

# KEERTHANA SENTHILNATHAN

[ksenthilnathan@ucsd.edu](mailto:ksenthilnathan@ucsd.edu) [linkedin.com/in/skeerthana02/](https://www.linkedin.com/in/skeerthana02/) +1(858)-518-3984

## EDUCATION

### University of California San Diego

*M.S in Data Science - 4.0/4.0*

Sep. 2024 – June 2026

*California, USA*

### National Institute of Technology Tiruchirappalli

*B.Tech in Electronics and Communication Engineering-3.9/4.0*

Nov. 2020 – May 2024

*TamilNadu, India*

## EXPERIENCE

### Alignerr LabelBox

*AI Tutor Generalist*

June 2024 - August 2024

*SanFrancisco , USA*

- Restructured AI model performance through advanced coding techniques and comprehensive data analysis; accomplished a **40%** increase in reliability and a **30%** improvement in overall system accuracy.
- Evaluated model performance, addressing diverse challenges to assist cutting-edge **generative AI models**. **Skills:** Prompt Engineering, Inferencing, Training Data Generation, Model Optimization

### Carnegie Mellon University

*Research Intern, Xu Lab*

Dec. 2023 - June 2024

*Pittsburg, USA*

- Conceived advanced classification models using LLMs and **Visual Linguistic Multimodal models** for Subtomogram classification of Cryo-ET images, resulting in a **15%** improvement in classification accuracy.
- Attained a ROUGE score of **0.782**, demonstrating significant accuracy and reliability for betterment of model.

### University of Hagen, Germany

*NLP Research Intern*

June 2023 - Dec. 2023

*Hagen, Germany*

- Re-estimated **visual-linguistic embedding space** by capturing implicit linkages in memes through innovative re-estimation techniques, resulting in a **20%** increase in embedding accuracy.
- Detected hate meme detection performance and contributed to a significant increase in accuracy by **18%** by validating new embeddings properties and performed **Meme Text Prediction (MTP)** task for bridging semantic gap
- Secured an accuracy of **84.57%**, after post-training the **Detectron** Visual Linguistic model.

### Samsung R&D Institute, India

*Generative AI Summer Intern 2023*

June 2023 - Aug. 2023

*Bangalore, India*

- Developed an LLM for **conversational dialogue summarization**. Worked with the Voice Intelligence team on **Flan T5** and **GPT** models, deploying zero and one-shot learning techniques.
- Succeeded a ROUGE score of **0.745**. **Skills:** NLP, Generative AI, Text Generation.

### National Institute of Technology, Tiruchirappalli

*Deep Learning Research Intern*

April 2023 - June 2023

*Tiruchirappalli, India*

- Refined hyperparameters of **RNN** and **LSTM** models for cloud resource usage prediction, achieving **30%** reduction in forecasting error and **25%** increase in computational efficiency.
- Conducted deep analysis to identify most affected instances, concluding with an RMSE score of **0.095**. **Skills:** Time series forecasting, hyperparameter tuning, Optimization Algorithms

### Samsung R&D Institute, India

*PRISM Intern*

June 2022 - Dec. 2022

*Bangalore, India*

- Materialized an open-domain **question-answering system on resource-constrained devices** employing an efficient information Retrieval method for keyword extraction, improving efficiency by **27%** for keyword extraction.
- Bagged an accuracy of **81.25%** with encoder-decoder transformer architecture (**BART**) for contextual retrieval and text generation. **Skills:** NLP, Generative AI, Text Generation.

## CAPESTONE PROJECT

### Novel Question-Answering System Using Hierarchical NER System - AGS Healthcare, USA Jan. 2024

- \* Spearheaded an innovative **entity-specific extractive QA** model for second-level entity recognition, focusing on granularity in information retrieval of symptom and disease.
- \* Got a Precision of **0.89** and a Mean Reciprocal Rank (MRR) of **0.767**. **Skills:** NER, Extractive QA, Healthcare Data Analysis

## PROJECTS

---

### LangChain Chat with my Data

June 2024

- Consolidated a **chatbot** responding to user queries based on content of specific documents, engaging **Retrieval-Augmented Generation (RAG)** for precise response accuracy.
- Implemented **LangChain** to integrate and process data, ensuring it leverages document-specific information rather than general training data, revamping information retrieval efficiency by **35%**.

