

Prashanth Kumar Gunda

(954) 560-2208 • gundapr@mail.uc.edu • Cincinnati, Ohio • [LinkedIn](#) • [GitHub](#) •

EDUCATION

University of Cincinnati, Carl H. Lindner College of Business

Master of Science in Business Analytics

Indian Institute of Technology, Hyderabad

Master of Technology in Thermo-Fluid Engineering

Cincinnati, Ohio

Expected: Aug 2024

Hyderabad, India

Aug 2021

SKILLS

- **Analytical, and Visualization tools:** SQL, R, NoSQL, Excel, Tableau and PowerBI
- **Programming tools:** C, C++, Python, MATLAB, Java, Scala, Streamlit, HTML, CSS
- **Deep Learning:** PyTorch, Keras, TensorFlow
- **NLP:** NLTK, RAG, LLM, Prompt Engineering, PEFT, LoRA, LangChain, GraphRAG
- **Machine Learning:** Regression, Clustering, Classification, Gradient Boosting, Neural Networks, Sklearn, Scipy
- **Cloud Technologies:** AWS, AWS ECS, Azure, Kubernetes, GCP

WORK EXPERIENCE

Data Scientist Intern.

April 2024- June 2024

GEN1E Life Sciences, INC

San Francisco, CA

- Developed **ML pipeline** using **MLOps** to identify ARDS patient endotypes.
- Identified the base-line parameters by **data manipulation** techniques using libraries such as **Pandas and Numpy** and performed advanced EDA on datasets of over **300,000** patients to extract features for endotype classification.
- Utilized **A/B testing** to compare models. Achieved the highest **ROC-AUC score of 0.98** with a **SVM** model.
- Used Logistic regression to check impact of HTE classification and endotype classification on model's effectiveness.

Graduate Assistant

Aug 2023-Dec 2023

University of Cincinnati

Cincinnati, OH

- Designed a **novel pipeline** using **Clinical BERT** for advanced **Named Entity Recognition (NER)** and **sentiment analysis**, optimizing the extraction of social determinants from diabetes patient records
- Analyzed a vast dataset of **10M** electronic health records (EHRs) using **SQL, PyTorch**, revealing SDOH for diabetes.
- Optimized **Clinical BERT** with **PEFT & LoRA** leading to **10% decrease** in computational time.
- Built a chatbot utilizing **RAG** and **LangChain** techniques and **Streamlit** to fetch patient records enabling user interaction about social determinants. This enhanced targeted care, **reduced time** to access patient insights by 25%.

Data Analyst

Mar 2022 – July 2023

Infosys

Hyderabad, India

- Orchestrated a **pipeline** for identifying **fraudulent** insurance claims using a database exceeding **100M** records, with data curated and compressed using **Hadoop** and **Hive**, thus reducing data processing time by **40%**.
- Utilized **ETL** tools and integrated **SAS** for advanced data processing and preprocessing.
- Conducted **EDA** using Python libraries, and employed **Spark** with **Hadoop** for **distributed data** processing.
- Implemented **ensemble** techniques on **Random Forest** and **XGBoost** to achieve a model of **78%** accuracy in fraud detection, contributing to a projected **\$1.5M** increase in annual company profit.
- Deployed the final models using **AWS** and **AWS ECS**, and integrated **MLflow** reducing deployment time by **35%**
- For claim amount data, visualized data trends and seasonality, reducing forecasting errors by **20%**, which improved financial planning and contributed to a **\$500,000** reduction in forecasting discrepancies.
- Built an **ARIMA** model using **statsmodels** for accurate forecasting, improving forecast accuracy by **30%**.

Data Analyst

Jun 2021 –Jan 2022

BYJU'S

Bengaluru, India

- Designed a web platform using **HTML, CSS**, and **JavaScript** to capture and log user engagement times with Maths content, tracking over **50,000** overinteraction patterns and employed **Git** for version control.
- Extracted and processed the engagement data using **SQL**, applied **K-Means clustering** to identify **5 distinct patterns** (**silhouette score of 0.84**), & optimized content delivery strategies, resulting in a **20%** increase in engagement.

Research Associate

Aug 2020 – July 2021

Indian Institute of Technology, Hyderabad

Hyderabad, India

Project: Machine Learning Application in Cavitation Induced Vortex Dynamics

- Developed unsupervised model (**K-means**) that achieved **86%** accuracy in predicting optimal cluster membership under varying environmental conditions like liquid density, temperature and viscosity on bubble interactions.