Chaitya Shah

 Chaitya0623 | Chaitya Shah | chaityashah.netlify.app | chaitya0623@gmail.com | +91 70456 31629

EDUCATION

Dwarkadas J. Sanghvi College of Engineering: B.Tech in Computer Engineering (CGPA: 8.89/10)	*Dec 2021 - Present*
Kishinchand Chellaram College: 12th HSC (95.5%)	*Jun 2019 - Aug 2021*
Gopi Birla Memorial School: 10th CBSE (93.8%)	*Apr 2014 - Mar 2019*

SKILLS

- **Programming**: C, C++, Python, HTML, CSS, TailwindCSS, JavaScript, Three.js, React, Django-Rest, MySQL, R, LaTex.
- **Technologies**: Git, Github, Robot Operating Systems, AWS, FastAPI, Docker, Terraforms, Google Cloud, Looker Studio, Arduino.
- **ML Tools**: Numpy, Pandas, Seaborn, Matplotlib, Scikit-Learn, Keras, Tensorflow, OpenCV, NLP, LLM's, Langchains, Haystacks.

EXPERIENCE

Research Intern | *Indian Institute of Technology, Patna*   *May 2023 - Present*
- **Research Paper**: Writing a Research Paper on Fraud Detection in Ethereum transactions using **Graph Neural Networks**.
- **Random Walk**: Implemented **Directed Random Walk** algorithm to traverse through nodes using the networkx library.
- **Feature Engineering**: Performed Data Analysis and Feature Engineering for every traversed node in the graph.
- **Temporal Graphs**: Created a **Temporal Graph Attention Layer** giving us relevant performance metrics based on the timestamps.
- **Model Performance**: Training the model on a larger recent dataset resulting in substantial reduction in false positives.

Research Engineer | *SimPPL, New York*   *Sep 2023 - Present*
- **Workshop Paper**: Working on Machine Learning models to reduce fake information and **promote trust on social internet**.
- **Data Extraction**: Wrote the code to extract posts from the Truth Social App that redirected to Wikepedia, Bluesky and Youtube URL's.
- **Content Sharing**: Working on NLP models to highlight how the sharing of content on the App has progressed over time.
- **Technologies**: Storing the relevant data on **Google Cloud** and Data visualizations were made using **Google Looker Studio**.

Artificial Intelligence Intern ↗ | *Infiheal*   *Jun 2023 - Sep 2023*
- **Technologies**: Worked on a Mental Health Chatbot, trained on Web Scraped Data & **Prompt Engineering**, tested using **FastAPI**.
- **Models**: Models were made using **Large Language Models (LLM's), LangChains, Haystacks** & **HuggingFace** Transformer models.
- **Deployment**: Deployed models on **Amazon Web Services (AWS)** via Linux using Terraforms, Docker, CloudWatch, ECR and ECS.
- **Mentorship**: Led a team of 8 newer interns, mentored and provided guidance to them about the used Chatbot Technologies.
- **Performance Enhancement**: Attained a 4% accuracy boost by fine-tuning the algorithms, also cutting response time by 1-2 seconds.

PUBLICATIONS

- **Sentence Restructuring with User-Controlled Difficulty using NLP** ↗: A custom synonyms dataset was curated using **Google Ngrams Viewer** word counts, replacing words in the given sentence while preserving context through the **Lesk Algorithm**. The sentence complexity was then adjusted by adding or reducing words, and finally, **Sentence Transformers** were used to reframe the sentence while maintaining contextual similarity. This paper was presented at the 14th International Conference on Computing, Communication and Networking Technologies (ICCCNT) 2023, IIT Delhi, **to be published** in **IEEE Digital Library Xplore**.

- **MedLedger - An ML Model for Disease Outbreak Prediction**: Currently working on a Research Paper that employs classification algorithms to predict localized disease outbreaks at the clinic level, considering factors such as locality, seasons, and age.

PROJECTS

- **StockSage** ↗: Developed an interactive Full-Stack Web application using **React and Django**, enabling users to manage their transactions for CAC40 stocks. The application employs **LSTM** for stock price predictions and recommends new stocks based on user holdings through **Association Rules and the Apriori algorithm**.
  Additionally, there's a feature for assessing the **Credibility of Financial Tips**. Data is initially scraped from Reddit, subjected to **Topic Modeling**, and if it significantly leans towards either a stock going up or down, the tip is deemed reliable. However, if it lacks a clear bias, predictions cannot be made, and the tip's validity is uncertain.

