Manvitha Kalicheti

Open to full-time roles starting June 2024

manvithakalicheti@gmail.com | (470) 917-5179 | mkalicheti | mkalicheti | www.mkalicheti.com

Education

Georgia Institute of Technology **Aug. 2022 - May 2024**
Master of Science, Computational Science & Engineering, CGPA: 4.0/4.0 Atlanta, GA, USA
Coursework: Machine Learning, NLP, Time Series Analysis, Algorithms, Big Data Systems, Data Mining & Stat. Learning

Indian Institute of Technology, Hyderabad **Aug. 2018 - May 2022**
Bachelor of Technology, Mechanical Engineering, CGPA: 9.18/10.0 Hyderabad, India

Skills

Computer Languages :	Python (NumPy, Pandas, Scikit-learn, Matplotlib, PyTorch, statsmodel), R (ggplot2, tidyverse), C/C++
RDBMS & Big Data :	MySQL, MS SQL, Teradata, Hadoop, Hive, PySpark
Statistics :	Regression, Forecasting, Hypothesis testing, Causal Inference, Dimensionality Reduction
Machine Learning :	Clustering/Segmentation, Ensemble Methods, Recommender Systems, Propensity Scoring
Tools :	Tableau, BeautifulSoup, Streamlit, MS Excel, Git, Bash

Work Experience

Blue Cross Blue Shield of IL, MT, NM, OK & TX (HCSC) **Jun. 2023 - Aug. 2023**
Senior Data Analytics Intern, BI Data Engineering Chicago, IL

- Engineered two critical views from claim/member/provider datasets (**15B+** records) using **HiveQL** and **PySpark**, integrating **1TB+** data from **4+** sources and executing data reconciliation strategies to ensure quality.
- Collaborated with cross-functional teams to develop dynamic **Tableau** dashboards on healthcare utilization patterns.
- Conducted meticulous data lineage tracking over **26** advanced SQL transformations, documenting the journey of **48 fields** from source to final tables and pushing a strategic re-evaluation of the project's scope.

Georgia Institute of Technology **Aug. 2022 - Present**
Data Science, Health Analytics Centre Atlanta, GA

- Utilizing microsimulation and advanced statistical techniques to determine opioid prevalence across the US.
- Built linear, logistic, and negative binomial **regression** models using **R** to examine opioid consumption patterns.
- Implemented **bootstrapping** techniques to account for potential inflation in statistical significance.

Data Engineering, Health Analytics Centre

- Building data extraction pipelines from **10TB+** Medicaid claims and demographics database using **MySQL**.
- Utilized advanced SQL functionalities (**stored procedures, window functions**) to streamline data preprocessing and enhance efficiency by reducing the reliance on external R/Python environments.
- Addressing missing data challenges in **1M+** records using stratified sampling and historical mode imputation.

CrystalBall AI Pvt. Ltd. **Jul. 2021 - Aug. 2021**
Optimisation Intern Hyderabad, India

- Saved over **4 man-hours** per job by automating drone image and GPS data postprocessing using a **Python** API.
- Identified workflow bottlenecks and collected time and resource consumption data to build an edge computing device.

Selected Projects

Analysis of Factors Influencing Life Expectancy of a Country

- Built an 82% R^2 multiple linear regression model in **R** to estimate life expectancy, incorporating diverse factors.
- Employed advanced techniques: correlation matrix, VIF for multicollinearity; backward stepwise regression, ANOVA comparisons for variable selection; and thorough residual analysis to test assumptions.

Computational Data Analysis

- Designed and implemented kmeans and kmedoids algorithms for image compression.
- Built a matrix factorisation based personalised recommender system on the "Netflix Ratings" dataset.

Honors/Publications/Certifications

Published at *The Journal of Pediatrics* and *Frontiers in Public Health*

- Statistical Analysis on Opioid Utilisation**:** Regression Modeling on Pediatric Cardiac Surgery and Adult Trauma Hospitalized Cases in the Medicaid insured population.

Fellowships: George Family Foundation Fellowship, 2023-24

- Awarded a competitive $5000 award for **research excellence** in health systems at the graduate level.

Winner, Best Emerging Health Hack @ Georgia Tech's annual hackathon [Devpost]

- Built a recommender system to suggest less processed alternatives within a food group based on search terms.

Teaching Assistantship @ Georgia Tech: Statistical Modeling & Regression Analysis

- Instruction and clarification of technical concepts to **300+ OMSA** students; conversion of demo code from R to Python.

Microsoft Certified - Azure Fundamentals (AZ-900)