### Genetic LSTM

March 2024

- Applied **Genetic Algorithms (GA)** to optimize LSTM model hyperparameters by encoding as genes, using **crossover** and **mutation** to generate new sets, and selecting best configurations based on performance metrics.
- Leveraged evolutionary principles, such as selection, crossover, and mutation, to iteratively refine **forecasting model**. Attained an RMSE Score of **0.087**.

### Sentiment Analysis on Customer Reviews

Jan. 2024

- **Introduced** a **sentiment analysis** system for classifying customer reviews, combining **SQL** for **data extraction** and preprocessing with Natural Language Processing (NLP) techniques for sentiment classification.
- Secured an accuracy of **0.835**, demonstrating effective sentiment classification and insight into customer opinions.

### Speaker Recognition System

June 2023

- **Initiated** a speaker recognition system with **Convolutional Neural Networks (CNNs)** to identify and verify speakers based on voice characteristics accurately,
- Received an F1 score of **0.89**, landing a high accuracy in distinguishing between different speakers.

## PUBLICATIONS

---

### A Novel Question Answering System using Hierarchical NER System for healthcare

*In review*

Keerthana Senthilnathan, Manasvin Surya B J, Rajeshwari Sridhar

Submitted to *Sadhana- Academy Proceedings in Engineering Sciences*, June 2024.

### Symptoms Annotated dataset for token classification in Medical Extracts

*In review*

Manasvin Surya B J, Keerthana Senthilnathan, Madhumitha Senthilnathan

Submitted to *International Journal of Data Mining and Bioinformatics*, May 2024.

### Cloud Resource usage forecasting using deep-learning based predictive framework

*Accepted*

Keerthana Senthilnathan, Jitendra Kumar

Submitted to *Springer Nature Singapore (Scopus indexed CCIS Series)*, April. 2024.

## TECHNICAL SKILLS

---

**Languages:** Python, SQL, Matlab, C, C++, Java

**Data Analysis:** Pandas, Numpy, Matplotlib, Seaborn

**Machine Learning:** PyTorch, TensorFlow, Keras, IoT, Computer Vision

**Developer Tools:** Jupyter notebook, Colab, Git, VS Code, Google Cloud Platform, Qiskit, Excel, Azure

**Mathematical Foundations:** Statistics, Multivariable Calculus, Linear Algebra, Hypothesis testing, Probability

## CERTIFICATIONS

---

**AWS:** Generative AI with Large Language Models

**Udemy:** Python for Data Science and Machine Learning

**Samsung R&D Institute:** PRISM Internship Certification

**IBM:** Data Analysis with Python | Introduction to Data Science | Introduction to AI Ethics

**Deeplearning.ai:** Natural Language Processing with Classification and Vector Spaces | Sequence Models | Neural Networks and Deep Learning | Supervised Machine Learning: Regression and Classification | Natural Language Processing in Tensorflow

## LEADERSHIP / EXTRACURRICULAR

---

### Pragyan Crossfire and Guest Lectures Team - Manager, NIT Trichy

May 2024 - Jan. 2021

**Ideated and implemented** Crossfire, **Contacted and coordinated** with panelists for topics of social importance. **Planned and managed** interactions between students and almost **20+** notable luminaries.

### Festember Workshops and Informals- Manager - NIT Trichy

May 2024 - Dec. 2020

**Coordinated** Festember, South India's largest cultural student-run non-profit event with **17,000+** attendees. **Collaborated with teammates** to contact renowned companies for professional shows.