- **TubeBoost** ↗: Engineered an ML framework aimed at **YouTube optimization**, focusing on identifying the optimal upload schedule for channels and forecasting the top 10 emerging content creators. Additionally, we undertook the classification of video quality into distinct categories and determined the optimal video duration, and deployed the application on **Streamlit**.

- **AVATAR - *(Aerial Vehicle for Advance Terrain Analysis and Reconnaissance)***: Engaged in developing a multi-functional drone capable of Alarm Generation and Payload Delivery. It employs **YOLO Models** and **Thermal Imaging** for distant human detection, while autonomous navigation is facilitated by **GPS Tracking**. In cases where humans cannot be detected visually, **Acoustic ML** will be used. It also supports voice commands implemented by topic modelling of speeach to text NLP Models.Integration is achieved through ROS, complemented by GUI visualizations. This versatile drone serves purposes in both military and civilian operations and has helped us earn a place as a finalist in the Smart India Hackathon 2023.

- **Short-a-Thon** ↗: An interactive platform meticulously tailored for hosting and participating in hackathons, providing users with the flexibility to either initiate or join these dynamic events. The platform harnesses a sophisticated machine learning model, specifically in the field of **NLP**, coupled with host-specified parameters to seamlessly streamline the process of team shortlisting for hackathons, enhancing the overall experience for participants and organizers alike.

- **VerifyU** ↗: Fulfilled the role of a Full Stack Developer (React, Django) for a Document Verification website, while also taking a proactive role in implementing a robust ML model combining **CNN, OCR, and Encryption through QR codes** to improve document authenticity validation. Effectively contributed to the end-to-end deployment process, ensuring a secure user experience.

- **Internship Portal** ↗: Developed a comprehensive Internship Portal as a Full Stack Django Developer at **DJ Unicode**, centering on facilitating seamless interaction between students and professors for internships and research opportunities. The project encompassed Figma for design, React for frontend, and Django for backend, with deployment on **Netlify** and **PythonAnywhere**.

EXTRACURRICULAR ACTIVITIES

Coding Team Member | *DJS Antariksh (Official Martian Rover Team)*   *May 2022 - Present*

- **Robot Operating System**: Worked on Rover Integration using **Arduino** & ROS on **Ubuntu (Linux)**, worked with **Gazebo, Rviz**.
- **Graphical User Interface**: Integrated a React website with ROS, aiding in camera visualizations, displaying data and Rover Control.
- **Autonomous Movement**: Worked on Autonomous functioning of the Rover & the Robotic Arm with **AI, ROS and OpenCV**.
- **International Competitions**: 3rd in International Rover Design Challenge 2022, 8th in International Rover Challenge 2023 (Onsite).

Machine Learning Head | *Synapse (AI/ML Committee)*   *Sept 2022 - Present*

- **Mentorship**: Designed and conducting training sessions on ML concepts, CV techniques, and NLP algorithms for mentees.
- **Guidance**: Assisting them with personalized guidance and mentoring for their Research Papers spanning various ML technologies.
- **Contribution**: Helped organizing events, **had speaker sessions**, delivered seminars, contacted sponsors and wrote Research Papers.
- **ML Models**: During my time as a mentee, I acquired proficiency in a wide range of machine learning techniques, including classification, regression, clustering, natural language processing (NLP), and deployment models.

ACHIEVEMENTS

- **2nd in European Rover Challenge 2023 (Remote)**: Was one of the operators for the Navigation Task (scored full), experimented with Autonomous movements of the Robotic Arm for the Maintenance Task using ROS, Arduino, MoveIt, OMPL, Docker and OpenCV, (practiced on simulation using RViz and Gazebo) and was a part of the Presentation Task and Final QnA.

- **Winner of X-tract 2022** ↗: An Inter-College Datathon where we comprehended and applied significant Machine Learning Algorithms to 7 distinct Data Science datasets in a limited Timeframe of 4 hours while competing with 200 shortlisted students to win the Event.

CERTIFICATIONS

- **Machine Learning with Python (with Honors)** ↗: Coursera   *Issued Apr 2023*
- **Introduction to Cloud Computing** ↗: Coursera   *Issued Oct 2022*
- **2022 Complete Bootcamp from Zero to Hero in Python** ↗: Udemy   *Issued Aug 2022*
- **Google Analytics for Beginners** ↗: Google Analytics   *Issued Jul 2022*
- **The Fundamentals of Digital Marketing** ↗: Google Digital Garage   *Issued Dec 2021*