Samyak Mehta

samyakmehta28@gmail.com linkedin.com/in/samyakmehta28 github.com/samyakmehta28

Work Experience

•Adobe Inc - Multimodal Content Lab *Research Intern*
Bengaluru, India May 2022 - Jul. 2022
  – A novel end-to-end method to supplement multimodal procedural documents with more multimodal granular details.
  – In a human evaluation on Amazon Mechanical Turk, 57.47% of annotators found the expanded content generated by our pipeline to be clearer, while 58.83% and 52.42% found it more textually and visually informative respectively.

Technical Skills

- **Programming Languages**: C, C++, Python, SQL, MATLAB, Latex, Verilog, Bash

- **Web Technologies**: HTML, CSS, Django, Node.js, Express.js, Nest.js, MongoDB, JavaScript, TypeScript

- **Tools/Framework**: Numpy, Pandas, Scikit-learn, Keras, Tensorflow, GIT

- **Operating System**: Windows, Linux ** Elementary proficiency*

Education

- **University of California San Diego** *Sept. 2023 - June. 2025*
  Master of Science in Computer Science and Engineering CGPA: 4.0/4.0
  – **Key Courses Taken**: Probabilistic Reason and Learning, Recommender System and Web Mining, Statistical Learning, Principles in Computer Architecture, Unsupervised Learning.
- **Indian Institute of Technology (IIT), Guwahati** *Jul. 2019 - May. 2023*
  Bachelor of Technology in Electronics and Communication Engineering CGPA: 9.13/10, Rank: 2/103
  – Minor in Computer Science and Engineering.
  – **Key Courses**: Data Structures and Algorithms, Software Engineering, Introduction to Computing, Digital Logic and Computer Architecture, Computer Systems, Computer networks, Deep Learning, Probability and Random Processes.

Thesis and Projects
* In progress

- **Prompt Response Logging System with Detailed Metrics** *March. 2024 - April. 2024*
  Personal t.ly/fJqv4
  – Implemented a **NestJS** proxy layer to forward requests to OpenAI, logging responses to ClickHouse database with detailed metrics. Included APIs for querying, filtering requests, and fetching aggregate metrics.
  – Developed a dashboard using **Streamlit** to showcase sample requests, token counts, and request graphs over time.

- **Node.js and Express.js Projects** *January. 2024 - March. 2024*
  Personal t.ly/QHkL0
  – Implemented a Task Manager application using Node.js and Express.js, allowing users to organize tasks efficiently.
  – Developed a RESTful Store API utilizing Node.js and Express.js, enabling seamless interaction with store inventory data.
  – Created a Job API to facilitate job listings and applications, enhancing job search functionalities.

- **WebBlog** *July. 2021 - September. 2021*
  Personal git.io/JCvUN
  – A web development project, developed using **Django framework** for blogging purpose.
  – The web-app has complete **user authentication** and it has functionalities like creating, reading, editing, deleting, liking or unliking blogs. The blogs are also categorized in case the user wants to read blogs of a specific category.

- **Optical Character Recognition of Textual Documents** *August. 2022 - May. 2023*
  B.Tech Project under Dr. Anirban Dasgupta Thesis
  – This project involves developing a pipeline for input document line segmentation, handwritten/print line classification, OCR processing for each line, and post-processing the OCR output using a novel NLP technique.
  – Our model produced a **WER of just 0.045 and CER of just 0.005** on OSCAR dataset, providing better accuracy than models like Textblob, SymSpell, Seq2SeqLM BART etc.

Achievement

- **Patent**: System and Methods for expanding sparse multimodal documents, Patent filed June 2023, Patent pending.
- **Publication**: A Novel Pipeline for Improving Optical Character Recognition through Post-processing Using Natural Language Processing. Aishik Rakshit, Samyak Mehta, Anirban Dasgupta. *Accepted to IEEE GCON 2023.*
- **IMS 2021**: Received the prestigious Institute Merit Scholarship for the year 2021.
- **Department Rank**: Department **Rank 2** in Electronic and Communication Engineering department.
- **Joint Entrance Examination Main**: Secured **99.79 percentile** among 1 million candidates, in 2019.
- **Joint Entrance Examination Advanced**: Secured **AIR 1808** among 0.17 million candidates, in 2019.

Extracurricular

- **Teaching Assistant**, for the course Introduction to Python during SP24.
- **Teaching Assistant**, for the course Data Science in Practice during WI24.
- **Organizer**, PR and Branding, Techniche, IIT Guwahati, 2019-2020.
- **Team Member**, Mess Management Committee, Brahmaputra Hostel, IIT Guwahati, 2019 - 2020.
- Part of **Electronics Club**, IIT Guwahati, 2020 while working on the "Bunny Hammer Game" project.