Vedant Saraswat

vedantiitkgp | pitess100 | portfolio | Vedant | vsaraswa@uci.edu | +1.9495618066

EDUCATION

University of California Irvine - Master of Science in Computer Science   Sept 2023 - Apr 2025
– Fully funded masters with **Graduate Teaching/Research Assistant** scholarships, focusing on Big Data and AI Applications.
– **Courses** - Advanced ML, Reinforcement Learning, Efficient ML, Information Retrieval, Computational Geometry, Networks

Indian Institute of Technology Kharagpur - Bachelor of Technology (Honors)   July 2017 - Apr 2021
– Minor Curriculum in **CS** - Programming & Data Structures, Algorithms, Social Computing, Image Processing, Probability, ML
– Department Rank 4 and awarded **RK Bramha Cup** for the Best final year (Technology) and recognition letter from Director

EXPERIENCE

Machine Learning Intern - (Quality of Engineering Team) - Viasat   June 2024 - Sep 2024
– Developed ML **forecasting models** on AWS to predict active network speed traffic patterns, optimizing bandwidth allocation.
– Designed ETL pipelines with **SQL** in Snowflake for models ranging from regressor to **t-Transformers** getting under 6 RMSE.

Data & Tech Fellow - (Professor Tom Andriola) - University of California Irvine   Sep 2023 - Mar 2024
– Developed a **Mixture of Experts LLM** chatbot from scratch to deliver student-oriented responses for diverse academic queries.
– Built GPU Optimized **RAG** platform to automate the creation of sustainability dashboards based on AASHE data, with web integration via Streamlit to facilitate efficient querying and drive informed decision-making to enhance sustainability initiatives

Computer Vision Research Engineer - (ZotBins) - University of California Irvine   July 2023 - Present
– Developed a waste classification pipeline, employing a modified EfficientNet architecture, achieving 92.7% multi-label accuracy.
– Deployed real-time waste classification using CLAHE, perception correction, and optimized YOLOv5 model with 87.2% mAP.

Software Development Engineer - (Web & Games Team) - Amazon MX Player   Jun 2021 - Aug 2023
– Orchestrated end-to-end creation of a **HLS Streaming Player** for various web OS with inbuilt DRM protection and **dynamic ad insertion**, utilizing VideoJs, Widevine, and delivering seamless experiences for 175 million people.
– Engineered full-stack user flows, covering **Authentication**, Subscriptions, **Push Notifications** and **Content Recommendation** from scratch utilizing TypeScript, React, Next.js, Spring, and an adaptive SageMaker-based recommender engine.

Undergraduate Researcher - (Professor Pawan Goyal - Thesis Project) - CNERG Labs   Jun 2020 - Aug 2021
– Detected protein interaction in Medical texts based on shortest **dependency knowledge graphs** and common sense encoders.
– Researched suicidal personality prediction on extracted multimodal Facebook data using **T-LSTM and Deep BOSE** network

Software Development Intern - (Transportation Tech team) - Amazon   Apr 2020 - Jun 2020
– Enhanced client onboarding in fleets by creating a **Spring** backend module to rectify errors, utilizing Postgres Database to manage **data disparities** between container and their Sable Database API Calls on the frontend developed with JSP.
– Formulated an algorithm to optimize **Transactions Per Second (TPS)** effectively reducing latency for the TCDA portal.

Data Science Intern - (IBUTUS team) - Tech Mahindra   May 2019 - Jun 2019
– Optimized Model Performance on 30 million VOIP instances solving label imbalance and contributed to the sk-multilearn library.
– Applied **cost-sensitive RL algorithms**, reducing SCUMBLE by 15% and enhancing macro-F1 score by 10.5% via ensemble.

Deep Learning Intern - (Professor Mrigank Sharad) - AgNext Research Labs   Dec 2018 - Jan 2019
– Utilized OpenCV for noise reduction and enhancing contours in 4 categories of maize data via **Otsu thresholding** and **RCNN**.
– Optimized **Mobilenet CNN** and **SSD** models using Transfer Learning, elevating recall and F-score from 60% to 86% with data augmentation, Blob detection etc, and deployed them in flagship product of **Spice Board of India** - Specx Vision Basic.

PROJECTS

Speculative Decoding - Experimented optimization pipeline to accelerate language model inference, achieving 3x speedups.

UCI Search - Engineered IR Search Engine on CS Database with PageRank, n-gram indexing, RAG, and RLHF/PPO techniques.

Smart Chat Assistant - Designed an AB Inbev Business and Customer-centric proactive Chatbot with interactive AI webhooks

Doctor's Prescription Scanner - Developed a Tesseract OCR based interface for multilayer handwriting correction for patients

SKILLS

Languages/Scripts	C, C++, Python, Go, Java, Javascript, Typescript, Rust, Octave, Matlab, HTML, Bash
Libraries	OpenCv, Numpy, Pandas, React, Next.js, Tkinter, Spacy, NLTK, Flask, Django, Beautiful soup, Spark
Frameworks/Databases	Tensorflow, PyTorch, ROS, Selenium, Spring, SQL, MongoDB, CUDA, AWS, Git, Azure, Docker