# Dhanavikram Sekar

+1(720)481-1463 | Boulder, CO | dhanavikram2000@gmail.com | GitHub | LinkedIn | Portfolio

## Education

**M.S. in Data Science**, University of Colorado Boulder — August 2023 – May 2025
**Courses taken**: Datacenter Scale Computing, Database Systems, Machine Learning, Deep Learning

**B.E. in Electronics and Communication Engineering**, Anna University — August 2017 – May 2021
**Relevant Courses**: Data Structures and Algorithms, Communication Networks

## Technical Skills

**Programming Languages**: Python, Go/Golang, SQL, BASH, Rust, R, Java, C, C++, HTML, CSS, Javascript
**Tools/Technologies**: Git, Docker, Kubernetes, Helm, Apache Kafka, gRPC, Terraform, Ansible, Nginx, MLflow, PowerBI, Databricks, Openstack, Apache Airflow, Unix/Linux(Systemd, Ping, Traceroute, Nmap), SonarQube
**Frameworks**: Apache Spark, Flask, Django, FastAPI, Polars, Dask, Ray, Pytest, PyTorch, TensorFlow, Streamlit, Langchain
**Databases & Storage**: MySQL, PostgreSQL, Neo4j, MongoDB, Cassandra, Snowflake, Amazon Redshift
**Concepts/Practices**: Object Oriented Programming, Infrastructure as Code (IaC), Continuous Integration and Continuous Delivery (CI/CD), Observability, Monitoring, Software Development Lifecycle (SDLC), Agile Methodology
**Certifications**: AWS Certified Cloud Practitioner

## Experience

**Platform Developer** — September 2024 – Present
*University of Colorado Boulder* — *Boulder, CO*

- Built an end-to-end **distributed network analytics platform** using Elasticsearch and Kibana and deployed it across a **10 node cluster**, to evaluate **13** years worth of broadband latency and throughput data.
- Revamped existing pipeline using **polars** and vectorized logic, reducing runtime from **30 hours** to **5-10 minutes**.
- Optimized data storage by **75%** and improved I/O speed by **80%** by transitioning from CSV to Parquet format.

**Data Engineer Intern** — May 2024 – August 2024
*Navajo Transitional Energy Company* — *Broomfield, CO*

- Built **ELT data pipelines**, transferring 500,000+ records from on-prem SQL Server to Azure Fabric OneLake.
- Developed **10+** executive dashboards in **PowerBI** to convey information at a glance for decision making.
- Automated web scraping of coal shipment data using Python, eliminating **5+ hours** of manual weekly tasks.

**Software Engineer** — July 2021 – July 2023
*Tata Consultancy Services* — *Chennai, India*

- Designed and implemented **scalable data processing and ML pipelines** with **Apache Spark (PySpark)**, achieving a **50%** reduction in processing times over legacy pandas/sklearn pipeline.
- Developed ML models (XGBoost, Prophet, ARIMA) for demand forecasting, as custom **Spark UDFs** and integrated them as **SparkML** components, enhancing forecasting performance by over 35%.
- Engineered a **RESTful API** for the pipeline using **Django** and tested its functionality using **Postman**.
- Containerized the pipeline using **Docker** and deployed it across a distributed cluster with inter-node networking.
- Built **CI/CD pipelines** with **Jenkins**, integrating automated unit testing using pytest reducing errors by **40%**.
- Implemented **model monitoring** and **model drift detection** using MLflow, Grafana, and Prometheus.

## Projects

**Micro-Demucs: Music Separation as a Service** | *Flask, Redis, Docker, Kubernetes, GCP* — GitHub

- Built a **GCP-based microservices backend** for music separation featuring a **REST API** frontend, **Redis** queuing, MinIO object storage, and Meta's demucs model.
- Containerized the components with **Docker** and orchestrated on separate **kubernetes** (**GKE**) clusters.
- Implemented an **Nginx** ingress controller for public internet access within the GCP environment.

**DocChat: Document Q&A Chatbot** | *Retrieval Augmented Generation, Huggingface, Streamlit* — GitHub

- Built a Mistral-7B based **RAG system** to answer queries from uploaded documents.
- Implemented MiniLM-L6-v2 embeddings from **Huggingface** to convert text to vector embeddings.
- Utilized **FAISS-CPU** for local vector storage and retrieval of document embeddings.