

# APALA THAKUR

Cupertino, CA 95014 | (408) 674 5893 | [apala.thakur@gmail.com](mailto:apala.thakur@gmail.com)  
<https://www.linkedin.com/in/apala-thakur-b6ab6428b> | <https://github.com/apalathakur>

*I am motivated to explore and apply Machine Learning techniques to Natural Language Processing. My goal is to augment my existing skills in NLP and ML to novel application domains, handling diverse datasets, creating innovative tasks, implementing efficient models and algorithms for solving complex issues.*

## EDUCATION

### University of California, Santa Cruz, CA

M.S. in Natural Language Processing

Summer 2023

B.S. in Computer Science

Spring 2021

**Relevant Coursework:** Machine Learning (ML), Artificial Intelligence (AI), Probability & Statistics, Natural Language Processing (NLP), Fundamental Linguistics, Advanced Machine Learning Application in NLP, Data Wrangling

## SKILLS

**Languages:** Python, C/C++, Java    **Tools & Library:** PyTorch, Data Wrangler, FAISS, HuggingFace, Haystack

**NLP:** LLM Transformers - BERT, GPT, Multi Modal Transformers - CLIP, BLIP, ViT, OpenFamingo

**ML:** DNN, CNN, RNN, LSTM, GRU

## HIGHLIGHTED PROJECTS & EXPERIENCE

### Improving Answering Accuracy in Multi-Modal QA with Retrieval-Augmented Models

Summer 2023

NLP Capstone, Mentor - Meta Inc.

- Selected WebQA and MultimodalQA datasets, then developed an API for efficient data loading, using CLIP/BLIP to generate embeddings of text and image in unified vector space for dense passage retrieval
- Engineered a document store using FAISS for information retrieval using cosine base similarity
- Evaluated the QA performance using F1 and EM scores on multi-modal OpenFamingo transformer
- Enhanced QA performance on OpenFamingo Multimodal Transformer through fine-tuning

### Child Mental State Detection through EEG Wave Pattern Analysis

Winter 2020

B.S. Capstone Project

- Precisely divided the EEG data for each child into 10-minute intervals to capture mood patterns of focus, defocus, and drowsiness, with each segment containing as many as 16,000 instances of 25 wave patterns
- Achieved 68% accuracy with RNN-based model on mental state detection using wave patterns

## LEADERSHIP & EXTRACURRICULAR ACTIVITIES

### Speech Evaluator, Intonation Specialist, Delivery Style Advisor

Jan 2021 – present

Toastmaster club (Easy-Speak), Cupertino, CA

- Provided constructive support to speakers to deliver impactful speeches with confidence
- Advised speakers for refining vocal dynamics and presentation techniques

### Tutoring Assistant for Machine Learning

Winter, Spring 2021

Department of Computer Science, UCSC, California

- Assisted students with implementing and evaluating models like CNN, RNN, SVM, Naïve Bayes, and K-NN
- Provided guidance on model evaluation utilizing metrics such as F1 score, EM score, and Confusion Matrix

### Tutoring Assistant for Python

Winter 2020

Department of Computer Science, UCSC, California

- Supported students in learning Python programming and guided them to acquire skills in Python data structures for effective problem-solving