PRATIK RUGHE

Studentendorf Schlachtensee, Wasgenstr 75 | Berlin, Germany
+49-15510407852 | pratek.aiml@gmail.com

PROFILE SUMMARY

Student Tutor - Assistant | AI-Native Networks, Machine Learning & Systems Research

Master's student in Data Science with a strong foundation in Computer Engineering and applied artificial intelligence. Research interests include Computer Vision, Large Language Models, Transformer-based architectures, and scalable deep learning systems. Experienced in Python-driven ML workflows, PyTorch-based model development, GPU-accelerated training, and experimental evaluation of AI models. Motivated to contribute to applied and fundamental research at the intersection of vision, language, and real-world intelligent systems, with a focus on robust, interpretable, and deployable AI solutions.

EDUCATION

UNIVERSITY OF EUROPE OF APPLIED SCIENCES — Potsdam, Germany
Master of Data Science - (GPA - 1.74) — *Sep 2024 – Current*

Focus *- Artificial Intelligence research with emphasis on Computer Vision, Large Language Models, Transformer architectures, Representation learning, and Applied deep learning systems.*

INDIAN INSTITUTE OF TECHNOLOGY - HYDERABAD — Hyderabad, India
Professional Certification in Artificial Intelligence & Emerging Technology - (Grade - A+) — *June 2022 – Dec 2022*

Focus - *Python for AI and Scientific Computing, Machine Learning & Statistical Modeling, Natural Language Processing (NLP), Computer Vision & Object Detection (YOLOv8), Signal Processing for AI Applications, Deep Learning with PyTorch, Large Language Models & Transformers,*

SAVITRIBAI PHULE PUNE UNIVERSITY — Pune, India
Bachelor of Computer Engineering - (GPA - 1.8) — *Aug 2017 – Jul 2021*

Focus - *Programming, Advanced Data Structures, Databases, Mathematics, and Statistics, Web Development, Cyber Security, Digital Electronics, Microprocessor.*

PROFESSIONAL EXPERIENCE

STUDENT TUTOR - DATA ANALYTICS — Potsdam, Germany
Student Tutor - Data Analytics — *Apr 2025 – Present*

- Supported 150+ bachelor's and master's students in applied machine learning, data analytics, and model-based reasoning, with focus on experimental design and result interpretation
- Conducted 20+ tutorial and lab sessions covering data preprocessing, exploratory analysis, basic statistical learning, and introductory ML concepts, using Python-based workflows
- Supervised and reviewed 50+ student projects and assignments, providing research-oriented feedback on method selection, feature design, model evaluation, and clarity of experimental results
- Guided students through end-to-end ML project pipelines, including problem formulation, data cleaning, baseline modeling, and structured result reporting
- Introduced foundational concepts in Computer Vision and Large Language Models, such as feature extraction, embeddings, and model limitations, within project guidance
- Contributed to measurable improvements in methodological rigor, reproducibility, and presentation quality of student project submissions.

ACCENTURE PVT LTD — Pune, India
Data Analyst — *Jan 2022 – Sept 2024*

- Worked with large structured datasets (100k+ records) to support analytical evaluation, reporting, and data-driven research tasks
- Performed systematic data validation, cleaning, and preprocessing using Python (Pandas, NumPy) and structured Excel workflows, ensuring data quality and reproducibility.
- Supported 5–7 concurrent analytics and ML-related projects, contributing to data preparation, exploratory analysis, and result consolidation.
- Assisted in the preparation of 10+ internal analytical reports and technical summaries, translating quantitative findings into structured insights for decision-making.
- Designed and optimized automated data preparation workflows, reducing manual processing effort by ~20% and improving consistency across projects.

PYSIST INDIA PVT LTD Pune, India
Data Analyst Intern *Jul 2021 – Dec 2021*

- Designed and implemented an NLP-based chatbot prototype for automated client query analysis and reporting, applying text preprocessing, intent classification, and response generation techniques.
- Improved operational efficiency by ~20% through automation of repetitive query-handling and reporting workflows.
- Collected, processed, and structured 10,000+ unstructured and semi-structured text records from multiple sources to create ML-ready datasets.
- Ensured data quality through systematic cleaning, normalization, and validation, supporting reliable downstream analytics and NLP experimentation.
- Documented data preparation steps and model behavior to support reproducibility and evaluation.

PROJECTS

Transformer-Based Anomaly Detection for Sustainable Systems

Technologies - *Python, PyTorch, NumPy, Pandas, scikit-learn, Transformers, CUDA, Linux*

- Implemented attention-based Transformer models for pattern recognition and anomaly detection in large-scale time-series data
- Designed energy-efficient training pipelines using mixed-precision GPU acceleration
- Evaluated model robustness using AUROC, precision-recall, and cross-validation on highly imbalanced datasets
- Focused on early fault detection to reduce resource waste and improve system sustainability

Vision-Based Intelligent Monitoring using OpenCV and Deep Learning

Technologies - *Python, PyTorch / TensorFlow, CUDA (basic), Linux*

- Developed computer vision pipelines for real-time visual anomaly detection and monitoring
- Integrated OpenCV-based preprocessing with CNN-Transformer hybrid architectures
- Optimized inference through adaptive resolution and frame selection techniques
- Applied models to sustainable infrastructure and environmental monitoring scenarios

Energy-Efficient Large Language Model Adaptation (Academic Project)
Technologies - *Python, PyTorch, Hugging Face Transformers, CUDA, Linux*

- Fine-tuned Transformer-based Large Language Models for domain-specific NLP tasks
- Implemented parameter-efficient training methods to reduce computational and energy cost
- Explored knowledge distillation for lightweight, deployment-ready language models
- Analyzed trade-offs between model performance, scalability, and environmental impact

SKILLS & PERSONAL

Technical Skills

Machine Learning & Artificial Intelligence : Machine Learning, Time-Series Forecasting, scikit-learn, TensorFlow, PyTorch, Generative AI concepts: Large Language Models (LLMs), embeddings, Retrieval-Augmented Generation (RAG) (applied projects)

Data Engineering & Analytics : Python, SQL, ETL / ELT pipelines, Pandas, NumPy, Polars, Containerization, Kubernetes (foundational to intermediate)

Databases : Relational Databases: PostgreSQL, BigQuery (working knowledge), NoSQL Databases

Systems, Containers & Accelerated Computing : Docker, Linux, Git, GPU-accelerated model training (PyTorch / TensorFlow – basic), CUDA fundamentals (basic understanding)

Big Data & Distributed Processing : Apache Spark / PySpark (basic to intermediate), Data platforms and large-scale analytical pipelines

MLOps & DevOps : Model lifecycle management, experiment tracking, CI/CD concepts, version control (Git), DevOps platforms

Networking Foundations : Fundamentals of 5G/6G architectures, Understanding of modern wireless communication systems

Languages : English – C2 (Fluent), German – Good knowledge (actively improving)

Driven by a polymath mindset, I explore ideas across machine learning, systems, and data to build innovative, responsible solutions with real-world impact.