Rachael Burris

+49 162 4552093 · Dusseldorf, DE · rachaelburris15@gmail.com · linkedin.com/in/RachaelBurris

SUMMARY

Machine Learning Engineer with 5 years of experience developing, deploying, and optimizing ML and generative-AI systems. Skilled in Python, PyTorch, and cloud-based model deployment with a focus on responsible, scalable, and fair AI applications. Adept at transforming data-science prototypes into production-ready models through CI/CD and close collaboration with product and engineering teams. Passionate about advancing equitable technology that improves decision-making and expands access to opportunity.

KEY ACHIEVEMENTS

- **Model Deployment:** Built and deployed production-ready ML models using PyTorch and scikit-learn, incorporating CI/CD automation for stable releases and reproducible workflows.
- **Cloud Engineering:** Adapted analytics infrastructure to cloud platforms (GCP/Azure, AWS-equivalent architecture) supporting model training, inference, and API integration at scale.
- **Generative AI Systems:** Designed retrieval-augmented and prompt-chained LLM applications integrating Hugging Face models and vector databases to deliver context-aware responses for users.
- **Deep Learning Research:** Developed a Siamese neural network for few-shot speaker identification and keyword spotting on edge devices, improving inference accuracy with optimized architectures.
- **Ethical AI Collaboration:** Partnered with cross-functional teams to embed fairness and transparency principles into model design and data-quality reviews, aligning technical outcomes with user impact.

PROFESSIONAL EXPERIENCE

Data Engineer, Director of Data Analytics | Muscular Dystrophy Association | 2020 - September 2025

- Develop and deploy ML and generative-AI models supporting analytics and research applications across clinical and operational domains.
- Implement retrieval-augmented LLM systems with Python, PyTorch, and Hugging Face; integrate with Streamlit front-ends for interactive insights.
- Establish CI/CD workflows to automate model testing, packaging, and release, reducing manual deployment time and improving version control.
- Collaborate with engineers and data scientists to translate prototypes into scalable, production-ready solutions that meet compliance standards
- Drive cross-team data governance initiatives ensuring fairness, transparency, and reproducibility in modeling and data reporting practices.

Teaching Assistant (Part Time) | University of Chicago | 2022-2023

- Supported graduate instruction in machine learning, model evaluation, and deep learning workflows using PyTorch and GCP.
- Mentored students on ML lifecycle best practices, reproducibility, and communicating technical results effectively.

Director of Clinical Support | Muscular Dystrophy Association | 2018-2020

- Created national data standards and reporting dashboards for healthcare programs.
- Led data visualization initiatives translating operational data into actionable insights for program improvement.

Associate Director of Clinical Support | Muscular Dystrophy Association | 2016–2018

- Produced data-driven insights for leadership and external partners using SQL and Python.
- Ensured accuracy and consistency across datasets, enabling longitudinal tracking and analysis.

SKILLS

- **Languages:** Python, SQL
- **AI & ML Frameworks:** PyTorch, TensorFlow, Scikit-learn, Hugging Face, LlamaIndex
- **Generative AI & LLMs:** RAG Systems, Prompt Engineering, Embeddings, Vector Databases (FAISS)
- **Data Engineering:** ETL (Talend, Qlik), Tokenization, Feature Engineering, CI/CD
- **Cloud & DevOps:** GCP, Azure, Docker
- **Visualization & BI:** Tableau, Matplotlib, Seaborn, Plotly
- **Other:** Agile Workflows, Cross-Functional Collaboration

PROJECTS AND PUBLICATIONS

- **Survival Analysis of Combinatory Treatments in Neuromuscular Disease**
  Developed Cox proportional hazards models to evaluate survival outcomes under multiple treatment combinations. Controlled for ambulation status, treatment type, and age of diagnosis to identify key predictors of extended survival.

- **Speaker Identification & Few-Shot Keyword Spotting for Edge Devices**
  Built Siamese neural networks for audio recognition and speaker verification using Google Speech Commands and VoxCeleb datasets. Optimized AI model for low-power edge computing devices through dimensionality reduction (PCA) and distance metric learning.

EDUCATION

Master of Science in Applied Data Science | University of Chicago, Illinois, US

- ***Recipient of Capstone Award for Best in Show***
- Relevant courses: Data Engineering Platforms, Statistical Analysis, Data Mining, Linear Non-Linear Models, Machine Learning, Time Series Analysis, Big Data platforms

Bachelors of Arts in Psychology | University of Washington, Washington, US

- Emphasis in genetics, childhood neurodevelopment