

Yeswanth Arasavalli

maneswararasavalli@gmail.com | +91 6303321322 | Hyderabad, India

Summary

Passionate professional in the field of Data Science which includes hands on skills in cloud computing, machine learning, and data visualization. Also very skilled in Python, R, SQL, and Tableau, also have a strong background in NLP, deep learning, and big data tech such as Hadoop and Spark. I am experienced in the development of large scale ML models, automation of data pipelines and in the process of generating insights which in turn produce measurable business results. Also very much into the idea of contributing to AI which is applied in the fields of product development, financial analytics, and enterprise solutions, a mix of technical skill with collaborative approach to problem solving.

Education

Masters in Computer Science (Big Data Analytics) | 2021 - 2023

Central University of Rajasthan, Ajmer, Rajasthan | **GPA:** 7.3/10

Relevant Courses: Machine Learning, Data Mining, Statistical Analysis, Big Data Analytics, Python, R, Operations Management

Academic Highlight: *Capstone Project:* Churn prediction using ensemble models (XGBoost, Random Forest); achieved 88% accuracy on telecom dataset with 200K+ records

Bachelors in Computer Science | 2018 - 2021

Adikavi Nannaya University, Rajahmundry, Andhra Pradesh | **GPA:** 8.37/10

Relevant Courses: Data Structures, DBMS, SQL, Java, C, HTML/CSS, Calculus, Mathematics, Excel & Power Point

Skills

- **Languages:** Python, R, SQL, Java, C
- **ML & AI:** Scikit-learn, TensorFlow, PyTorch, Keras, NLP, Time Series, CNN, LSTM
- **Data Wrangling:** Pandas, NumPy, Feature Engineering, Cleaning, ETL, BigQuery
- **Visualization:** Tableau, Power BI, Matplotlib, Seaborn
- **Cloud & Big Data:** AWS, Azure, GCP, Hadoop, Spark Streaming, PySpark
- **Deployment Tools:** MLflow, Docker, Kubernetes
- **Databases:** SQL Server, PostgreSQL, MongoDB
- **Soft Skills:** Communication, Problem-Solving, Team Collaboration, Time Management

Work Experience

Data Annotator & Prompt Engineer (Freelance) | *Telus International*

- Accurately labeled and categorized diverse datasets to train and improve AI algorithms, ensuring high data quality.
- Designed and refined prompts to optimize model performance and enhance the precision and relevance of AI-generated responses.

Data Science Intern | *Postulate IRT*

- Cleaned and processed raw speech NLP sets (150K+ records) which sees a 30% improvement in data integrity.
- Developed out of the box models for marketing which in turn improved our forecasting accuracy by 22%.
- Created Tableau reports that support cross functional strategies and improve campaign performance.

Data Analyst Intern | *Edunet Foundation*

- Built out cloud and on premise data pipelines with Python and SQL which in turn reduced query latency by 25%.
- Developed decision tree and random forest models.

- Visualized which customer segments and model performance via Power BI dashboards.

Data Science Intern | PHN Technology Pvt. Ltd.

- Automated ETL processes for large scale log datasets which in turn simplifies data processing.
- Built out interactive dashboards in Tableau and Excel which saw a 40% reduction in reporting time.
- Applied statistics in R and used IBM Watson Studio for model implementation.

Course Certifications

- *Google Data and Advanced Data Analytics from Coursera*
- *Johns Hopkins University Data Science Specialization from Coursera*
- *Pyspark, Microsoft PowerBI from Infosys Springboard*
- *Tableau from iNeuron*
- *MongoDB Python Developer Associate from MongoDB University*
- *Machine Learning / Artificial Intelligence Engineer from Skillsoft & Infosys Springboard*

Projects

Retail Image Recognition with CNN

- Developed a CNN-based image classifier for 50+ retail product classes, achieving an accuracy of 92%.
- Launched via cloud pipeline and visualized in Tableau dashboards reducing inventory errors by 35% and increasing supply chain reliability.

Predictive Maintenance with LSTM

- Forecast we see 87% accuracy in which we predict machine breakdowns two weeks in advance via use of LSTM models on time series sensor data.
- Enabled over \$1.15M in annual savings with proactive scheduling and anomaly detection dashboards which also did that of predictive analytics that increases operational efficiency.

Multimodal Emotion Detection for Personalization

- Engineered an AI which identifies emotions in text, audio, and video streams with 85% accuracy.
- Enabled personalized recommendations which increased user engagement by 24% and increased revenue by 14% - all through data driven personalization.

Amazon Food Review Analysis

- Performed end-to-end sentiment analysis on **568K+ Amazon food reviews**, including preprocessing, label engineering, and exploratory analysis.
- Built baseline TF-IDF + Logistic Regression model achieving **~89% accuracy**, then fine-tuned **DistilBERT**, improving performance to **~94% accuracy and F1 score**.
- Developed inference pipeline and deployed an interactive Gradio web app on **Hugging Face Spaces**, enabling real-time review sentiment prediction.

Electric Load Prediction

- Built out models which achieved a MAPE of 2.5% for forecasting hourly electricity demand which we did via regression and time series forecasting.
 - Identified out of the ordinary load patterns for use in resource optimization and financial metric validation in energy and utility domain
-