CHINMAYI HEGDE

San Jose, CA 95126 | +1 669-649-3847 | chinmayilokeshwar.hegde@sjsu.edu | GitHub | LinkedIn | Portfolio

MS in Computer Science - San Jose State University, San Jose, CA — Expected May 2024
BE in Computer Science - PES University, India — Aug 2020
Relevant Coursework - Big Data Processing, Cloud Computing, Artificial Intelligence
Publications - Vehicle Trajectory Prediction using GANs, Autonomous Defense Device with GSM and NLU

TECHNICAL SKILLS

Programming: Python, R Programming, Javascript, Java
Databases: SQL, BigQuery, NoSQL - MongoDB, Cassandra, Neo4j
Tools and Technologies: Git, REST, Flask, Google DataStudio, TensorFlow, Apache Spark, PySpark, Spacy, Pandas, Numpy, Sci-kit, Keras, OpenCV, Streamlit, Data visualization - Tableau, Power BI, Looker, Matplotlib, Airflow, Vertex AI
Cloud Technologies: Google Cloud (GCP), Microsoft Azure, Docker
ML/AI: Supervised Learning, Unsupervised Learning, Entity Extraction, Natural Language Processing, Sentiment Analysis, Recommendation System, Face Recognition, Audience Development, Image processing, Video processing

PROFESSIONAL EXPERIENCE

WebMD – NY, USA (Remote) — Jun 2023 – Aug 2023
Data Science Intern
- Engineered **machine learning pipelines** for recommendation systems using **BigQuery,** Python on **Vertex AI,** and **Airflow**, to accurately target potential customers.
- Conducted **clustering** and **time series forecasting** using **Python** for customer segmentation for targeted marketing campaigns, increasing market penetration by 15%
- Devised a data extraction solution using **GCP's Generative AI API** and **Natural Language Processing (NLP)** to extract details from SQL **BigQuery** database of 30k products, resulting in extraction accuracy of 95%, with exploratory analysis on **Tableau.**

Merkle – Bangalore, India — Sep 2020 – Jun 2022
Data Scientist
- Integrated an ML **anomaly detection** feature into an in-house tool, for **time series** data with ARIMA and **PySpark**
- Orchestrated the **migration** of critical ETL scripts from Analytics Workbench to the **Google Cloud Platform (GCP)** for scalability, demonstrating expertise in cloud-based data engineering
- Integrated Python **Natural Language Generation** and **NLP** scripts for automated KPI reports in a **Tableau** dashboard
- **Automated** a Python script for data analysis, and prioritization in data products, for the development of a data dictionary
- Developed **ETL data pipeline** with GCP **Airflow** and **Python** for 2 million records across 60 tables from BigQuery (SQL), and Adobe Analytics to ensure data quality

Merkle – Bangalore, India — Jan 2020 – Mar 2020
Software Development Intern
- Embedded a ResNet **object detection** and tracking model with 89% accuracy, using **Keras** and **Streamlit**
- Conducted fake review detection for client's brand analytics using Selenium web scraping to gather customer reviews and utilized neural networks (CNNs) with **PySpark** for large-scale classification in **Python**

PROJECTS AND PUBLICATIONS

Multi-Lingual Image Description Assistant | *Hugging Face, LLM, Image processing, NLP* — GitHub
- The application takes an image input to produce its description in audio output in the selected language
- Employed Hugging Face's image-to-text model, OpenAI's LangChain for language translation, and Hugging Face's text-to-speech model to create a multilingual audio description from user-uploaded images

Deep Learning-based Cyberbullying Prevention using NLP | *Transformers, Neural networks, Flask* — GitHub
- Utilized BERT, PySpark, and various deep learning models (LSTM, RNN, CNN) for proactive cyberbullying detection. Named Entity Recognition (NER) and Topic Modelling (LDA) for context mining/extraction with Flask for real-time blocking, with Python.
- Achieved 92.7% accuracy.

Travel Plan Recommendation Web App | *Python, ReactJS, Flask, Cassandra* — GitHub
- Built a ReactJS and Flask-driven web app with Cassandra (NoSQL), showcasing API integration, database utilization (CRUD)
- Integrated ML recommendation system into Flask backend for travel suggestions

Vehicle Trajectory Prediction using GANs | *Image/Video Processing, Neural networks, TensorFlow* — GitHub
- Devised a collision avoidance model using video processing with TensorFlow YOLOv3 for object detection and tracking, and Generative Adversarial Networks to predict vehicle trajectories. Achieved 88-93% accuracy.

LEADERSHIP AND VOLUNTEERING

- Student organizer - IEEE ICACCI, Genesis DevCon
- Head of Design - TEDxPESU (2018, 2019), IEEE GirlsGeekHack and inGenius
- Website Optimization Student Assistant - Department of Computer Science