

# ANUKEERTHI REDDY POTHEPALLI

[apothepalli1@student.gsu.edu](mailto:apothepalli1@student.gsu.edu) | (470) 838-5598 | [LinkedIn](#) | [Portfolio](#)

## WORK EXPERIENCE

### Georgia State University

Atlanta, Georgia, U.S.A

Data Science Research Assistant

Aug 2023 - Current

- Performed time series analysis using the Prophet model to evaluate new product seasonality, trends and forecasted product post-launch sales using AutoTS ARIMA, leading to a 20% improvement in sales target achievement.
- Queried the cube datastores using SQL, and utilized Pyspark for analyzing, modeling and extracting insights from large-scale data, refining retailer order quantities, and enhancing post-launch effectiveness by 10%.
- Designed a Tableau dashboard to analyze product launch curves leveraging financial data, Python NLTK and BERT based embeddings to analyze unstructured product review text data, resulting in a 20% improvement in market performance.
- Developed and fine-tuned AI-driven sentiment analysis(NLP) pipelines leveraging BERT-based models; improved customer feedback analysis workflows resulting in a 40% boost in response strategy effectiveness within three months of deployment.
- Optimized statistical models with R programming using Maximum Likelihood Estimation (MLE) to predict customer satisfaction stages and identify factors driving customer choices, enhancing decision-making processes by 15%.

### Tata Consultancy Services (TCS)

Bangalore, India

Data Analyst, Team Lead

Apr 2021 - Jul 2023

- Built and maintained data-driven dashboards in Power BI and Tableau, providing actionable insights on Metrics, KPIs, and complaint closure timeliness, resulting in a 15% improvement in stakeholder decision-making efficiency.
- Implemented Unsupervised Machine Learning and Text Analytics models using Scikit and SpaCy packages to analyze unstructured service data of medical devices, reducing file processing time by 10% and improving accuracy by 18%.
- Managed a team of 4 analysts to extract manufacturing data from SAP database using SQL Queries and conduct feature engineering on medical device service records, enhancing post-market surveillance accuracy by 20%.
- Optimized daily operations by automating tasks with Microsoft Excel Macros and Salesforce ECM Web automation using Python Selenium and advanced analytics, achieving a 30% improvement in Resource Utilization.
- Led a top-performing team achieving zero escalations for 11 months and improving customer satisfaction by 40% while utilizing advanced AI models for Quality System compliance and regulatory adherence in the medical device industry.

## PROJECTS

### ECG Classification with CNN and LSTM

Georgia State University, Atlanta

- Created and implemented a robust AI model (Utilizing CNN + LSTM) for Multi-class ECG classification, which allowed real-time monitoring of heart rate patterns; the algorithm processed data from over 1,000 ECG recordings to ensure reliability in clinical settings with an accuracy equaling the SOTA(99%).

### Scalable Recommendation System Optimization with PySpark and HDFS

Georgia State University, Atlanta

- Engineered a scalable recommendation system with PySpark and RDDs, enhancing computation efficiency, enabling analysis and recommendation modeling on a large dataset of around 5 million data points. Used HDFS for reliable storage, reducing data access latency by 30% and cutting cloud infrastructure costs by 25%.

## EDUCATION

### Georgia State University, J. Mack Robinson College of Business

Atlanta, GA, United States

Master's in Data Science and Analytics (Business Concentration)

Aug 2023 - Dec 2024

Courses: Machine Learning, Data Programming, Statistical Foundations, Scalable Analytics, Predictive Analytics, Deep Learning Analytics, Business Intelligence, Data Management, Data Visualization, Text Analytics, Graph Analytics

## SKILLS

**Competencies:** Linear Regression, Gradient Boosted Machines, Random Forests, Text Processing, Transformers, Word Embedding (Word2Vec, BERT), SARIMAX, Prophet, LIME, SHAP, Partial Dependence Plots, Parameter estimation, Hypothesis Testing, ANOVA, Statistical Distributions, Large Language Models(LLM), Neural Networks

**Programming Languages:** Python, SQL, R-Programming, NoSQL

**Tools/Frameworks:** Power BI, Microsoft Excel, Pandas, Numpy, StatsModels, Matplotlib, Seaborn, PySpark, Sklearn, TensorFlow, Keras, Tableau, Google Cloud Platform, Azure Machine Learning Studio, MS Fabric, Asana, Git

## CERTIFICATIONS

- DataCamp: Data Analyst Associate Certificate

Jan 2024