KRISHNA APURVA

Dallas, TX | (469) 880-8745 | krishnaapurva469@gmail.com | linkedin.com/in/krishnaapurva469 | github.com/krishnaapurva

EDUCATION

The University of Texas at Dallas May 2025
Master of Science - Business Analytics and Artificial Intelligence **GPA: 4.0**
Awarded Lars Magnus Ericsson Fellowship in Management (Endowed) for academic excellence and leadership

Medicaps University May 2020
Bachelor of Technology - Information Technology **GPA: 3.6**

TECHNICAL SKILLS

- **Languages**: SQL, NoSQL, Python (Pandas, Matplotlib), R, Apache Spark, Java, Linux
- **Databases**: PostgreSQL, MySQL, MongoDB, Teradata, PL/SQL, Data Modeling
- **ETL:** Talend, ETL Design, Airflow, TAC Server, Control-M, Data Warehouse
- **Cloud:** Azure (ADF, Databricks, Data Lake), AWS (S3, Lambda, Redshift, Glue)
- **Data Science:** Statistical Analysis, Regression, Classification, Clustering, Feature Scaling, EDA, T-test
- **Tools/IDE:** Tableau, Power BI, MS Excel, Git

EXPERIENCE

Data Engineer | *Innova Solutions* **May 2021 - June 2023**

- Orchestrated data pipelines utilizing TAC Server for over 20+ Talend jobs, ensuring smoother execution and reducing processing time through sequential task management
- Engineered ETL pipelines with Talend staging and parallel processing, achieving faster processing and accurate business reports from over 10 million records
- Managed creation and modification of tables, views, functions, and stored procedures in PostgreSQL, MongoDB, and Teradata, aligning with project needs and optimizing database functionality for efficient data operations
- Spearheaded automation of quarterly report generation using Talend and Control-M, streamlining workflow efficiency by 30% and ensuring timely, consistent delivery
- Collaborated with cross-functional teams to enhance information interoperability, reduce integration issues, and improve overall system performance, leading to smoother operations and fewer errors

Data Analyst | *J. J. Enterprises* **August 2020 - May 2021**

- Designed and maintained Tableau dashboards for monitoring KPIs, refined customer segmentation, and contributed to a ~10% rise in business revenue
- Conducted in-depth analysis with SQL queries and Tableau to evaluate sales figures and market research information, leading to strategic product decisions and ~8% increase in quarterly revenue

Data Science Intern | *Medicaps University* **June 2019 - August 2019**

- Implemented a real-time object detection system using a YOLO V3 deep learning model in Python, exhibiting 78% accuracy in detecting and monitoring suspicious activities

PROJECTS

Tokyo Olympic Azure Data Pipeline | *Azure Data Factory, Azure Data Lake, Databricks, Power BI*

- Engineered data ingestion using ADF, automated extraction of 11k records from GitHub to Azure Data Lake
- Transformed data with Azure Databricks using Spark and crafted Power BI dashboards

Olist Ecommerce ETL Data Pipeline | *Python, Pandas, ETL, PostgreSQL, Git Desktop*

- Pioneered ETL pipelines utilizing Pandas to extract, transform, and load ~1M records into PostgreSQL
- Led PostgreSQL scaling by implementing indexing, views, and 3NF, enhanced performance and data integrity

Car Insurance Claim Prediction | *R, EDA, Data Mining, Data Cleaning, Oversampling, Undersampling*

- Conducted EDA and mitigated skewed datasets by implementing oversampling to imbalanced datasets
- Employed various predictive modeling techniques (Logistic Regression, Random Forest, Decision Tree, and Neural Network) for Claim prediction and achieved 82% accuracy and 61% AUC