Isita Polamarasetti

ip22@rice.edu | (832) 806-9945| Github | Linkedin | PortFolio

Education

Rice University- GPA: 3.8 / 4.0 **08/2023 - 05/2025**
Master's in Computer Science *Houston , TX*

Gitam University - GPA: 9.43 / 10.0 **06/2019 - 04/2023**
Bachelor of Technology in Computer Science and Engineering *Visakhapatnam, India*
- Consistently awarded a **25%** scholarship each year for academic excellence at Gitam University.
- Top **10%** in leetcode.

Experience

Intelligent Medical Objects - IMO Health **05/2024 - Present**
NLP Engineering Intern *Houston , TX*

- **Rare Disease Prediction Project:** Conducted statistical analysis on individuals diagnosed with PAH (Pulmonary Arterial Hypertension) and those without. Created charts to understand the stages of PAH, compared groups to identify mutual information, assessed co-occurring conditions, and performed statistical tests to draw meaningful comparisons. The analysis is featured in the abstract of the article which is being presented in the **ISPOR** Europe 2024 Conference.
- **AILA Project:** Enhanced the efficiency of systematic literature review processes, reducing the time by **35%**. Contributed to making extraction and prompt engineering processes more efficient, corrected bugs, and developed a model to retrieve information from pdfs, images and plaintext using **Pydantic** and **LLM chain**.

Rice University **12/2023 - Present**
Teaching Assistant (COMP 543) | Grader (Comp 502) *Houston , TX*

- Collaborated with **Professor Christopher Jermaine** to provide personalized support during office hours for the course **'Grad Tools and Models for Data Science,'** addressing inquiries, guiding on assignments, and managing grading responsibilities with constructive feedback. Additionally, served as a Grader for **'Neural Machine Learning I (Comp 502)'** under **Erzsébet Merényi.**
- Assisted students in mastering data science tools, including relational database systems, Hadoop, Apache Spark, TensorFlow, fostering a comprehensive understanding of data manipulation techniques. It's noteworthy that I also assumed a significant role, handling **15%** of the class responsibilities, for a cohort of **300** students.

Footsteps LLC (Amazon DSP) **02/2023 - 07/2023**
Software Engineering Intern (Remote) *California, USA*

- Used statistical modeling, machine learning, including LOF and gradient boosting, to derive actionable insights from extensive Amazon driver data, leading to an **11%** cut in delivery times and a **13%** boost in fuel efficiency.
- Developed advanced automation solutions to optimize operations on the Amazon logistics website. Integrated APIs to automate operations, and designed custom scripts, saving **21** hours per week in manual data entry and reduction in errors considerably.
- Leveraged OpenStreetMap API to develop a logistics website with React, JavaScript, SCSS, Node.js, embedding dashboards that boosted efficiency by **33%** and increased weekly bonuses through improved delivery tracking.

Pharmapro Tivra Health LLP **10/2022 - 07/2023**
Data Science Intern (Remote) *Mumbai, India*

- Implemented NLP classifiers and matchers for complex biomedical clinical trial data, achieving a **35**% notable increase in accuracy through the application of machine and deep learning algorithms like Token-based Matching and Bert.
- Created interactive dashboards and data-driven applications with database connectivity using tools such as Tableau and Python libraries to visualize insights related to medical specialties, detailing the distribution of specialized doctors across states and cities in India, their experience, demographics and other factors.
- Orchestrated integration of healthcare APIs from industry-leading sources like **nih.gov.in** and **credihealth.com** to automate extraction of structured medical data, resulting in a **85%** improvement in data acquisition speed and accuracy.

Projects

- **Hotel Search Engine:** Led the development of a Unified Hotel Search Engine website with Apache Solr, implementing NLP techniques (TF-IDF, Word2Vec, BERT, LSA, NER, Sentiment Analysis), dimensionality reduction(PCA) and inverse indexing, achieving a **35%** increase in search speed and **27%** higher recommendation accuracy.
- **Efficient Hashing Using Huffman Coding (COMP 582 Project):** Introduced an innovative integration of hashing and Huffman coding to enhance data manipulation efficiency and optimize space utilization by **47%**. Demonstrated proficiency in advancing hashing techniques, resulting in a substantial reduction in memory usage for HashMaps through extensive research and algorithm implementation.
- **PixPlate:** An innovative app offering personalized recipes based on users dietary needs and health conditions, allowing them to select specific products or take a picture of their fridge. Utilized TensorFlow for object detection to identify items in fridge images and implemented large language models (LLM) for generating customized and context-aware recipe recommendations. The website is built using React Native and deployed on AWS.

Skills

Languages: Python, Java, JavaScript, Typescript, Matlab
Web Development: React.js, HTML, CSS, Node.js, jQuery, Bootstrap, REST, MVC
Data Science: NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch, Keras, Plotly, Dash, NLTK, Spacy, Pydantic, Matplotlib
Databases: MySQL, PostgreSQL, NoSQL
Tools/Platforms: GIT, Linux, Pytest, BeautifulSoup, Selenium, Hadoop, Docker, Tableau, CI/CD, AWS, Microsoft Tools