Soyab Karki

Starkville, MS | 662-352-1863 | sk2078@msstate.edu | linkedin.com/in/soyab-karki/ | github.com/SoyabKarki |

Education

Mississippi State University Starkville, MS
Bachelor of Science in Computer Science, Minor in Data Science and Mathematics *Aug. 2023 – May 2027*

- **GPA:** 4.0/4.0
- **Shackoul's Honors College**
- **Relevant Coursework:** Operating Systems, Data Structures & Algorithms, Computer Organization, Parallel Algorithms, Software Development, Data Science, Data Visualization, Machine Learning, AI Robotics

Technical Skills

Languages: Python, Go, C/C++, JavaScript/Typescript, SQL
Backend & Frontend: FastAPI, React.js, Express.js, Node.js, PostgreSQL, Redis, HTML/CSS
Data & MLOps: Airflow, Spark, Kafka, PyTorch, Pandas, Scikit-Learn, LangChain
Cloud & DevOps: Docker, Kubernetes, AWS, Git, GitHub, Prometheus, Grafana

Experience

Data Analytics Research Assistant Aug. 2024 – Present
Mississippi State University *Starkville, MS*

- Developed data pipelines for 300GB+ consumer review datasets using batch processing and Parquet storage, improving data processing speed and storage consumption by 10x.
- Automated filtering of 2,500+ research papers using OpenAI API and custom keyword matching in Python, reducing manual review time from days to minutes.
- Co-authored "Modeling consumer ratings: The trade-off between quality and consumer preference," published in the European Journal of Operational Research (https://doi.org/10.1016/j.ejor.2026.01.001).

Systems Software Engineer Fellow Jan. 2024 – Present
High Performance Computing Collaboratory *Starkville, MS*

- Parallelized 10+ algorithmic problems using MPI in C++, achieving up to 5x speedup over sequential code.
- Deployed Kubernetes using K3s on a lightweight 3-node Linux cluster to prototype HPC workflows.

Data Scientist Intern May 2024 – July 2024
United States Department of Agriculture - Agricultural Research Service *Starkville, MS*

- Developed 6 risk-analysis ML models using Gradient Boosting and Logistic Regression in Python, achieving baseline accuracy of up to 97%.
- Processed 1.5 years of meteorological data from Google Earth Engine using Python API for ML training.
- Presented findings at 3 regional conferences, clearly communicating complex AI workflows to diverse audiences.

Projects

Distributed LLM-As-A-Service | *Kubernetes, FastAPI, React.js, NGINX, Github Actions* Jan. 2025

- Architected a horizontally-scalable microservice platform on Kubernetes, implementing an API gateway to route and load balance requests to distributed LLM workers.
- Developed a full-stack AI chat application using React.js and an event-driven backend, automating deployments using CI/CD pipeline for zero-downtime updates.

Reddit Data Pipeline | *Python, Airflow, PostgreSQL, Docker, Prometheus, Grafana* Aug. 2025

- Built a containerized ETL pipeline that pulls stock discussions from 3 Reddit forums, runs FinBERT sentiment analysis, enriches with news/stock data, and stores in PostgreSQL with materialized views for fast queries.
- Designed real-time dashboards using Prometheus and Grafana to improve system observability.

Resume Radar | *FastAPI, Redis, React.js, Chrome Extension API, Docker* July 2025

- Built a Chrome extension that analyzes resumes against job descriptions using OpenAI API.
- Optimized FastAPI backend with REST APIs and Redis caching, reducing system latency by over 90%.

Leadership

Data Science Club at MSU | *President* Feb. 2024 – Present

- Recruited 50+ members and raised $3,000+ to run datathons and workshops.

ACM SIG-HPC | *Founding Member* Sep. 2023 – Present

- Configured mini HPC clusters using Raspberry Pi; competed in 2 major supercomputing competitions.