Dheeraj Rahul Reddy Piduru

Dallas, Texas, 75252 | +1 (945-308-0889) | pidurudheeraj52@gmail.com | LinkedIn

EDUCATION

The University of Texas at Dallas **Dec 2025**
Master of Science, Business Analytics GPA: 3.72
Relevant Coursework: Advanced Statistics for Data Science, Applied Machine Learning, Applied Deep Learning, Predictive Analytics for Data Science, Prescriptive Analytics, Big Data Analytics, Applied Econometrics

Mahindra University, Hyderabad **Aug 2023**
Bachelor of Technology, Computer Science & Engineering GPA: 3.00

SKILLS

Programming Languages: Python, R, SQL, Shell Scripting, Java
Data Analysis & Visualization: Excel, Tableau, Power BI, Looker, Matplotlib, Seaborn, DAX
ETL & Data Warehousing: Apache Airflow, AWS Redshift, AWS S3, Snowflake, Google BigQuery, Informatica
Database & Version Control: MySQL, PostgreSQL, SQL Server, Git
Data Techniques: Exploratory Data Analysis (EDA), Hypothesis Testing, Data Wrangling, KPI Development

WORK EXPERIENCE

Business Analyst Intern | *Skiaverse Private Limited, Hyderabad, India* **Jan 2023 – Aug 2023**

- Utilized Python and SQL to analyze over **50,000 rows** of historical and real-time market data, uncovering behavioral patterns that supported the development of a region-specific Go-To-Market (GTM) strategy.
- Executed a data-driven GTM strategy based on analytical insights, achieving a **10% increase in customer engagement** within the first quarter across multiple product categories.
- Collaborated cross-functionally with a **5-member marketing team** to identify and define **3 high-conversion** customer segments, optimizing campaign targeting and improving marketing ROI by **12%**.
- Designed and deployed a Tableau dashboard that automated reporting across **7 core KPIs**, improving team efficiency by **15%** and saving over **8 hours per week** of manual effort.

PROJECTS

Home Credit Default Risk | *Python, Power BI, SQL, Machine Learning*

- Built a classification model using **Logistic Regression and Random Forest** to predict client loan repayment behavior, achieving **91% accuracy** and enabling informed risk-based lending strategies for financial institutions.
- Applied advanced **feature engineering techniques** with Python and SQL, transforming raw credit data into predictive features that enhanced the model's sensitivity to behavioral patterns.
- Tuned model performance using **10-fold cross-validation** and visualized training metrics with **TensorBoard**, identifying optimal hyperparameters and reducing overfitting across iterations.
- Ranked top 20% in the Kaggle competition with a **72.5% score**, supported by dynamic **Power BI dashboards** that communicated model insights to non-technical stakeholders.

Employee Retention Analytics Platform | *Power BI, Random Forest, Python*

- Designed and deployed a fully-interactive **Power BI dashboard** to visualize employee sentiment and engagement patterns from surveys, enabling leadership to identify emerging attrition risks in real time.
- Trained a **Random Forest classifier** on IBM HR datasets, integrating demographic and survey features to achieve **92% accuracy** and an **F1-score of 0.87** for predicting high-risk employees.
- Utilized **Pandas, NumPy, and Matplotlib** in Python to preprocess, analyze, and visualize trends across multiple HR metrics, helping uncover drivers behind employee dissatisfaction and turnover.
- Automated data refreshes and used **DAX expressions** in Power BI to provide continuously updated KPIs for HR managers, improving proactive decision-making and team-level planning.

Weather and Air Pollution Analysis for NY State | *Hadoop, Hive, Tableau, Python*

- Processed over **48 GB of unstructured Integrated Surface Data (ISD)** using **Hadoop MapReduce and Hive**, generating county-wise aggregates that revealed long-term weather and air quality patterns.
- Created **interactive Tableau dashboards** to visualize pollution metrics, precipitation trends, and temperature anomalies across **62 counties**, aiding government decision-makers and researchers.
- Developed and deployed a **Flask-based REST API** to expose filtered datasets for targeted retrieval, reducing average data latency by **40%** and supporting efficient downstream analysis.
- Integrated SQL-based filtering and Python scripts to automate aggregation pipelines, reducing manual processing time by **60%** and supporting **environmental impact studies** and data transparency across **62 NY counties**.