Ben Clement Kishore Yadala

benjaminyadala@gmail.com | LinkedIn
(541) 250-6522 | Littleton, Colorado

Professional Summary

Data and Machine Learning Engineer with a Master's degree in Computer Science and experience building scalable data pipelines and applied ML systems, measured by production reliability and downstream analytics adoption, by combining Python, SQL, cloud platforms, and ML workflows across structured and unstructured data.

Technical Proficiencies

- **Programming and Querying:** Python, SQL
- **Data Engineering and Systems:** Data pipelines, ETL, ELT, data modeling, data warehousing, data quality, data validation, monitoring, troubleshooting
- **Cloud Platforms and Orchestration:** Google Cloud Platform, BigQuery, Cloud Storage, workflow orchestration, Airflow, Cloud Composer, AWS
- **Machine Learning and AI:** Applied machine learning, natural language processing, large language models, retrieval-augmented generation (RAG), LoRA fine-tuning, prompt engineering, feature engineering, model evaluation
- **ML and Data Libraries:** Scikit-learn, PyTorch, Hugging Face Transformers, FAISS
- **Analytics and Collaboration:** Exploratory data analysis, Tableau, Looker, stakeholder communication, Agile, cross-functional collaboration

Professional Experience

Sports Media Inc. – Remote **August 2025 – Current**

Data Science Intern

- Built cloud-native data pipelines supporting analytics and ML use cases, measured by stable ingestion and validated datasets, by integrating external APIs into analytical data stores using Python and SQL.
- Developed NLP workflows for sentiment analysis and LLM-assisted processing, measured by improved insight extraction from unstructured text, by applying tokenization, embeddings, and prompt engineering techniques.
- Performed exploratory data analysis and feature preparation, measured by faster iteration on predictive use cases, by cleaning, transforming, and validating large-scale structured and unstructured datasets.
- Delivered stakeholder-facing analytics assets, measured by adoption in business reviews, by translating ambiguous questions into metrics, dashboards, and data-ready outputs.
- Improved reliability of analytics workflows, measured by reduced data issues, by implementing validation checks, retries, and monitoring.
- Earned the **Outstanding Performance Award**, measured by cross-team recognition, by owning end-to-end execution and delivery quality.

Cognizant – Hyderabad, India **July 2022 – April 2023**

Data Engineer

- Built and maintained production-grade data pipelines, measured by consistent downstream availability, by implementing ETL workflows using Python and SQL on cloud data platforms.
- Prepared feature-ready datasets for analytics and ML teams, measured by reduced rework, by applying data transformations, validation rules, and quality checks.
- Optimized pipeline performance and stability, measured by improved execution reliability, by troubleshooting failures and tuning queries.

- Supported analytics and data science use cases, measured by on-time delivery, by collaborating with stakeholders to translate requirements into data models.
- Operated within regulated enterprise environments, measured by compliance adherence, by following strict data quality and operational standards.

Cognizant – Hyderabad, India **Feb 2022 – July 2022**

Data Engineer Intern

- Executed SQL-based data extraction and transformation workflows, measured by accurate analytics inputs, by preparing query-ready datasets.
- Validated and documented datasets, measured by reduced downstream issues, by performing data checks and testing.

Projects

Retrieval-Augmented Generation System (EB5 KnowledgeRAG)

- Designed and implemented an end-to-end RAG system, measured by improved response relevance, by combining vector search with LoRA fine-tuning on a transformer-based language model.
- Improved factual accuracy of generated responses, measured by qualitative evaluation, by implementing context-grounded prompting and embedding-based retrieval.
- Refined model outputs through experimentation, measured by reduced hallucinations, by performing iterative error analysis and prompt tuning.

Computer Vision Road Damage Detection (YOLOv8)

- Built an object detection pipeline, measured by balanced precision and recall, by training and evaluating a YOLOv8 model on labeled image data.
- Improved model performance and efficiency, measured by validation metrics, by tuning training parameters and data preprocessing steps.

Cloud-Based NLP Pipelines

- Developed cloud-hosted NLP workflows, measured by successful text processing at scale, by designing modular pipelines for analysis and experimentation.
- Evaluated latency and quality tradeoffs, measured by comparative outputs, by testing multiple prompt and model configurations.

Education

Oregon State University, Corvallis, OR **June 2025**

Master of Engineering in Computer Science

GPA 3.6 / 4.0

CVR College of Engineering, Hyderabad, India **May 2022**

Bachelor of Technology (BTech) in Mechanical Engineering

GPA 8.11 / 10

Certifications & Achievements

- Databricks Fundamentals Accreditation
- HackerRank SQL Intermediate Certification
- Outstanding Performance Award, Sports Media Inc.