

DHEERAJ CHILLAMCHARLA

Boston, MA (Open to relocation) | (469)-943-6727 | chillamcharladheeraj@gmail.com | [LinkedIn](#)

SUMMARY

Data Scientist with 4+ years of experience in developing and deploying Machine Learning models, Predictive modelling, LLMs, GenAI. Proficient in Python, SQL, TensorFlow, and cloud platforms, with a focus on personalized learning, customer behaviour analysis, and improving business decision-making through data-driven insights. With a strong foundation in working with relational database management systems (RDBMS) like MySQL and SQL Server, ensuring data integrity and accessibility and in Tableau and Power BI to translate complex data into clear, engaging, and actionable visualizations.

TECHNICAL SKILLS

Languages: Python, SQL, Java, Scala

Data Science Technologies: Machine Learning, Generative AI, Feature Engineering, Supervised Learning, Unsupervised Learning, Decision Trees, Classification, SVM, Random Forests, Naive Bayes, KNN, K Means, CNN

Databases: Oracle, Snowflake, PostgreSQL, Cassandra

Python Packages: Pandas, NumPy, Matplotlib, SciPy, Scikit-Learn, SeaBorn, PyTorch, ggplot2, Plotly, keras

Data Analytics: Data Manipulation, Data Cleaning, Data Visualization, Exploratory Data Analysis, Data Analysis

Cloud Platforms: AWS, Azure (Databricks)

Tools: Tableau, Power BI, Excel, Visual Studio, GIT, Jupyter, Alteryx, Looker, Salesforce, Docker, Kubernetes

Methodologies: Agile/Scrum, Software Development Life Cycle (SDLC)

PROFESSIONAL EXPERIENCE

Abecedarian, Boston, USA | Data Scientist

Feb 2024 - Current

- Designed and developed an innovative academic enhancement product using **LangChain** and **Large Language Models (LLM)**, improving teaching workflows and automating content correction.
- Integrated **OpenAI GPT models** and fine-tuned **custom LLMs** with domain-specific academic datasets, enhancing personalized learning experiences and curriculum adaptation.
- Applied advanced **prompt engineering** techniques for dynamic content generation, such as question formulation, text summarization, and automated feedback.
- Implemented **GenAI workflows** utilizing **Python APIs** to enable real-time assessment, automatic grading, and tailored learning experiences for students.
- Leveraged **NLP evaluation metrics** (BLEU, ROUGE, perplexity) and **LangChain** to optimize conversational tutoring models and enhance educational content delivery.

Peapod Digital Labs, USA | Data Scientist

Jan 2023 - Jun 2023

- Collaborated in an Agile team to deliver high-performance data pipelines using **Spark** and **SQL** in **Databricks**, processing over 10 billion rows of grocery Ad data for forecasting and analysis.
- Engineered predictive models for ad effectiveness using advanced techniques, achieving an **R2 score of 0.85** for three product categories across two banners, helping shape business decisions.
- Optimized **SQL queries** to reduce dashboard data retrieval time by **60%**, significantly cutting cloud **resource consumption** and improving overall processing speed.
- Developed an interactive dashboard that visualized key customer metrics and spending patterns using a newly introduced feature on **purchase frequency**, enabling better analysis for marketing and inventory management.

HealthKart, Bengaluru, IN | Data Scientist

May 2019 - Jul 2021

- Developed predictive models** using **Random Forest**, **XGBoost**, and **Gradient Boosting**, improving product demand forecasting accuracy by **18%**, optimizing inventory management, and reducing stockouts by **15%** through the analysis of **transactional data**, **customer demographics**, and **product reviews**.
- Engineered new features such as **time-based shopping patterns**, **customer health profiles**, and **regional trends**, leading to the successful implementation of personalized health supplement recommendations and increasing average order value by **20%**.
- Built and deployed a **churn prediction model** using **Logistic Regression** and **SVM**, analyzing customer behaviour to reduce churn by **12%**, and integrated **NLP techniques** for **sentiment analysis** on unstructured data, improving customer satisfaction insights by **30%**.
- Leveraged **TensorFlow** and **Keras** to automate nutritional product image classification through deep learning, reducing manual processing time by **40%**, and integrated **real-time IoT device data** (fitness trackers) to enhance personalized health recommendations, increasing customer engagement by **22%**.
- Led **A/B testing** and **statistical analysis** using **Python** and **R** on marketing campaigns, identifying key drivers of customer retention, which resulted in a **15%** improvement, while designing **Tableau dashboards** for business leaders to track sales and user engagement metrics, improving decision-making by **40%**.

EDUCATION

Master of Science in Data Analytics Engineering, GPA: 3.96

Dec 2023

Northeastern University, Boston, MA

Bachelor of Technology in Computer & Communication Engineering

Jul 2020

Manipal Institute of Technology, Manipal, India