom dashboards for both internal and external projects, resulting in improvement in data visualization effectiveness.

- Contributed to data-driven decision-making processes by integrating Machine Learning into SAC planning, leading to increase in decision accuracy.
- Specialized in crafting HR and Finance analytics dashboards tailored to address specific departmental needs.
- Enhanced overall decision-making by integrating key performance indicators (KPIs) into comprehensive data-driven solutions, contributing to a significant improvement in strategic insights.
- Tech Stack: **SAP Analytics Cloud, Data Warehouse Cloud, Airtable, SQL, Microsoft Power BI, S/4HANA**.

05/2022 – 10/2022
Berlin, Germany

Data Science Research Assistant
Fraunhofer FOKUS

- Improved overall workflows and operations by 25% through the adoption of advanced MLOps methodologies.
- Created systems to keep track of metrics, parameters, and files for the IML4E project, making it 30% more efficient to monitor machine learning models.
- Conducted thorough data quality assessments utilizing ISO25012/25024 metrics to validate the integrity of input data for machine learning models, resulting in a significant 20% enhancement in performance.
- Tech Stack: **MLflow, Kubeflow, Python**.

02/2022 – 05/2022
Munich, Germany

Data scientist
IMSP - Institut für Marktforschung, Statistik und Prognose

- Achieved expertise in predictive model development for customer behavior using R and Python.
- Conducted impactful ad-hoc data analyses to derive actionable insights.
- Demonstrated proficiency in automating reports and processes to enhance operational efficiency.
- Programmed custom functions for streamlined task execution.
- Tech Stack: RStudio, R packages, Python.

10/2021 – 03/2022
Dortmund, Germany

Data Science Research Assistant
TU Dortmund

- Instructed 'Big Data Analytics' course as a Tutor, facilitating the application of machine learning algorithms through Python and providing guidance on coursework assignments.
- Served as a Research Assistant in the Business Informatics department, contributing to data mining projects.
- Tech Stack: **Microsoft Azure Machine Learning Studio, Python**.

08/2019 – 12/2019
Hyderabad, India

Quality Specialist
Amazon India

- Developed and implemented Python automation scripts, achieving a 50% reduction in testing time and proficiently utilizing Python, NumPy, and Pandas for validating and analyzing extensive datasets.
- Collaborated seamlessly with cross-functional teams, including Product Managers, Developers, System Administrators, and Database Administrators, ensuring smooth updates in analytics, software releases, and timely client deliverables.
- Tech Stack: **Python, NumPy, Pandas**.

Education

04/2020 – 10/2023
Potsdam, Germany

Master of Science in Data Science
University of Potsdam
Thesis: "Enhancing Air Cargo Supply Chain Performance Through Accurate Demand Forecasting: A Machine Learning Approach"

09/2014 – 07/2019 **Bachelor of Technology in Computer Science Engineering**
Hyderabad, India *Sreenidhi Institute of Science and Technology*
Thesis: "Domain Ontology based Semantic Search for Efficient Information Retrieval through Automatic Query Expansion"

05/2017 – 05/2019 **Technology Entrepreneurship Programme**
Hyderabad, India *Indian School of Business*

Skills

Python

SQL

Machine Learning

Scikit-Learn

NumPy

Computer Vision

Operations Research

Predictive Analytics

Ggplot2, Dplyr, Plotly, Mlr3

DWC

SAP BI

Time Series Analysis/Prediction/Forecasting

Java

Project Management
Google Project Management Certificate

Digital marketing

Statistical Analysis
Statistical modeling
Hypothesis testing
A/B testing
Experimental design

Version Control
Git
GitHub
GitLab

Programming Languages
Python
R

Database Management
MySQL
PostgreSQL
MongoDB
Oracle
SAP HANA

Communication Skills
Ability to explain complex concepts to non-technical stakeholders
Data storytelling
Presentation skills

Time Series Analysis

R

Machine Learning & Deep Learning
Scikit-learn (for machine learning in Python), TensorFlow, Keras, PyTorch

