SOYAB KARKI

+1 (662) 352-1863 | sk2078@msstate.edu | Starkville, MS, USA | linkedin.com/in/soyab-karki/

EDUCATION

Mississippi State University — **August 2023 - May 2027**
Bachelor's, Computer Science — *GPA: 4*
- Minor in Data Science and Mathematics
- Judy and Bobby Shackouls Honors College
- Relevant Courses: Methods & Tools in Software Development, Data Structures & Analysis of Algorithms, Computer Organization, Data Wrangling, Linear Algebra, Statistical Inference, Probability

SKILLS

Programming Languages: Python, C/C++, JavaScript, TypeScript, SQL, HTML/CSS
Libraries/ Frameworks: Pandas, NumPy, Scikit-Learn, Tensorflow, FastAPI, React.js, Express.js
Data Tools: Airflow, Apache Spark, Apache Kafka, AWS, Databricks, PostgreSQL, Tableau
Developer Tools: Git, Github, Docker, Node.js

PROFESSIONAL EXPERIENCE

Mississippi State University — **Starkville, MS, USA**
Undergraduate Research Assistant — *August 2024 - Present*
- Designed and implemented ETL pipelines for 300GB+ consumer review datasets using streaming, batch processing, and Parquet storage, improving data processing speed 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.
- Developing an R package for clustering visualization and insight extraction to support targeted marketing analysis.

High Performance Computing Collaboratory (HPCC) — **Starkville, MS, USA**
Student Worker — *January 2024 - May 2025*
- Parallelized 10+ algorithmic problems using MPI in C, demonstrating 5x performance gains over sequential code.
- Performed maintenance and assisted in building a $100,000 supercomputing cluster for the SC24 competition team, ensuring system stability and readiness for high-performance benchmarking tasks.

United States Department of Agriculture - Agricultural Research Service (USDA-ARS) — **Starkville, MS, USA**
Data Scientist Intern — *May 2024 - July 2024*
- Extracted and validated 1.5 years of meteorological data from Google Earth Engine using Python APIs and quality-control filters, enabling robust training datasets for machine learning models.
- Collaborated with 5 researchers to implement Gradient Boosting and Random Forest models in Python, improving predictive accuracy by 25% and providing actionable insights to reduce crop contamination risk in field studies.
- Presented research findings at 2 regional conferences, clearly communicating complex ML workflows to academic and industry audiences, and receiving commendations for clarity and practical impact.

LEADERSHIP & EXTRACURRICULARS

Data Science Club at MSU — **Starkville, MS, USA**
President — *February 2024 - Present*
- Coordinate with the executive team to host data science events, including datathons, workshops, and outreach, successfully recruiting 50+ members and raising $3000+ for club events.

Association for Computing Machinery (ACM) — **Starkville, MS, USA**
Founding Member — *September 2023 - Present*
- Build and configure mini supercomputing clusters using Raspberry Pi and MPI to demonstrate parallel processing.