MAHAVIR CHANDALIYA

San Jose, CA, USA | +1(669)-213-7346 | mahavir.chandaliya@sjsu.edu | https://www.linkedin.com/in/drmahavir/ | https://docmhvr.github.io/

EDUCATION

Master of Science in Computer Engineering, Machine Learning Engineering **(Jan 2022 – Dec 2023)**
San Jose State University, San Jose, CA

Bachelor of Engineering in Computer Science and Engineering **(Aug 2016 – Nov 2020)**
D.J. Sanghvi College of Engineering, University of Mumbai, Mumbai, India

EXPERIENCE

Research Assistant at SJSU (**Feb 2024 – Current)**
- Contributed to AI cloud research for a surveillance dashboard, leveraging AWS for deployment and integration data from IoT stations, drones, CCTV's.
- Trained and fine-tuned a CNN-based model for crime detection and classification from CCTV videos of SJSU campus, improving accuracy by 15%.
- Authored and published a research paper on the "Powerline Inspection System," presented at the Big Data Conference in 2024.

Software Developer at Remajin **(July 2021 – Dec 2021)**
- Worked on a mobile retail app, a sales representative application, and an admin panel, used by over 1000 retailers and 13 sales representatives.
- Enhanced the sales app and admin panel by collaborating with software and business teams, adhering to quality standards and client needs.
- Resolved over 10 software bugs and added features to the mobile application, leading to the onboarding of 32 manufacturers and over 100 retailers.

Software Engineer Intern at Reckon Energy **(July 2019 – Sept 2020)**
- Transitioned the company to an online presence, by enhancing the company's website and optimizing data management with SQL, resulting in a 35% increase in online traffic. Utilized AWS services and third-party providers for deployment, significantly reducing downtime.
- Spearheaded database development, implementing efficient data management and retrieval with SQL, improving operational efficiency by 40%.

RELEVANT PROJECTS

Ayurbot (Graph RAG, Prompt Engineering, Generative AI**)** **(Aug 2024 – Sept 2024)**
- An Ayurveda chatbot using Langchain, Neo4j, and OpenAI to generate a knowledge graph from an Ayurveda book PDF for precise GPT-driven responses.
- Optimized the Neo4j graph vector store retrieval with Cypher queries, custom entities, relations, advanced prompts, hybrid search for enhanced RAG.

YouTube Data Analytics (AWS S3, Glue, Lambda, Tableau) **(July 2024 – Aug 2024)**
- Architected and deployed an end-to-end data pipeline utilizing AWS S3, AWS Glue, AWS Lambda, and AWS Athena to process, transform, and analyze YouTube trending video statistics for real-time analytics, facilitating data-driven decision-making for content optimization.
- Designed and developed a dynamic Tableau dashboard for interactive data visualizations from the processed YouTube data.

AI-Powered Research Expert (Text Generation, Text Summarization, Open-source LLM's, OpenAI**)** **(June 2024 – July 2024)**
- Utilized Llama3 and Hugging Face, to find, summarize, and recommend research papers based on user input, generating research approaches based on relevant papers. Implemented with the Langchain framework, Arxiv searches, and LLM chains, running LLMs locally using Ollama.
- Developed an OpenAI version using GPT models via the OpenAI API for summarization and generation tasks, hosted on Streamlit Community Cloud.

Book Genre Prediction (Natural Language Processing, Data Mining, Data Visualization, Machine Learning) **(Oct 2022 – Dec 2022)**
- Developed a web application using deep neural networks to predict literature genres from book summaries, achieving over 90% accuracy. Conducted data cleaning, EDA, feature extraction, data augmentation and created a Flask web app to display results.
- Applied and compared decision trees, Naive Bayes, and various word embeddings, leading to a 10% improvement in classification accuracy.

PUBLICATIONS

UAV-based Powerline Fault Detection and Classification (Link) **(Dec 2023 – July 2024)**
- Devised a Powerline Inspection System using pre-trained YOLOv8, trained on drone captured images of powerline components for fault detection, achieving 87% accuracy. Applied data augmentation and transfer learning techniques, utilized Roboflow for data annotations and pre-processing.
- Presented the research paper at the 10th IEEE International Conference on Big Data Computing and Machine Learning Applications, 2024.

Social Distancing Detection using Computer Vision (Link) **(Dec 2020 – May 2021)**
- Engineered a Computer Vision application using TensorFlow Graphs, Object Detection API, OpenCV, and Python to enforce social distancing by analyzing live CCTV video. Utilized OpenCV and evaluated CNN, RCNN, Faster-RCNN and YOLO models for object detection.
- Presented the paper at the 5th International Conference on Computing Methodologies and Communication 2021, published in IEEE Xplore.

TECHNICAL SKILLS

- Programming Languages: Python, C++, SQL, C, R, HTML, CSS, JavaScript, TypeScript
- Data Engineering: ETL tools, Hadoop, Spark, Airflow, Kafka, AWS CDK, AWS (Redshift, S3, Glue), GCP, Google AI Studio
- AI and Machine Learning: Pandas, NumPy, Scikit-learn, PyTorch 2.0, OpenCV, TensorFlow, nltk, SpaCy, AWS Bedrock, LangChain, HuggingFace
- Data Visualization: Excel, Tableau, Power BI, Matplotlib, Seaborn, Looker, Quarto, Plotly, Jupyter Notebooks, Colab, Streamlit, Gradio
- Software Development: React, Node.js, Flask, Django, REST APIs, Jira, Pytest, Fast API, Docker, Kubernetes, Git, GitHub, CI/CD
- Databases: MySQL, PostgreSQL, SQLite, MongoDB, Cassandra, DynamoDB, Neo4j, Pinecone, Chroma DB, FAISS

OTHER ACHIEVEMENTS

- Secured first place in Machine Learning Track at Intel's AI for Social Good Hackathon. Building an entire pipeline for ML problem in the 9-hour event.
- Completed OpenCV University's Computer Vision/Deep Learning Professional Certification for Deep Learning with Pytorch 2.0.
- Earned the OCI 2024 Generative AI Professional Certificate by Oracle.