

# Ajay Therala

☎ (623) 286-1552 | ✉ atherala@asu.edu | 🌐 ajay-therala7 | 🌐 Ajaytherala

## EDUCATION

Arizona State University, Tempe, AZ, United States

August 2023 - May 2025

Master of Science in **Computer Science**

GPA: 4.0/4.0

**Course Work:** Topics in Natural Language Processing, Knowledge Representation & Reasoning, Foundations of Algorithms, Data Mining, Data Visualization, Statistical Machine Learning.

Jawaharlal Nehru Technological University, Hyderabad, TS, India

August 2017 - July 2021

Bachelors of Technology in **Information Technology**

GPA: 9.21/10.0

**Course Work:** C, C++, Java, Data Structures & Algorithms using Python, Java & C++, Business Intelligence, Design & Analysis of Algorithms, Advanced Linux Programming.

## TECHNICAL SKILLS

**Languages :** C, Python, SQL, HTML, CSS, JavaScript, D3.js

**Tools/Technologies :** Machine Learning, Deep Learning, Natural Language Processing, Hadoop, Pyspark, Spark SQL, Git, Django, Flask, Azure ML Studio, AWS, Elastic Search, Generative AI, Streamlit, HDFS, LangChain, OCI

**Database/Server :** SQL Server, MongoDB

## CERTIFICATIONS

- Microsoft Certified: Azure Data Scientist Associate [Verify Here](#)
- Oracle Certified : Oracle Cloud Infrastructure 2024 Generative AI Certified Professional [Verify Here](#)

## PUBLICATIONS

[1] Evaluating Multimodal Large Language Models across Distribution Shifts and Augmentations, Publisher : IEEE/CVF Conference on CVPR Workshops [Paper Link](#)

[2] NetMD- Network Traffic Analysis and Malware Detection, Publisher: IEEE [Paper Link](#)

## PROFESSIONAL EXPERIENCE

Tata Consultancy Services Limited, Hyderabad, India

August 2021- August 2023

*Systems Engineer (ML Developer), Digital Research and Innovation*

- Collaboratively engineered key components of TCS Cognitive Product Support, an Intelligent Domain-Specific Search Engine.
- Crafted Data Lens, a component for training custom NER models, achieving an impressive 80% - 90% accuracy by leveraging Ontology.
- Designed and executed BrickWork to manually annotate entities and relationships from PDF documents, producing ontologies and datasets for training NER and NLU models.
- Demonstrated expertise in Generative AI, Embeddings, and Prompt Engineering by developing advanced GPT-powered bots. Delivered impactful client demos, earning high praise from esteemed clientele.
- Acquired robust web development skills, including web scraping with Selenium and Chrome WebDriver.
- Mastered AWS services, particularly S3 Bucket & Textract, to enhance data storage & automate document processing. Strategically planned & developed a POC for extracting key-value pairs from handwritten forms & streamlined data management.

Tata Consultancy Services Limited, Hyderabad, India

March 2021-June

2021

*Research Project Intern*

**Title:** NetMD - Network Traffic Analysis and Malware Detection.

- Investigated data refinement techniques and assessed ML and DL algorithms on malware detection datasets like NetML, CICIDS2017, and non-vpn2016.
- Achieved a top 5 position in the NetML - Network Traffic Analytics Challenge 2020, surpassing baseline metrics.
- Presented research findings at ICAIIC 2022, sharing key insights with peers and industry professionals.

## PROJECTS

**Text Similarity Model for Answer Validation for Online Learning Platforms**

*Tech stack : Python, NLP, Siamese Networks, LLM*

- Engineered cost-effective automated answer validation systems for e-learning platforms, eliminating manual labor.
- Developed a sophisticated "Text Similarity Model for Answer Validation" utilizing deep learning technologies.
- Implemented a dual validation approach integrating Siamese Networks and Large Language Models for enhanced accuracy.

**Visual Analysis of Sleep Studies**

*Tech stack : HTML, CSS, D3.js, Flask, Streamlit, Python*

- Developed a visual interface for visualizing the sleep cycles and occurrences of apnea events using the Polysomnography (PSG) Data using Streamlit & Flask.

**Analyzing and Mitigating Hallucinations in Multi modal LLM's**

*Tech stack : Python, Instruct BLIP, LLM's*

- Analyzed behavior of Multi modal LLM's such as Instruct BLIP across varied question types.
- Utilized Mistral LLM to curate QA pairs from COCO Captions, for querying the model & detect hallucinated responses generated by the model.
- Focused investigative efforts on count and color-based questions to unveil and document fabricated instances