

Chinmayi Hegde

+1 669-649-3847 | chinmayi13hegde@gmail.com | [GitHub](#) | [LinkedIn](#) | [Portfolio](#)

WORK EXPERIENCE

Senior Data Scientist

Jan 2025 – Present

Walmart

Bentonville, AR

- Built an AI workflow using Python to help non-developers assess use case risk, increasing adoption of internal AI tools and reducing friction. This was built as a part of the Global Governance Hackathon 2026 and was selected as Top 5.
- Worked on AI Fairness Standard to compute disparity metrics and recommend mitigation techniques for model training for models in fraud, employment, and finance domains, while collaborating with People Tech and employment legal teams.
- Evaluated 200+ machine learning models to identify performance informing product deployment decisions and improving reliability across use cases.

Data Scientist (Contract)

Nov 2024 – Jan 2025

NASA AI for Life in Space (via Crowdplat)

Pleasanton, CA

- Built a deep learning model (Python - pandas, numpy, scikit-learn) to classify rodent cells from spaceflight vs. ground control using image and tabular data.
- Implemented explainable AI (XAI) functionality to improve model transparency and interpretability using Streamlit.

Data Scientist II

Sep 2024 – Nov 2024

Chegg Inc.

Santa Clara, CA

- Built causal impact models with Python, and SQL to assess marketing campaign effectiveness across 11 social media platforms.
- Automated reporting for perks activation among subscribers using AWS Redshift (SQL), Python, and Tableau.
- Oversaw 5 A/B tests with Optimizely for the refund and payment pages to analyze key metrics, providing insights to reduce churn.

Software Development Intern

Jun 2023 – Aug 2023

WebMD

New York City, NY

- Used Python, SQL, and JupyterLab on historic customer data to ML propensity/regression and customer segmentation models with 92-94% accuracy, optimizing outreach for the email marketing team.
- Built an ML pipeline (Python, BigQuery, Vertica) for time series forecasting, delivering insights to 300K+ healthcare professionals.
- Finetuned BERT models with PyTorch and utilized Llama API Prompt Engineering for sequence labeling of 26,000+ product names.

Data Scientist

Jan 2020 – Jun 2022

Merkle

Bangalore, India

- Automated anomaly detection (Python, SQL) on 2.5 TB of behavioral data using time series forecasting, reducing time-to-alert and enabling proactive operational decision-making.
- Delivered 7+ Tableau dashboards powered by time series forecasting, and text summarization (NLP) using Python and SQL to optimize decision-making for advertisement and marketing.
- Designed data pipelines using Airflow, Python, and SQL, to integrate unstructured data from 20+ systems and fix data quality issues.
- Authored web scraping scripts with Python (Selenium, Scrapy) for data mining and used entity recognition for text preprocessing.

EDUCATION

Master of Science in Computer Science | San Jose State University, CA

May 2024

Bachelor of Engineering in Computer Science | PES University, India

Aug 2020

SKILLS

Languages: Python, SQL, R

Databases: PostgreSQL, Cassandra, MongoDB, Hive

Cloud Platforms: MySQL, GCP (Vertex AI, BigQuery, Looker Studio), Vertica, AWS (Redshift, DynamoDB, S3, Lambda)

Tools & Packages: Docker, Git/GitHub, Tableau, Power BI, Excel, Jira, Hugging Face, Spacy, Sci-kit learn, NLTK (NLP), Pandas, ScraPy

Certifications: [Architecting with Kubernetes](#), [Machine Learning](#), [Google Cloud Infrastructure](#)

[View all](#)

PERSONAL PROJECTS

Finance Analytics with Llama2 — Python, Tableau

- Built a financial forecasting tool (Python, Llama 2) for automated transaction categorization and predictive spend analysis, with Tableau dashboards for planning and trend monitoring.
- Created interactive Tableau data visualizations for spending trends and future spending forecasts.

Customer Churn Prediction — Python, SQL, ML Classification

- Developed a customer churn prediction model using XGBoost and random forests in Python, identifying key features of churn and achieving 87% model accuracy. Performed customer lifetime value analysis to segment customers, and identifying high-value customer groups.

PUBLICATIONS

Vehicle Trajectory Prediction using Generative Adversarial Networks (GANs)

[Read here](#)

- Built a Generative AI-driven system to predict the path of a moving vehicle to avoid collision using Tensorflow, YOLOv3, and GANs.

Autonomous Defense Device with Internet of Things (IoT) and Natural Language Understanding (NLU)

[Read here](#)

- Developed a safety device for self-defense with features such as AI voice assistant, and automatic SOS location reporting with GPS.