

Dheeraj Chillamcharla

Boston, MA | (469) 943-6727 | chillamcharladheeraj@gmail.com | www.linkedin.com/in/dheeraichillamcharla

A data enthusiast fuelled by caffeine; a traveller who loves code. Proven track record of excelling in challenging projects and thriving in fast-paced work environments. Strong team player with preference for hands on learning and enthusiasm for new opportunities.

EDUCATION

Northeastern University, Boston, MA

Master of Science, Data Science and Analytics

GPA: 3.96/4

Coursework: Database Management, Data Mining, Data Visualization, Machine Learning, MLOps

Manipal Institute of Technology, Manipal, India

Bachelor of Technology, Computer and Communication Engineering

Coursework: Data Mining and Predictive Analysis, Python Programming, Applied Mathematics, Data Structures

TECHNICAL SKILLS

Programming languages: Python, SQL, R

Data Techniques: Data Management, Data Analysis, Data Mining, Data Visualization, ETL pipeline, Business Intelligence

Libraries: Pandas, NumPy, SciPy, Scikit Learn, Matplotlib, Seaborn

Tools: Databricks, GitHub, MS Excel, Tableau, Power BI, RStudio, BigQuery, PyCharm, Git, GCP, Airflow

Certification: Google Data Analytics Professional Certificate

WORK EXPERIENCE

Peapod Digital Labs, Boston, MA

January 2023 – June 2023

Data Analyst Co-op [SQL, Python, Databricks]

- Architected and executed **high-performance data pipelines** to process massive datasets exceeding 10 billion rows. Leveraged big data technologies Spark and SQL in Databricks to effectively source and compile grocery Ad data for forecasting and analysis.
- Constructed highly accurate **predictive models** to predict ad effectiveness for three distinct product categories across two banners. Achieved an outstanding R2 score of 0.85, enabling persuasive storytelling and driving impactful business decisions.
- Developed **interactive dashboard** that visually showcases diverse customer metrics, like spending patterns across customer types, facilitating in-depth analysis of purchase behaviour and enabling strategic marketing and inventory management.
- **Optimized SQL queries** to streamline dashboard data retrieval, reducing the number of queries by over 60%. This optimization led to decreased cloud resource utilization and substantially improved run time.

Lucida Technologies Pvt. Ltd., India

January 2020 – July 2020

Data Analyst Intern [Python, Machine Learning Models, NLP, GCP]

- Developed a Python and OpenCV-based **Machine Learning model** for image analysis, enabling automatic table detection and saving. Semi-automated solution drastically reduced manual processing of paper-based bills, saving both time and space by 50%.
- Developed **back-end functionalities** for the multi-team project 'Edumonics', that extracts keyword phrases from text given Django API and implements Natural Language Processing techniques to identify "Subject-Verb-Object (SVO)" combinations.

PROJECTS

INCOME CLASSIFIER BASED ON DEMOGRAPHIC DATA [Python, Classification, Analysis]

- Conducted thorough **Exploratory Data Analysis (EDA)** to uncover patterns, inconsistencies, biases, outliers, and missing values.
- Executed meticulous **data cleaning** to address missing and duplicate data, handle data type inconsistencies, and eliminate irrelevant columns based on Variance Inflation Factor (VIF) values, enhancing the accuracy of machine learning models by 2-3%.
- Developed and optimized range of **machine learning algorithms** including KNN, Decision Trees, Random Forest, Logistic Regression, and Neural Networks. Achieved an optimal classification accuracy of 86.54% and an impressive F1 score of 92%.

DESIGN AND DEVELOPMENT OF SPORTS DATABASE ([Link](#)) [SQL, Python, MySQL Workbench, Microsoft SQL Server]

- Established a **database** on Microsoft SQL Server to manage player, staff, team, match, and country information. Database employed functions for automated calculations, triggers for data integrity, and table-level constraints for data consistency.
- Developed a detailed **Entity Relationship Diagram (ERD)** comprising 14 tables, including 1-to-1 and one-to-many relationships.
- **Automated data import** processes using Python scripts. These scripts dynamically generated Data Definition Language (DDL) statements based on Excel data, streamlining data integration into the database.

DATA ANALYTICS AND VISUALIZATION OF IPL PERFORMANCE ([Link](#)) [Tableau, Python, Data Visualization, Dashboard]

- Transformed **complex data into actionable insights**, effectively communicating the strengths and capabilities of players across 14 IPL seasons using four key metrics.
- Created executive level interactive **data visualizations** to show a team win percentage against other teams
- Developed insightful **Tableau dashboards** that provided overviews of team performance, player statistics, and team preferences.
- Ensured **data integrity** through meticulous data cleaning, eliminating missing, duplicate, and irrelevant data, standardizing data types, and harmonizing team information.