

Aryan Sharma

✉ aryanraj@umich.edu ☎ 7344892551 📍 Ann Arbor 🌐 Github 🔗 LinkedIn 📁 Portfolio

Education

- | | |
|--|------------------------------------|
| University of Michigan , Masters in Data Science 🔗 <ul style="list-style-type: none">GPA: 3.53/4.0Course Highlights: Machine Learning, Large Language Models, Information Retrieval, Database Management Systems, Data Manipulation and Analysis. | 08/2023 – 05/2025
Ann Arbor, US |
| PES University , Bachelor of Technology in Computer Science and Engineering 🔗 <ul style="list-style-type: none">GPA: 3.53/4.0Course Highlights: Machine Learning, Statistics for Data Science, Cloud Computing, Big Data, Data Analytics.Awards/Honors: Prof. M R Doreswamy Merit Scholarship for excellent academic performance. | 08/2019 – 06/2023
Bengaluru, IN |


Professional Experience

- | | |
|---|------------------------------------|
| Amazon , Software Development Engineer Intern 🔗 <ul style="list-style-type: none">Engineered a distributed locking mechanism using AmazonDynamoDB, preventing 100% of concurrent AWS Batch Scheduler assignments, eliminating job queue blockages and curtailing vCPU calculation errors by 65% during grey failures.Architected operational monitoring metrics via AWS CloudWatch, providing engineering teams with stage level insights for AWS Batch Scheduler's Managers Bootstrap process, slashing average troubleshooting time by 30% and improving system reliability by 5%. | 05/2024 – 08/2024
Seattle, US |
| University of Michigan , Graduate Student Instructor - Statistics 250 - Introduction to Statistics and Data Analysis 🔗 <ul style="list-style-type: none">Facilitated and led interactive weekly discussions focusing on fundamental concepts (hypothesis testing, regression analysis, visualization in R), systematically measuring improvement through quiz and exam performance and achieving a 20% increase in student comprehension.Created, administered, and graded lab assessments using real-world datasets in R and Python; personalized feedback enhanced student concept understanding and improved overall lab assignment performance. | 01/2024 – 05/2025
Ann Arbor, US |
| Genpact , Data Scientist 🔗 <ul style="list-style-type: none">Built interactive dashboards using Tableau and developed customer segmentation models in Python (Scikit-learn, Pandas) to identify churn drivers, enabling targeted customer strategies resulting in 30% lower churn and 25% higher average order value. | 02/2023 – 08/2023
Bengaluru, IN |
| PES University , Teaching Assistant - Statistics for Data Science 🔗 <ul style="list-style-type: none">Created and conducted interactive practical sessions utilizing Python (Pandas, NumPy) and R with real-world datasets for data manipulation, exploratory analysis, and inference.Enhanced student grasp of applied data science concepts, evidenced by a 25% increase in participation and a 15% improvement in exam performance. | 06/2022 – 11/2022
Bengaluru, IN |
| Smarthub.ai , Machine Learning Engineer Intern 🔗 <ul style="list-style-type: none">Developed a Linear Regression model (Python: Scikit-learn, Pandas) to identify optimal positioning angles for robotic arms in automated vehicle paint application for TVS Motors.Validated model predictions through comparative experiments against fixed-angle processes, demonstrating a 98% reduction in paint wastage and decreasing operational costs and environmental impact. | 06/2022 – 10/2022
Bengaluru, IN |

- Extracted and analyzed large-scale user operational logs (MongoDB, PyMongo), constructing a process similarity matrix to identify redundant user tasks.
- Delivered actionable insights that facilitated workflow streamlining, leading to a 30% reduction in operational redundancies and a 20% improvement in team productivity.


Publications

RtTSLC: A Framework for Real-Time Two-Handed Sign Language Translation, 06/2023

Springer, DOI: https://link.springer.com/chapter/10.1007/978-981-99-0769-4_62 

- Pioneered a deep learning framework (CNN, Siamese Networks) achieving 98% accuracy and <100ms latency for real-time, two-handed Indian Sign Language translation; findings documented in **Springer** and presented at an **International Conference**.

Sign Language Translation Systems: A Systematic Literature Review, 10/2022

IGI-global, DOI: <https://www.igi-global.com/gateway/article/311448> 

- Synthesized insights from over 200 peer-reviewed papers on Sign Language Interpretation, pinpointing 5 key research gaps and defining 4 novel methodological avenues; **IGI Global** publication forms a foundational resource to guide future research.

Cardiac anomaly detection models for wearable devices, 10/2021

Student Research Symposium (SRS)

- Architected and validated a deep-learning based architecture (TensorFlow-Keras) for detecting and classifying cardiac anomalies using wearable-collected 2-lead ECG signals, demonstrating 98% predictive accuracy. Presented detailed analytical methodology and evaluation results at the **13th HiPC Student Research Symposium (SRS), part of the 28th IEEE International Conference** on High-Performance Computing, Data, & Analytics.

Projects

GDINOSAUR, 01/2024 – 05/2024

Grounding DINO with Spatial Awareness for REC, EECS 545 Course Project | Winter 2024

- Co-engineered and executed the data pipeline for RefCOCO-3DS, sourcing and preparing 3D object models (30 COCO categories) to generate a 7,000+ image synthetic dataset with Blender.
- Advanced Grounding DINO's Referring Expression Comprehension by fine-tuning with RefCOCO-3DS and curated 2D datasets, achieving a 15% mAP uplift in spatial/non-spatial understanding via Elastic Weight Consolidation (EWC).

ECG Classification, CHIPS (Centre for Heterogeneous and Intelligent Processing Systems) 08/2021 – 12/2021

- Built, trained, and evaluated a deep learning predictive model (Python, TensorFlow-Keras) to classify cardiac conditions from wearable-sourced 2-lead ECG data, achieving 98% classification accuracy, enabling informed wearable-device-based cardiac monitoring decisions.

Skills

Machine Learning: TensorFlow, Keras, PyTorch, Huggingface, Deepspeed, NumPy, Pandas, **Data Engineering**

Tools: Apache Spark, Airflow, AWS data services, **Data Processing and Distributed Workload:** Docker, Kubernetes,

SQL/NoSQL: MySQL, MongoDB, CosmosDB, Cassandra, PyMongo, **Programming:** Python, R, C/C++, C#, Java, Perl,

JavaScript, Kotlin, Go, Typescript, KornShell, SAS, Stata, MATLAB, Unix/Linux, **Cloud Technologies:** AWS, GCP, Azure,

Data Visualisation/Business Insights: Power BI, Tableau, AWS QuickSight, **Front-End, Full Stack, Web**

Development Stack: HTML, CSS, Javascript, React, Vue, Angular, **Big Data:** Hadoop, Apache Spark, Kafka, Kinesis