

# Sherry Thomas

GitHub — Kaggle — LinkedIn — Portfolio  
*Indian Institute of Technology Madras*

## Introduction

Highly motivated Data Science student at IIT Madras (BS in Data Science and Applications, expected 2025) with a passion for Large Language Models (LLMs) and Natural Language Processing (NLP). Proven ability to contribute effectively through research, teaching, and open-source projects. Eager to leverage strong technical skills and a background in philosophy to solve real-world problems.

## Education

Program	Institution	CGPA	Year
BS in Data Science and Applications	IIT Madras	9.94	2024
BA in Philosophy	Christ University, Bengaluru	9.57	2021

## Work Experience

### Data Engineer Intern

AI4Bhārat-IIT Madras (Feb 2024 - Present)

- Contributing to the TTS Lab by cleaning and preparing data for open-source text-to-speech (TTS) models.
- Working with a team to build TTS models for the 22 official languages of India.

### ML Research Intern

IIT Madras (Sept 2023-Jan 2024)

- Developing content bridging theory and practice in classical ML.
- Creating high-quality ML notebooks exploring model internals and guiding effective EDA and model building.

### Machine Learning Techniques (CS2007) TA

IIT Madras (May-Aug 2023)

- Prepared comprehensive course notes adopted as standard course material.
- Mentored students, conducted sessions on classical ML topics, and created tutorial content.
- Recorded companion sessions, enhancing course accessibility.

## Projects

### Cat Image Classifier using Neural Networks from Scratch

Link: <https://github.com/SherryS997/Cat-Image-Classification-Neural-Network-using-NumPy-and-SciPy>

- Implemented custom neural network in Python with NumPy.
- Independently coded deep learning algorithms, starting from scratch.
- Tech Stack: Python, NumPy, SciPy, PIL

### Blog-Lite v2 (Vue3, Python-Flask)

Link: <https://github.com/SherryS997/Blog-Lite-v2>

- Awarded 'Best Project'. Part of course-work.
- Platform for text/image blogs, enabling user interactions, utilizing Flask Rest API with Celery & Redis.
- Tech Stack: Vue3, Python, Flask, SQLite, Redis, Celery

### Machine Learning Algorithms (Quarto, Markdown, Latex)

Link: <https://github.com/SherryS997/Machine-Learning-Algorithms>

- Developed a comprehensive resource documenting the mathematical foundations of classical ML algorithms.
- Adopted by IIT Madras as course notes for the Machine Learning Techniques Course.
- Tech Stack: Python, Scikit-Learn, Pandas, Tensorflow

## Certifications

### Deep Learning Specialization

Completed 5 courses mastering CNNs, RNNs, LSTMs, Transformers, and optimization techniques in neural networks for applications like speech recognition and NLP.

Link: <https://coursera.org/share/6dfedef63e8fbd72284092cb7c5b0da9>

### Generative AI with Large Language Models

Acquired learning in LLM lifecycle, training, fine-tuning, and deployment strategies for real-world applications, understanding business implications from industry insights.

Link: <https://coursera.org/share/abd98e71f4b17faa5c391bc2aceff982>

## Positions of Responsibility

### Senate Member

Christ University, 2019-20

- Actively contributed as a member of the Senate, the highest governing body.
- Participated in crucial meetings addressing important matters.

### Student Council Member

Philosophy Department, Christ University, 2020-21

- Helped out in Department Budgeting.
- Organized the Annual Fest and various departmental events throughout the academic year.

## Skills

**Languages:** Python, JavaScript, Java, SQL

**Web:** HTML, CSS, Vue.js

**Tools:** TensorFlow, PyTorch, OpenCV, Scikit-learn, Pandas, Flask, Django, SQLAlchemy

**NLP:** llama.cpp, Langchain, Llama LLM, Mistral LLM, Bloom LLM, Flan-t5