Pandas

Matplotlib

Seaborn

Natural Language Processing (NLP)

Artificial intelligence

Azure ML

Recommender System

SAC

PyTorch

Microsoft Power BI

Data Analysis

Tableau

Data Manipulation and Analysis
Pandas (Python library)
NumPy (Numerical Python)
SQL (Structured Query Language)
MySQL, NoSQL

Data Visualization
Matplotlib
Seaborn
Plotly
Tableau
PowerBI
Qlik Sense

Cloud Platforms
AWS (Amazon Web Services)
Azure (Microsoft)
GCP (Google Cloud Platform)

Natural Language Processing (NLP)
NLTK (Natural Language Toolkit)
SpaCy
Transformers (Hugging Face)

Optimization Techniques
Linear Programming
Integer Programming

Collaboration Tools
Jupyter Notebooks

Languages

German	English
Learning A2	B2

Projects

Walmart sales prediction
- Conducted in-depth analysis on markdown impact on sales, considering variables like fuel prices, temperature, unemployment, and CPI.
- Applied Python for simple and multiple linear regression models, enhancing sales forecasting accuracy.

Ontology-based information retrieval system
- Implemented optimized data warehousing strategy for an 80% improvement in retrieval speed.
- Achieved a 50% increase in retrieval accuracy by leveraging Java as the interface and Ontology Web Language (OWL) for efficient execution.

Prediction of used car prices
- Employed AIC and BIC for optimal variable subset selection in dataset analysis.
- Analyzed models based on lowest BIC values, emphasizing coefficient interpretation, statistical significance, confidence intervals, and goodness of fit for the best feature subset.

Resistor's price prediction
- Applied ML algorithms for price prediction using Azure ML Studio.

Creditworthiness of customers
- Used ML algorithms to predict creditworthiness of the customers for Bank(client).
- Executed steps from cleaning of dataset, feature engineering, deploying ML algorithms to communicating the results.

Quit Instagram: How social media controls our lives
- Led targeted research on Instagram's impact on daily life, analyzing 200 survey responses.
- Categorized data into clusters, including negative posts, social factors, and psychological elements.
- Highlighted the substantial influence of unimportant posts, resulting in 36.7% unfollowing and 41.3% muting actions on Instagram.
- Identified additional factors such as disagreement and inactivity, enriching the overall study insights.

Artifact logging using MLFlow
- Implemented artifact logging with MLflow in a project for enhanced reproducibility and tracking of machine learning experiments.
- Strengthened development and analysis processes through streamlined artifact management.

Vibration Anomaly Detection using ML
- Enhanced production planning efficiency by 85% through the implementation of ML algorithms.
- Utilized MLflow for anomaly detection on vibration data, optimizing process reliability and minimizing disruptions.

Eurostat Insights Dashboard
- Developed and implemented a Power BI dashboard to extract actionable insights from Eurostat data.
- Streamlined data visualization, enabling intuitive interpretation and decision-making.
- Enhanced data-driven decision processes by presenting key Eurostat metrics in an accessible and user-friendly format.

Publications

2026 **From Black Box to Glass Box: Designing Transparent Multi-Agent Systems for Enterprise Users**
Accepted at HCI International 2026 (*to appear*)

2025 **Predicting Tech Readiness through Bibliometric Analysis using Unsupervised Machine Learning**
International Society for Professional Innovation Management (ISPIM Conference) 2025

2025 **Multi-agent Retrieval-Augmented Generation for Enhancing Answer Generation and Knowledge Retrieval**
Progress in Artificial Intelligence. 24th EPIA Conference on Artificial Intelligence, EPIA 2025. Proceedings. Part II

2025 **Fallstudienbasierte Analyse der Integration generativer KI-Werkzeuge in die strategische Vorausschau.**
Symposium für Vorausschau und Technologieplanung

Declaration

I hereby declare that the above-mentioned information is accurate to the best of my knowledge and belief.

Bhavesh Mahender Jain
Heilbronn, 30.04.2026