

Sai Yeshwanth Mekala

7206594782 | yesh20@icloud.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#) | Boulder, Colorado.

PROFESSIONAL EXPERIENCE

JOHNS MANVILLE

Littleton, Colorado

Data Scientist Intern

May 2024-Present

- Performed Sequential and Supervised Deep Learning (TensorFlow, Keras) predictive modeling including Long Short-Term Memory (LSTM), Bidirectional LSTM and Gated Recurrent Unit (GRU) on scaled time series, gaining 70% R-Squared.
- Developed state-of-the-art Deep Learning Transformer with attention mechanism in Python (Pytorch), achieved 0.09% MAPE.
- Revamped current process with 31% improvement by employing Explainable AI (SHAP Explainer), reported KPIs and waterfall charts enhancing Anomaly detection yielding actionable insights for Technical and Non-Technical Stakeholders.
- Spearheaded automated categorization of operator comments for Fiber Cut Quality through Zero-shot text classification in Python using RoBERTa-MNLI AI model and reduced manual inspection time by 100%.
- Led the Data Analysis in R through Correlation analysis, Upper Tailed and Lower tailed Hypothesis Testing (z-test) and Hypothesis Testing with Randomization at 5% significance level and found 100% evidence of Feature Independence.
- Led the debugging and optimization of SQL Stored Procedures and Logs in SSMS and SAP, achieved 100% accurate Stats Reports for Non-Technical Stakeholders.
- Architected an efficient ETL pipeline, integrating large data from sources including SQL Servers, Salesforce, SAP, Seeq API, and Historians into Jupyter Lab, resulting in a 75% improvement in data collection and preprocessing time.

TATA CONSULTANCY SERVICES

Hyderabad, India

Systems Engineer

July 2021-June 2023

- Developed SQL, PL/SQL queries in Oracle and test plans as a Database Administrator, leveraging existing ETL pipeline (SQL servers and Jenkins) as a Data Engineer for incident analysis and gained 100% resolution time.
- Expedited KT process by 50% by documenting cross-functional Standard Operating Procedures (SOP) in Word with SQL scripts.
- Collaborated with manager and created Dashboards in Excel for SMEs, resulting in 96% Client Satisfaction Index.
- Earned On the Spot award in 2022 for exceptional client feedback for Data Analysis and automation and 15% boost in CSI.

PROJECTS

CloudSteps: Personalized Steps Analytics using Apple Health Data on Google Cloud Platform

- Implemented a cloud-based service and processed over 1GB XML data, leveraging BigQuery, Message Queue and Cloud Functions.
- Deployed the containerized REST API based Flask app using Docker on Kubernetes, achieving 99.9% uptime and handling 100+ concurrent requests and delivering real-time analytics with interactive Plotly visualizations.

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

- Executed RF classifier in Pyspark, achieving an accuracy of 94.9% by Hyperparameter tuning and K-fold cross validation.
- Streamlined real-time Big Data ingestion by integrating Apache Spark and Kafka framework, saved 50% processing time.

Data-Driven Meteorology: Unveiling Insights from Weather Data

- Applied K-Means and Hierarchical clustering for 100% optimal clusters (5) leveraging Elbow and Silhouette Method.
- Generated top 10 rules by Lift and Confidence by applying Unsupervised learning using Association Rule Mining (Apriori).
- Executed Supervised Machine Learning algorithms including Gaussian and Multinomial Naïve Bayes, Support Vector Machine, Decision Tree and Artificial Neural Network (ANN) and recorded 70.28% Recall/sensitivity.

PUBLICATIONS

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

April 2020

- Performed Exploratory Data Analysis, implemented Logistic Regression classification ML model with 99.3% Accuracy.
- Collaborated with a professor and published the open-source research and gained 24 citations highlighting impactful research.

EDUCATION

UNIVERSITY OF COLORADO BOULDER

Boulder, Colorado, United States

Master of Science in Data Science

August 2023-May 2025

SASTRA DEEMED UNIVERISTY

Thanjavur, Tamil Nadu, India

Bachelor of Technology in Information and Communication Technology

June 2017-July 2021

SKILLS

Programming: Python, R, Java, C++, Pandas, NumPy, Mathematics, Statistical Modeling, Hypothesis Testing, KNN, Predictive Analytics, Software Development, Natural Language Processing (NLP), Relational Databases, MySQL, MongoDB, Data Structures, Scikit-Learn, Git, IBM Watson Studio, GPC, Web scrapping, Azure ML, Databricks, OpenCV, ML Ops

Tools: Jupyter Notebook, R Studio, Eclipse, VS Code, Spyder, Minitab, SQL Server Management Studio, Hadoop, AWS

Data Visualization: Tableau, Microsoft Power BI, Matplotlib, Seaborn, Plotly, ggplot2, Excel, Azure Data Studio

Interests: Statistics, Data Mining, Transformers, LLM, RAG, Multimodal Models, Artificial Intelligence (AI), Big Data Analytics