Rohit Chandra

rohitv.chandra@gmail.com | +1 (669) 302 - 0588 | https://www.linkedin.com/in/rohitv-chandra/ | https://github.com/rohit-chandra

EDUCATION

San Jose State University (SJSU), San Jose, USA   **Aug 2021 - May 2023**
 Master's in Artificial Intelligence
Coursework: ML, DL, NLP, Reinforcement Learning, AI, Statistics, Data Mining

Jawaharlal Nehru Technological University, Hyderabad, India   **Aug 2012 - May 2016**
Bachelor of Technology, Information Technology
Coursework: C, C++, Java, Algorithm and Data Structure, Operating System, Computer Architecture, DBMS, Multi-Threading

TECHNICAL SKILLS:

- **Languages:** Python, SQL, R, Java, Javascript
- **Tools/Framework:** Tensorflow, Pytorch, OpenCV, Spacy, YOLO, Huggingface, Transformer, BERT, Sci-kit Learn, Keras, Pandas, Numpy, Scipy, Matplotlib, Seaborn, Flask, Streamlit, FastAI, Word2vec, Boto3, Docker, AWS - Redshift, S3, Angular, Pyspark, Airflow
- **Data Science & Miscellaneous Technologies:** Data Science Pipeline - Data cleaning, Data wrangling, Data visualization, Predictive modeling, Statistical analysis, Interpretation, Model deployment, Model monitoring, Ensembling, Statistics, Time series, Sentiment Analysis, Text classification, Anomaly detection, Hypothesis testing, OOP, Git, GitCLI, GitHub actions, MLOps, Web scraping - Beautifulsoup, Selenium, Scrappy, Tableau, PowerBI

ACADEMIC PROJECTS:

Project: Sign Language Detection   **Aug 2022 - Present**
- Leading a team of 3 in developing an end-to-end sign language recognition web application in **Javascript/React** by processing large-scale video datasets and trained deep neural networks - (**CNN, LSTM, Single Shot Detector**) to recognize the American sign language in real-time
- Implemented **NLP text-to-speech API** to covert the words recognized from hand gestures to speech
- Created and annotated custom image dataset using labelImg to train a baseline model

Project: AI Blog Summarizer/Generate Blog Posts   **Aug 2022 - Present**
- Web scrapped AI blog posts using Beautifulsoup and trained **Huggingface Transformer**, **T5**, and **BERT** models to summarize the blogs.
- Developed an end-to-end web application using streamlit to display the summarized output.

Project: FaceID App   **Feb 2022 - May 2022**
- Implemented Siamese Neural Network for One-Shot Image Recognition research paper to develop a face ID app using **Kivy, TensorFlow object detection pipeline, and OpenCV**.
- Created a custom dataset of over 4000 images and applied data augmentation to improve the performance of the face detection model.

Project: Sentiment Analysis   **Feb 2022 - May 2022**
- Crawled yelp reviews for different products, performed data preprocessing, and trained **BERT** model to identify how the sentiment of the products varies over time.

Project: Customer Churn Analysis   **Sep 2021 - Dec 2021**
- Wrangled **125 million** records of customer data collected from a telecommunication company. Handled **imbalanced** data using different sampling techniques, trained multiple classifiers, and compared the performance using ROC/AUC, and F1-score
- The model predicted the churn with **98% precision** and based on the analysis we found that the complaints and customer status(active/inactive) are strongly associated with churn. Also, recommended ways to reduce the attrition rate

Project: Automatic Number Plate Recognition   **Sep 2021 - Dec 2021**
- Trained a **TensorFlow object detection model** with an accuracy of **96%** to detect license plates from images and in real-time from video, apply EasyOCR to extract the plate number, and leveraged a size filtering algorithm to grab the largest detection region

PROFESSIONAL EXPERIENCE:

Appen, Sunnyvale, USA | **Data Engineer Intern**   **Jun 2022 - Aug 2022**
- Developed **automated table design optimization pipelines** using **python,** and **Boto3** to find current distribution, sort keys, and skewed tables of all schema present on **AWS Redshift**. The automation script also generates the query plan and flags the high distribution cost columns.
- Proposed new distribution and sort keys to optimize the execution time of over 220 SQL queries running on the AWS redshift cluster and **improved** the performance by 15%

Infosys, Hyderabad, India | **Senior Data Scientist**   **Jun 2019 - Aug 2021**
- Developed numerous multi-classification models with up to **94% accuracy** to predict customer satisfaction ratings based on operational metrics such as ticket age, service type, customer's geo-location, and service providers for Microsoft
- This led to a **20% increase** in revenue of Microsoft service providers

PurpleTalk Inc., Hyderabad, India | **Software Engineer**   **Jun 2016 - Nov 2018**
- Upshot.ai: Developed **Android Software Development Kit (SDK)** customer engagement platform in **Java/kotlin** to capture customer activities, analyze data, and make data-driven decisions. It led to an **increase** in customer base by **15%**