

Anshumaan Chauhan

achauhan@umass.edu | 321-361-9962 | [linkedin.com/in/anshumaan-chauhan/](https://www.linkedin.com/in/anshumaan-chauhan/) | github.com/Anshumaan-Chauhan02

Education

University of Massachusetts Amherst

2022 – 2024

Master of Science (MS) in Computer Sciences GPA- 3.93/4

- Relevant Coursework: Algorithms for Data Science, Systems for Data Science, Machine Learning, Artificial Intelligence, Natural Language Processing, Reinforcement Learning, Software Engineering, Neural Networks.

Birla Institute of Technology and Science Dubai

2018 – 2022

Bachelor of Technology (BTech) in Computer Sciences GPA- 9.83/10

- Awarded merit scholarship based on my GPA.
- Awarded Bronze Medal for an outstanding academic performance and standing third amongst the batch.

Professional Experience

Bose Corporation, Framingham, MA - Data Engineer CoOp

2024

- Designed and implemented secure event-driven ETL pipelines using AWS serverless components (Lambda, Step Functions, Secrets Manager), achieving a 5% cost reduction through efficient secret consolidation and provisioning automation.
- Optimized data extraction processes for PowerBI Cloud, reducing operational costs by 10% while enhancing data loading into Snowflake, utilizing SQL transformations and Snowpark for seamless Python integration.
- Developed a Streamlit app for JIRA ticket creation, leveraging AWS pipelines for automated data fetching, resulting in a 20% increase in onboarding productivity.
- Created alert and failure monitoring pipelines with a Streamlit dashboard for Snowflake-deployed apps, leveraging telemetry logs to reduce downtime by 25% through daily queries via cron jobs.

Amazon, Amherst, MA – Graduate Student Researcher

2024

- Conducted experiments to enhance small language models (SLMs) with custom reasoning chains, identifying limitations in zero-shot performance.
- Proposed a decoupled three-step process (plan, reason, answer extraction) that improved GSM8K benchmark scores by +0.57 for 3B models and +3.32 for 7B models.
- Developed and tested a Partial Self Consistency (PSC) method for ensemble analysis, demonstrating that PSC outperforms standard self-consistency by up to +1.85 across various reasoning techniques.

Florida Institute of Technology, Melbourne, FL – Machine Learning Researcher

2022

- Analyzed and extracted the representation of the specifications in a subset of English language using Natural Language Processing (NLTK library) and designed a compiler for translating it to AADL.
- Proposed a Double Deep Q-Network for NAS, reducing scalability and time complexity while maintaining 97% of SOTA performance with ~35% fewer layers.
- Developed an unsupervised DBSCAN-based approach for medical image segmentation, achieving Dice Coefficients of 0.784 and 0.88 for kidneys and lungs, while avoiding the need for large, labeled datasets.

Tata Communications, India – Software Engineer Intern

2020

- Developed automated system in Python using Flask, Urllib and requests libraries/frameworks, to improve the customer targeting and user experience based on clicks per second and user heatmap on the website.
- Performed cross functional evaluation and strategy testing along with a team of 5 developers and marketing analysts, to increase the SEO rankings of the websites while reducing hosting costs and marketing spend by >14%.

Projects

Recipe Infusion

2023

- Created a Recipe Infusion framework featuring Recipe Generation and Style Transfer components, enhancing culinary creativity and personalization.
- Fine-tuned the DistilGPT model on combined RecipeNLG and RecipeBox datasets, improving BLEU and Perplexity scores by +0.04 and 24.5 points, respectively, and successfully implemented style transfer for various celebrities using T5-small models.

Skills

Programming Languages: Python | SQL | Java | HTML | CSS | TypeScript | LaTeX

Databases: MySQL | PostgreSQL | MongoDB | DuckDB

Amazon Web Services: EC2 | Lambda | Batch | S3 | DynamoDB | Athena | Step Functions | SNS | SES | SDK | CDK

Applications/Tools: Tableau | Git | Postman | Docker | Airflow | Snowflake | Apache Spark | Linux | Streamlit | Kubernetes

Productivity Software: JIRA | Confluence | Lucidchart

Soft Skills: Problem Solving | Critical Thinking | Time Management | Team Collaboration | Project Management