

Dheeraj Kadadalamari Bayareddy

Data & Backend Engineer | MSc Data Science | Aspiring Machine Learning Engineer

Berlin, Germany | +49 175 7094332 | dheerajbayareddy@gmail.com | [linkedin.com/in/dheerajreddykb](https://www.linkedin.com/in/dheerajreddykb) | github.com/dheeraj-bayareddy

PROFILE

MSc Data Science student transitioning into machine learning after 3+ years as a professional software engineer on a live insurance platform. Brings production engineering discipline API design, database architecture, and shipping reliable software under real deadlines to applied ML work spanning deep learning, classical ML, and model evaluation. Hands-on with Python, TensorFlow/Keras, Scikit-learn, and Large Language Models through self-directed projects in medical imaging and healthcare analytics. Looking to combine strong software engineering experience with data and ML skills in data-driven and machine learning roles in the European market.

SKILLS

Data & Machine Learning:	Python, Pandas, NumPy, Scikit-learn, TensorFlow/Keras, deep learning (CNNs, transfer learning), feature engineering, classification, clustering, model evaluation, LLMs (learning)
Databases & Data Engineering:	SQL Server, MySQL schema design, stored procedures, query optimization, data pipeline design
Backend & APIs:	C#, ASP.NET Core, RESTful API design, microservices
Tools & Collaboration:	Jira, Git, Postman, structured logging, Agile/async remote collaboration

DATA / ML PROJECTS

Diabetic Retinopathy Severity Grading (Deep Learning)

2026

Python · TensorFlow/Keras · Gradio

- Built a 5-class DR severity grading system on the APTOS 2019 dataset (3,662 retinal images); trained and compared a custom CNN, ResNet50, and MobileNetV2 (transfer learning).
- Best results: ResNet50 QWK 0.67, custom CNN 67.9% accuracy, MobileNetV2 0.89 macro-AUC; used Ben Graham preprocessing, class weights, and Grad-CAM; deployed as a 3-model ensemble in a live Gradio demo.

Healthcare Fraud Detection (Machine Learning)

2026

Python · Scikit-learn · Pandas

- Compared 6 classification algorithms (Logistic Regression, Decision Tree, KNN, Random Forest, Gradient Boosting, SGD) on a synthetic health-insurance claims dataset; Random Forest gave the strongest fraud/legitimate separation.
- Feature analysis identified provider-patient distance and procedure count as the strongest fraud indicators — directly usable for a rules-based pre-screen.

Multiclass Pathology Classification in Voice Signals

2022

Python · Deep Learning

- Built an end-to-end audio-classification pipeline (preprocessing, feature extraction, model training, evaluation) across multiple pathology classes.

WORK EXPERIENCE

Software Engineer, C#/.NET Backend

Nov 2022 – Feb 2026

Paktolus (client: Bolt Insurance), Mumbai, India — Remote

- Built and maintained REST APIs and SQL Server data logic end-to-end for 3 insurance platform modules with no dedicated data-layer owner, including schema design and query optimization on large, time-sensitive datasets.
- Automated manual batch processes (policy renewals, document generation) via Windows Services, using structured logging for systematic root-cause debugging.

Full-Stack Web Development Intern

Sep 2021 – Oct 2021

Tequed Labs Pvt. Ltd., India

- Delivered backend and frontend tasks independently in a 1-month internship — server-side logic, API integration, and database interactions.

EDUCATION

MSc Data Science
University of Europe for Applied Sciences, Potsdam, Germany
Coursework: Machine Learning · Marketing Analytics · Data Engineering · Cloud Computing · Data Visualization

Mar 2026 – Present

Bachelor of Engineering (B.E.), Computer Science & Engineering
RajaRajeswari College of Engineering, Bangalore, India
Coursework: Python · SQL · Java · C · DSA · Databases · Software Engineering · Final-year project: voice-signal pathology classification (deep learning)

2018 – 2022

LANGUAGES

English (Professional Proficiency) · German (A2, learning)

INTRESTS

Applied Machine Learning · Data-Driven Strategy · LLMs & Generative AI · Continuous Learning