Dhanunjaya Elluri Thimmaraju

Gerwigstr. 56, 76131 Karlsruhe
+49-1782652888 dhanunjayet@gmail.com linkedin.com/in/dhanunjaya-elluri github.com/Dhanunjaya-Elluri

PROFILE SUMMARY

Data Scientist and Machine Learning Engineer with over 5 years of applied data science, machine learning, and software development experience. Advanced knowledge of software design patterns, fluent in Python, strong in object-oriented and functional programming paradigms, and in technical writing. Excels at applying AI/ML theory to commercial development in high-volume data processing, visualization, and model development.

SKILLS

Languages:	Python, R, SQL, C/C++, Shell Scripting
Cloud/DevOps:	Azure, AWS, Databricks, Docker, Kubernetes, Git(CI/CD), Terraform
Database:	MySQL, NoSQL, PostgreSQL, MongoDB, Redshift, Cassandra, BigQuery
BI/ETL Tools:	Power BI, Tableau, Airflow, Kafka, Azure Data Factory, Grafana
Frameworks:	PyTorch, TensorFlow, MLflow, Flask, Langchain, LlamaIndex, HuggingFace, Scikit-Learn, NLTK, Pytest

WORK EXPERIENCE

Validaitor UG **January 2024 – Present**
Machine Learning Engineer *Karlsruhe, Germany*

- Mentored a team of 3 junior developers. Participated in code reviews, resulting in a 30% increase in milestone completions and a 20% decrease in coding errors.
- Built adversarial attack pipelines for testing LLM models to mitigate bias, fairness, privacy, and other safety measures.
- Optimized autoscaling mechanism of celery workers for sending millions of prompt requests, reducing latency by 70% and cutting ECS costs by approx. €9000 per month.
- Designed and built an efficient ETL pipeline in AWS using Airflow & Spark to validate, and transform 2 TB of data.
- Transformed existing Validaitor LLM platform into GDPR, ISO 27001 complaint, and managed platform infrastructure in AWS and Azure using Terraform, delivering 99.9% uptime.

Munich RE **October 2022 – December 2023**
MLOps Engineer Intern *Munich, Germany*

- Development and operation of in-house MLOps platform for compliant one-click hosting of ML models reducing €0.5M+ operational costs in Q2 and Q3 of 2023.
- Proposed and integrated both Databricks & Dataiku Feature Stores into MLOps platform enabling data versioning, governance, and optimized data pipelines.
- Built an NLP platform for automating the underwriting process, increasing time to delivery of Claims and Underwriting use cases by 70%.
- Built and integrated company-specific RAGs with re-ranking in Databricks enhancing policy issuance processes with an 80% improvement in efficiency and accuracy.

Technische Universität Dortmund **August 2021 – October 2022**
Research Assistant *Dortmund, Germany*

- Conducted comprehensive statistical analysis and applied clustering and real-time anomaly detection on time series sleep data successfully diagnosing sleep disorders and sleep apnea in 1000+ patients.
- Research and development of advanced pruning techniques to select the best subset of a trained ensembles to minimize memory footprint and maximize accuracy.
- Developed a content and collaborative-based recommendation system for TU Dortmund Bibliothek, improving the recommendation of similar e-books and e-journals from extensive collections.

Tata Consultancy Services Ltd **June 2018 – May 2021**
Data Scientist *Bangalore, India*

- Led client consultations, transforming complex datasets into actionable strategies, culminating in a 70% increase in customer satisfaction.
- Created an information retrieval pipeline from PDFs using Graph Neural Networks, and BERT for text analysis.
- Saved ∼500 hours in document processing by creating an efficient NER API using Python, TensorFlow, and MLflow, automating document annotation.
- Optimized supply chain operations with a cold start problem with DeepAR model resulting in a 45% improvement in inventory turnover, and a 25% decrease in stockouts.
- Conducted strategic A/B tests for one of Fortune 500 client's platforms, resulting in a 30% uplift in conversion rates, and a 20% rise in average order value.
- Designed and executed ETL pipelines with 1M+ transactions every day and created Power BI dashboards for enhanced monitoring and reporting.

EDUCATION

Technische Universität Dortmund **April 2021 – March 2024**
M.Sc. in Data Science *Dortmund, Germany*

Sri Siddhartha Insitute of Technology **August 2014 – June 2018**
BE in Computer Science *Tumkur, India*

RELEVANT PROJECTS

RAG Testing Framework **Validaitor**

- Developed a RAG testing framework to evaluate LLM applications with both lexical and semantic based metrics.
- Allows to create custom text data generations with configurable metrics.
- Generates comprehensive test results and benchmarks to rank the best LLM application.

Transformers for Quantized Time Series Forecasting Link **Master's Thesis**

- Implemented PAA, SAX, and kernel-SAX for time series dimensionality reduction, optimizing data for in-depth analysis and forecasting.
- Evaluated diverse transformers on quantized series, identifying Poisson distribution trends with a lambda of 1.7 in prediction discrepancies.
- Demonstrated a 15% reduction in mean squared error (MSE) compared to traditional models, further supporting the effectiveness of the applied techniques.

Hosted Model Environment (HOME) **Munich RE**

- Developed an in-house MLOps product, facilitating the seamless deployment of ML models to Azure ML as scalable endpoints.
- Enabled one-click, secure and compliant with APIM integration, ensuring rapid updates and operational readiness.
- Integrated Databricks and Dataiku feature store, and comprehensive monitoring capabilities, including model tracking, and data drift analysis.

Demand Forecasting: Supply Chain Management **TCS**

- Managed and optimized a data enrichment pipeline, conducting extensive Exploratory Data Analysis (EDA) on over 2 TB of supply chain data using SQL, Power BI and Python for informed model training.
- Extracted actionable business insights for dynamic pricing and inventory management, and implemented Amazon's DeepAR multivariate forecasting model to project future market trends and demand.
- The insights and forecasts derived from the model directly supported strategic decisions, resulting in a a 45% improvement in inventory turnover, and a 25% decrease in stockouts.

A/B Testing Python Library **TCS**

- Developed a lightweight Python library for streamlined A/B testing, allowing data scientists and analysts to efficiently design, execute, and analyze A/B tests within their projects.
- Implemented user-friendly APIs for randomization, hypothesis testing, and result visualization, simplifying the entire A/B testing process and reducing implementation time.
- Utilized across 100+ projects within the organization, enabling data-driven decisions.

LANGUAGES

English: C1 (Fluent)
German: A2 (Improving)
Telugu & Hindi: C2 (Native)