

Krishna Apurva

+1 (469)-880-8745 | krishnaapurva469@gmail.com | linkedin.com/in/krishnaapurva469 | github.com/krishnaapurva

EDUCATION

The University of Texas at Dallas

Master of Science, Business Analytics

Courses: Advanced Statistics, Business Analytics with R, Database Foundation for Business Analytics, Big Data, Prescriptive Analytics, Predictive Analytics for Data Science

Expected May 2025

GPA: 4.0

Medicaps University, Indore

Bachelor of Technology, Information Technology

May 2020

GPA: 3.6

TECHNICAL SKILLS

Languages: R, Python (NumPy, Pandas, Matplotlib, Scikit-learn), Java, C/C++, HTML, Shell scripting

Databases: SQL, PostgreSQL, Oracle, Hadoop, Hive, NoSQL, MongoDB, Teradata, PL/SQL, ER Model

ETL: Talend, ETL Design and Development, Databricks, DataStage, Control-M, Data Warehouse

Tools/IDE: Azure, MS Office, Power BI, Tableau, Git, Spark, AWS, JIRA, R Studio, Jupyter, VS Code

Data Science: Statistical Analysis, Data Visualization, Regression, Feature Scaling, Decision Tree, Random Forest, Convolutional Neural Networks, Classification, Probability, Clustering

WORK EXPERIENCE

Data Engineer | Innova Solutions

May 2021 - June 2023

CAPE – Centene Alternate Payment Engine Project

- Engineered a solution using stored procedures, CTE, and partitioning clauses, complimented by microservices, to filter the data flowing into **EDW downstream (Teradata)**, ensuring data integrity and boosting overall efficiency by 23%
- Leveraged **Talend** optimization functionalities with parallel processing to streamline complex **ETL pipelines**, resulting in a 20% decrease in data processing time and enhancing report generation speed
- Enhanced data storage efficiency by 35% through designing and implementing a **streamlined Talend workflow** that facilitated **NoSQL (MongoDB)** storage in **multi-tiered nested JSON** files and shortened query response time by 25%

Provider RTR Project

- Optimized **DataStage ETL** workflows for Healthcare data, reducing processing time by 27% through strategic redesign of transformation processes and implementing parallel processing methodologies
- Automated processes using **PL/SQL procedures** to streamline daily tasks such as data backups, indexing, and performance monitoring, resulting in an 18% improvement in overall operational efficiency within the database

Business Analyst | J. J. Enterprises

August 2020 - May 2021

- Extracted and analyzed data from SQL and Excel sources, utilizing advanced analytics tools, to ensure comprehensive and accurate datasets
- Implemented ad-hoc SQL queries and Tableau for robust **data analysis and visualization**. Evaluated sales and market research insights to inform strategic decision-making on new products, resulting in an 8% revenue increase
- Designed and maintained multiple **Tableau dashboards** for real-time monitoring of key performance indicators (KPIs), refining customer segmentation and contributing to a 10% rise in business revenue

Data Science Intern | Medicaps University

June 2019 - August 2019

Campus Surveillance Using Drone Project

- Achieved precise identification and tracking of individuals and moving objects through the implementation of OpenCV, enabling real-time monitoring of suspicious activities
- Orchestrated the development of the Yolo V3 object detection algorithm and achieved the best performance with a processing speed of 45 frames per second, surpassing algorithms by a 200% faster rate

PROJECTS

Car Insurance Claim Prediction (R, EDA, Data Mining, Data Cleaning, Oversampling, Undersampling)

- Balanced dataset by addressing a **6% (Claim)** and **94% (No claim)** data imbalance through **oversampling techniques**
- Employed various predictive modeling techniques, including **Logistic Regression, Random Forest, Decision Tree, and Neural Network**, for Car Insurance Claim prediction with 82% accuracy and 61% AUC

HealthCare Management System (PostgreSQL, Entity Relationship Modeling, Database Design)

- Integrated hospital operations with ER Modeling, DDL/DML implementation, and advanced queries, Triggers, Procedures, Views, and Functions for efficient healthcare management system reporting