Dheeraj Chillamcharla

chillamcharladheeraj@gmail.com • (469) 943-6727
LinkedIn • Website • Boston, MA

Data Science Profile

Enthusiastic and results-oriented professional capable of harnessing advanced data analysis techniques and machine learning algorithms to extract actionable insights and deliver solutions that enhance decision-making, maximize productivity, and deliver measurable business growth. Equipped with a Master's Degree in Data Science & Analytics coupled with proven strengths in leading end-to-end data-driven projects, from problem formulation to solution deployment. Detail-oriented; adept at prioritizing tasks, managing timelines, and promptly addressing emergent issues. Proficient in driving collaboration across teams and distilling complex technical findings into clear recommendations for both technical and non-technical stakeholders.

Areas of Expertise

- Data & Statistical Analysis
- Business Intelligence & Planning
- Documentation & Reporting
- Machine Learning
- Data Governance & Visualization
- Data-Driven Decision Making
- Data Engineering & Modelling
- Model Evaluation & Validation
- Cross-Functional Collaboration

Technical Proficiencies

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

Education

Master of Science in Data Science & Analytics, GPA: 3.96
Northeastern University, Boston, MA, Dec 2023

Bachelor of Technology in Computer & Communication Engineering
Manipal Institute of Technology, Manipal, India, 2020

Career Experience

Peapod Digital Labs, Boston, MA   Jan 2023 – Jun 2023
Data Science Co-op

Built and executed high-performance data pipelines to process massive datasets exceeding 10B rows. Employed big data technologies, such as Spark and SQL in Databricks, to source and compile grocery Ad data for forecasting and analysis.

- Engineered highly accurate predictive models, predicting ad effectiveness for three distinct product categories across two banners.
- Attained R2 score of 0.85, enabling compelling storytelling and influencing impactful business decisions.
- Streamlined dashboard data retrieval by refining SQL queries, slashing their quantity by over 60%, resulting in a significant reduction in cloud resource consumption and markedly enhancing overall processing speed.
- Crafted an interactive dashboard that visually presenting diverse customer metrics, unveiling spending patterns across customer types, thereby facilitating in-depth analysis of purchase behavior and empowering strategic marketing/inventory management.

Lucida Technologies Pvt. Ltd., India   Jan 2020 – Jul 2020
Data Science Intern

Devised a Machine Learning model using Python and OpenCV for image analysis, facilitating automatic table detection and preservation. The semi-automated solution significantly cut down manual processing of paper-based bills, resulting in a 50% reduction in both time and space requirements.

- Spearheaded the development of back-end functionalities for the multi-team project 'Edumonics', implementing a Django API to extract keyword phrases from text and applying Natural Language Processing techniques to identify "Subject-Verb-Object (SVO)" combinations.

Key Projects

Real Estate Analytics Project: Leveraging Machine Learning for Market Insights

- Steered the implementation of a resilient MLOps (Machine Learning Operations) pipeline, integrating tools such as GitHub Actions, Docker, Apache Airflow, DVC (Data Version Control), AWS, MLflow, and Flask for streamlined development, deployment, and maintenance of machine learning models.
- Instituted and upheld a continuous integration and deployment (CI/CD) pipeline on AWS, leveraging GitHub Actions. Resulted in seamless model deployment and automated monitoring, guaranteeing optimal performance and accuracy in real-world scenarios.

Design and Development of Sports Database (Link)

- Established and optimized a Microsoft SQL Server database for comprehensive management of player, staff, team, match, and country information.
- Implemented functions for automated calculations, triggers for data integrity, and table-level constraints for enhanced data consistency.
- Formulated an intricate Entity Relationship Diagram (ERD) featuring 14 tables, delineating 1-to-1 and one-to-many relationships.
- Automated data import procedures using Python scripts, dynamically generating Data Definition Language (DDL) statements based on Excel data. This streamlined and expedited the data integration process into the database, enhancing overall efficiency.

Data Analytics and Visualization of IPL Performance (Link)

- Translated complex data into actionable insights, articulating the strengths and capabilities of players over 14 IPL seasons using four key metrics.
- Crafted executive-level interactive data visualizations, illustrating team win percentages against opponents.
- Engineered insightful Tableau dashboards delivering comprehensive overviews of team performance, player statistics, and team preferences.
- Maintained data integrity through rigorous data cleaning, eliminating missing, duplicate, and irrelevant data, standardizing data types, and harmonizing team information.

Income Classifier based on Demographic Data

- Performed in-depth Exploratory Data Analysis (EDA) to unveil patterns, inconsistencies, biases, outliers, and missing values.
- Executed precise data cleaning procedures to rectify missing and duplicate data, address data type inconsistencies, and eliminate irrelevant columns based on Variance Inflation Factor (VIF) values. Resulted in an enhanced accuracy of machine learning models by up to 3%.
- Developed and optimized a spectrum of machine learning algorithms, including KNN, Decision Trees, Random Forest, Logistic Regression, and Neural Networks, attaining an optimal classification accuracy of 86.54% and F1 score of 92%.

Licenses & Certifications

Google Data Analytics Professional Certificate