

CHINMAYI HEGDE

San Jose, CA 95126 | +1 669-649-394 | chinmayilokeshwar.hegde@sjsu.edu | [GitHub](#) | [LinkedIn](#) | [Website](#)

SUMMARY:

I am a graduate student at San Jose State University with a passion for Data Science. I have extensive experience designing machine learning solutions from ETL pipelines, algorithms, databases, and websites, with a track record of successful projects in cloud-based data science and software development.

MS in Computer Science - San Jose State University, San Jose, CA

Expected May 2024

- Website Optimization Student Assistant - Department of Computer Science

BE in Computer Science - PES University, India

Aug 2020

- Student organizer - IEEE ICACCI, Genesis DevCon
- Head of Design - TEDxPESU (2018, 2019), IEEE GirlsGeekHack and inGenius

Relevant Coursework: Design & Analysis of Algorithm, Database System Principles, Cloud Computing, Machine Learning

TECHNICAL SKILLS

Languages: Python, R Programming, JavaScript, Java, TypeScript

Machine Learning: Keras, TensorFlow, Matplotlib, scikit-learn (sklearn), Pandas, NumPy, Apache Spark

Databases: SQL, BigQuery, NoSQL - MongoDB, Cassandra, Neo4j

Tools: Git, Google Cloud (GCP), Streamlit, Data visualization - Tableau, Power BI, Looker

Web development: Django, Flask, ReactJS, HTML, CSS

PROFESSIONAL EXPERIENCE

WebMD – NY, USA (Remote)

Software Development Intern, Data
2023

Jun 2023 – Aug

- Engineered **ML pipelines** for recommendation system with feature engineering, and **classification** models using BigQuery (SQL) and **Python** on Vertex AI and **Airflow**, to accurately target potential customers
- Conducted **clustering** techniques and multivariate **time series forecasting** using **Python** for customer segmentation and behavior prediction, resulting in targeted marketing campaigns, increasing market penetration by 15%
- Devised a data extraction solution using **GCP's GenAI API** and **Natural Language Processing (NLP)** to extract details from SQL **BigQuery** database of 26000+ products, resulting in extraction accuracy of 95%

Merkle – Bangalore, India

Data Scientist
2022

Sep 2020 - Jun

- Integrated an **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
- Developed **ETL data pipeline** with **Python** and Google Cloud Platform **Airflow** for 2M cross-functional web analytics records from BigQuery (SQL), Adobe Analytics, and Vertica to ensure data quality for 55+ tables improving scalability
- Integrated R and Python **Natural Language Generation** and NLP scripts for automated KPI reports in a **Tableau** dashboard
- Automated** a sophisticated **Python** script for data extraction, analysis, and prioritization in data products, for the development of a data dictionary
- Created and led courses in R, Python, and SQL focused on ARIMA forecasting, and data analytics for new analysts

Merkle – Bangalore, India

Software Development Intern

Jan 2020 - Mar 2020

- 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
-

ACADEMIC PROJECTS

- Multi-Lingual Image Description Assistant ([GitHub](#)):** Developed a **Streamlit web application** that employs 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
- Real-time Cyberbullying Prevention using NLP and Classification ([GitHub](#)):** Developed a Python-based system using **NLP**, leveraging **Spark** and **deep learning models (LSTM, RNN, CNN)** for proactive cyberbullying detection. Showcased strong software development skills with **Flask integration** for real-time blocking
- Travel Plan Recommendation Web App ([GitHub](#)):** Built a ReactJS and Flask-driven web app with **Cassandra (NoSQL)**, showcasing API integration, database utilization (CRUD) and end-to-end development skills. Integrated **ML recommendation system** (collaborative filtering-based) into Flask backend for personalized travel suggestions
- Vehicle Trajectory Prediction using GANs ([IEEE Publication](#)):** Devised a collision avoidance model using **video processing** with **TensorFlow YOLOv3** for object detection and tracking, Generative Adversarial Networks to predict vehicle trajectories