

Arashdeep Mehroke

✉ arashdeepmehroke@work@gmail.com 📞 +1-209-205-6044 🔗 [linkedin.com/in/arashdeep-singh-020398251/](https://www.linkedin.com/in/arashdeep-singh-020398251/) 🏠 GitHub

Profile Summary

Data Scientist & ML Engineer with 3 years of experience in AI, cloud computing, and big data. Pursuing an M.S. in Data Science at Georgia Tech, building on a strong foundation in statistics, exploratory data analysis, and machine learning. Experienced in ML pipelines, cloud-based AI solutions, and spatial data analysis. Skilled in Python, SQL, TensorFlow, PyTorch, and AWS.

Education

Master of Science in Data Science , Georgia Institute of Technology, (Atlanta, GA)	Expected May 2026
Bachelor of Science in Computer Science , University of California, (Merced, CA)	May 2024

Experience

Machine Learning Engineer , Office of Performance and Data Analytics – Stockton, CA	May 2024 – Present
--	--------------------

- Developed and optimized PyTorch-based deep learning models to enhance urban safety initiatives.
- Designed and deployed ETL pipelines using Azure Databricks, improving efficiency and ensuring data integrity.
- Created an interactive dashboard for real-time analytics on city data, enhancing accessibility and decision-making.
- Deployed AI models using AWS SageMaker, S3, and Lambda for scalable, cloud-based inference.
- Built a large-scale feature engineering framework, increasing predictive model accuracy by 15%.

Technical Lead , Western Digital – Merced, CA	January 2023 – May 2024
--	-------------------------

- Utilized Python and OpenCV to generate synthetic silicon wafer images with defects for model training.
- Improved defect classification models by increasing dataset diversity, boosting detection accuracy by 25%.
- Designed and implemented a data lake using S3 for efficient storage and retrieval of image data for model training.
- Optimized distributed model training and hyperparameter tuning with Apache Spark, significantly reducing training time.

Web Application Developer , Sierra Nevada Research Institute – Merced, CA	January 2022 – Dec. 2022
--	--------------------------

- Developed a dynamic front-end application using React and JavaScript, collaborating with end users to refine functionality and user experience.
- Implemented a Django-based backend for data storage, ensuring ACID compliance in the ETL process.
- Designed and deployed RESTful API endpoints to efficiently serve imported raster data to the front-end application.

Data Analyst , Blu-Lite Inc. – Ripon, CA	May 2018 – Dec. 2021
---	----------------------

- Developed regression models to forecast sales and inventory trends, driving a 5% increase in revenue.
- Created dashboards and visual reports with Power BI and Excel, improving data accessibility by 25%.
- Automated routine data reporting processes, reducing manual effort by 20%.
- Created insightful business intelligence dashboards using Power BI and SQL.

Projects

Emotion Detection System	GitHub
---------------------------------	--------

Real-time DL model for facial emotion recognition with 88% accuracy.

Real-Time Theft Detection System	GitHub
---	--------

Edge-based CV system using YOLOv8n and pose tracking to detect theft via live RTSP camera feeds.

Skills

Languages: Python, R, SQL, Java, C++

Technologies: TensorFlow, PyTorch, Docker, Kubernetes, Power BI, ArcGIS

Cloud Platforms: AWS (SageMaker, Lambda, S3), GCP (BigQuery, Vertex AI)

Certifications

AWS Certified Machine Learning Specialty Issued February 2025

IBM Full Stack Software Developer Professional Certificate: Issued August 2023