Deeja Chhabra
chhabradeeja@gmail.com | (817)-264-1595 | LinkedIn | GitHub

Summary:

Detail-oriented – Data Analyst with 4 years' experience with data. Passionate about creating dashboards to gain meaningful insights. Seeking to apply expertise in a team trying to change the meaning of Data.

Education:

- University of Texas, Arlington**, Masters (MS) in Computer Science**, **GPA: 4.0** **Dec 2023**
- Mody University, Bachelor of Engineering in Computer Science, GPA: 3.23 **May 2018**

Experience:

Machine Learning Engineer Intern: Green Watt — Jun 2023-Aug 2023

- Engineered a robust chat assistant, seamlessly transitioning between text and voice interactions by integrating Microsoft Azure, AudioCodes, **OpenAI's ChatGPT** (LLM) and Whisper technologies, contributing to **80%** rise in user engagement.
- Implemented **ML Ops** best practices for continuous model monitoring, deployment, and optimization.
- Presented actionable insights derived from data analysis to **senior leadership** with information crucial for decision-making.
- Spearheaded the communication of pertinent **insights** to technical and non-technical **audiences** through **Power BI dashboards** using data **storytelling** techniques. Tracked and visualized a **15%** YoY growth in conversions.
- Mentored a team of **5**, in implementing **Power business intelligence** visualization tools, dashboards and reports.

Data Analyst: Infosys — Jun 2018-Nov 2021

- Applied data processing, data cleaning, **descriptive analytics**, and **predictive modelling** on ride share application data, identifying **market trends**, **patterns**, **gaps** through **clustering**, **regression** that guided a product redesign, resulting in **20%** user growth.
- Developed **product recommendation engine** for customers based on association rule mining, and customer segmentation.
- Utilized **DAX** functions to create relationships, calculated columns, and measures for effective data representation.
- Led **database management processes** including integration, database design, data processing and data modelling.
- Achieved a **20%** improvement in data processing by optimizing **SAS** programs for data manipulation, analysis, and reporting.
- Conducted **data analytics** to solve complex business problems using **SQL** on large scale datasets, to **extract** actionable insights.
- Integrated **Tableau data-driven alerts** and **notifications**, enabling the tracking of crucial **KPIs** and cutting response times by **25%**.
- Employed **statistical modelling** in retail to enhance experiences, predict trends, and drive a **30%** increase in **sales**.
- Developed and customized **real-time Power BI dashboards** for **finance** and **marketing** departments**,** identifying crucial **KPIs.**
- Thorough **documentation** of machine learning techniques, covering aspects such as Data, code, and evaluation metrics
- Expertise in **SQL** and languages like **R** and **Python,** with a focus on troubleshooting and improving data processes.
- Fired **SQL** queries for **ad-hoc analysis** on **large scale datasets**, to extract **meaningful insights,** showcasing analytical skills.
- Created **Pivot Tables** and **Charts**, **VLOOKUP's**, formulas, etc. in **microsoft excel** based on business users' need.

Academic Projects:

Price Estimation with RFC and XGBoost | TensorFlow, NumPy, Matplotlib

- Implemented Random Forest and XGBoost Regressor models for fare estimation, achieving an impressive R2 score of over **0.85**.
- Transformed categorical features into numerical ones using label encoding, leading to a **25%** reduction in model training time.
- Leveraged exploratory data analysis (**EDA**) techniques to uncover insights, leading to a **30%** reduction in **feature dimensionality.**

Forecast Monthly Average Temperature in DFW | TensorFlow, Python, NumPy, Pandas

- Employed **ARIMA** and **FBProphet** machine learning Model for accurate temperature **forecasting**, involving **feature engineering**.
- Developed informative **data visualizations** using Power BI to illustrate historical trends and **predict future** temperature values.
- Achieved significant accuracy with max monthly average temperatures projected at **85.74**(ARIMA), **90.61**(FBProphet) for 2024.

Early Detection of 3D printing issues | scikit-learn, pytorch, computer vision, SciPy

- Employed **data augmentation** on images and visually inspected the results by plotting them to validate their efficacy.
- Utilized multiple CNN models, tuning them, and including leveraging **transfer learning** with models such as **VGG16** to exploit the dataset's scale and achieve notable performance improvements with a success rate of **99%** in **anomaly detection**.
- Extensive research on ML, Deep Learning to be able to apply **image segmentation** to improve model accuracy by at least **10%**.

Skills & Certifications:

- Languages: Python, R programming, DAX, SQL, Spark
- Data Science: Analysis, Diagnosis, Data Visualization, Statistics
- Databases: SQL Server and MySQL, PostgreSQL, Oracle DB
- Cloud Frameworks: MS Azure, Amazon Web Services (AWS) Sage maker, Snowflake, GCP
- Data Tools: Alteryx, Tableau, Microsoft Power BI, MS Excel, power point, Databricks
- Certifications: Microsoft Certified Data Fundamentals, Microsoft Certified Azure Fundamentals, Microsoft Certified AI Fundamentals, Microsoft Technology Associate (MTA), DeepLearning.AI TensorFlow Developer specialization