

# Sai Yeshwanth Mekala

7206594782 | yesh20@icloud.com | <https://www.linkedin.com/in/sai-yeshwanth-mekala-061917176/> | Boulder, CO.

## EDUCATION

### Master of Science in Data Science

University of Colorado Boulder, Boulder, Colorado, United States

GPA: 4/4

August 2023 – May 2025

### Bachelor of Technology in Information and Communication Technology

SASTRA Deemed University, Thanjavur, Tamil Nadu, India

GPA: 8.3/10

June 2017 – July 2021

## SKILLS

**Programming:** Python, R, Java, C++

**Data Visualization:** Tableau, Power BI, Excel

**IDE's:** Jupyter Notebook, R Studio, Eclipse, VS Code

**Others:** Statistics, Data Mining, Data Structures, Communication, scikit-learn, pandas, TensorFlow, Keras, Git

**Database:** SQL, MySQL, Oracle, MongoDB

**Cloud:** IBM Watson Studio, Google Cloud, AWS

**Languages:** English, Telugu (Native), Hindi

## PROFESSIONAL EXPERIENCE

### Systems Engineer

July 2021 – June 2023

Tata Consultancy Services, Hyderabad, India

- Collaborated with the Backend Team and developed a Virtual Career Fair web page for Stevens Institute of Technology by applying HTML, CSS, JavaScript, and PHP skills, contributing to 50% of the Project.
- Maintained and monitored the Student Information System of Excelsior University using SQL, PL/SQL and Java Skills and gained 100% response and resolution time.
- Guided interns in executing Broken Links Automation Testing for the course catalog page and eliminated broken links, yielding a 100% working links rate.
- Created Kanban board and Dashboards in Excel to showcase effectiveness, resulting in a 96% client satisfaction index.
- Accelerated the Knowledge Transfer process by at least 50% through the creation of Document of Understanding and Standard Operating Procedures (SOP) by collaborating with Data Engineers and Database Administrators.

## PROJECTS

### Heart Disease Identification from Patients' Social Posts, Machine Learning Solution on Spark

- Implemented predictive models by creating advanced ML algorithms in Pyspark on Cleveland Heart disease dataset, achieving an accuracy of 94.9% for the Random Forest Classifier through Hyperparameter tuning and K-fold cross validation.
- Developed an ETL pipeline to extract real-time data from the Twitter API using Apache Kafka for efficient ingestion and processing.
- Utilized Apache Spark to transform the streaming data into feature vectors and made predictions in real-time.

### Applying Dimensionality Reduction in Collaborative Filtering Recommender Systems.

- Created Singular Value Decomposition (SVD) and Non-Negative Matrix Factorization (NMF) models on the Movie Lens 1M dataset to predict Top N movies for a specified user ID.
- Conducted model evaluation using cross-validation and Grid Search and achieved 0.87 RMSE for SVD.

### Sports Retail Data Insights

- Performed statistical tests to verify the Stationarity of the data and conducted Time Series Analysis with ARIMA and developed an optimal model with (2,1,5) configuration to predict the operating profit.
- Applied Association Rule Mining (ARM) using the Apriori algorithm with minimum support of 50% and minimum confidence of 70% and generated frequently purchased item sets and associations which provides actionable insights.

### Coffee contains more Caffeine than an Energy Drink

- Performed Web Scrapping on Caffeine Informer database using BeautifulSoup and extracted caffeine and calories content.
- Conducted Lower and Upper tailed Two Sample Hypothesis testing at a significance level of 5% and found strong evidence to conclude that there is more caffeine content in Coffee compared to an Energy Drink.

## PUBLICATIONS

### Breast Cancer Detection Using Machine Learning ([ijatcse12922020.pdf](#))

April 2020

- Performed Exploratory Data Analysis on Breast Cancer Wisconsin dataset and implemented a high performing Logistic Regression model with an accuracy of 99.3%.
- Collaborated with a professor and published the research work in the International Journal of Advanced Trends in Computer Science and Engineering and gained 24 citations highlighting impactful research.

## AWARDS

- Achieved On-the-Spot award for excellent client feedback.

March 2022

## INTERESTS

- Machine Learning, Big Data Analytics