Dhairya Jayesh Chheda

Ready to Relocate | 📞 +1 930-333-5591 | ✉ dhairyajayeshchheda@gmail.com | in linkedin.com/chheda-dhairya | ○ GitHub | ⊕ Portfolio

EDUCATION

Master of Science (M.Sc.) in Data Science — August 2023 – May 2025
Indiana University – Bloomington, Indiana, USA — *GPA: 3.9/4.0*

Bachelor of Engineering (B.Eng.) in Information Technology — August 2017 – May 2021
University of Mumbai – Mumbai, Maharashtra, India — *GPA: 9.62/10.0*

TECHNICAL SKILLS

Languages: Python, R, SQL, C/C++
Databases: Data Warehousing, Database Querying, MS SQL, MySQL, PostgreSQL, NoSQL, Salesforce, Snowflake, Databricks
Frameworks: Numpy, Pandas, Matplotlib, RShiny, ggplot2, Scikit-learn, NLTK, Spacy, Tensorflow, Keras, PyTorch, Streamlit, API
Data Science: Statistics (Descriptive & Inferential), Hypothesis Testing, A/B Testing, Data Processing, EDA, Data Visualization, Business Intelligence Reports and Dashboards, Key Performance Indicators (KPI), Feature Engineering, Dimensionality Reduction, Data Mining, Machine Learning (Supervised & Unsupervised), Regression, Classification, Predictive Modeling, Recommendation Systems
Developer Tools: Stata, Tableau, Microsoft Power BI, Alteryx, ETL (Extract Transform Load), Google Cloud Platform (GCP), Amazon Web Services (AWS), GitHub Actions, Jenkins, JIRA, Microsoft Excel (Pivot Tables & VLOOKUP), Microsoft PowerPoint

PROFESSIONAL EXPERIENCE

Machine Learning Engineer Intern — September 2024 – December 2024
*Jiseki: an early stage AI **startup*** — *San Mateo, CA, USA*
- Optimized **ETL pipelines using AWS** Glue, Amazon Redshift, and Amazon S3 to integrate and harmonize data from various sources into a **centralized data warehouse**, **enhancing data pipeline scalability and efficiency by 50%**.

Data Scientist Intern — May 2024 – July 2024
Indiana University — *Bloomington, IN, USA*
- **Extracted and gathered** 50+ reports and refined **metric computations by monitoring** insights, streamlining data export process to Microsoft Excel and relational databases, achieving **50% improvement in data processing efficiency**.
- Conducted **data modeling** and **exploratory data analysis (EDA)** on **1M+** data points to create **dynamic, interactive and meaningful Tableau dashboards**, enabling data-driven reporting and presenting **actionable KPIs** to stakeholders.
- **Communicated** findings from analyses to consultants through **visuals, data tables and models**, and assisted with the development of final reports and presentations to be shared with stakeholders, fostering effective **teamwork** and project success.

Software Engineer — June 2021 – July 2023
LTIMindtree — *Mumbai, Maharashtra, India*
- **Spearheaded** a team in achieving **75% acceleration in project completion** by utilizing the Six Sigma approach and **SCRUM project management**, and following **SDLC** and **Agile** methodologies, minimizing project duration from 40 days to 10 days.
- Streamlined **CRUD operations** with Python and SQL **automation**, saving over 4 days of manual effort.
- Constructed Python scripts for **data migration and cleaning**, particularly for **Teradata** and **PeopleSoft DB2** transfers.

PROJECTS

SalesOptimizer: E-Commerce Sales Analysis ⬈ — *Excel, Python, Statistical Analysis, EDA, SQL, Microsoft Power BI*
- Designed a real-time **sales dashboard to track KPIs** and monitor data accuracy, increasing **operational efficiency by 15%.**
- Performed **financial planning and strategic analysis** and conducted time series analysis, **identifying trends, key growth factors and forecasting** sales for the next 15 days, contributing to a **30% increase in annual sales growth**.

EduPredict: Student Performance Forecaster ⬈ — *Decision Tree, Random Forest, XGBoost, Ensemble Learning*
- Enhanced model performance by applying **ensemble learning** techniques, **boosting F-1 score to 0.86**.
- Implemented a **REST API** to enable seamless **integration** with other applications and **deployed** the model via Streamlit, **optimizing performance, scalability, and reliability**, and increasing user engagement.

Airbnb Insights: Booking Analysis & Recommendations ⬈ — *Data Warehousing, SQL, BigQuery, Looker, PySpark*
- Utilized PySpark for **distributed data processing** and transformation, **reducing data processing overhead by 40%**.
- **Visualized data** on BigQuery and Looker to **uncover trends and insights**, and **recommend strategic business-decisions**.

PowerForecast: Electricity Demand Predictor ⬈ — *A/B Testing, Analysis of Variance (ANOVA), ARIMA/ARMAX*
- Applied **advanced statistical techniques** and developed **predictive forecasting models** to **forecast short-term and long-term** electricity usage patterns, aiding in efficient resource allocation.
- **Optimized model performance** by incorporating external factors such as seasonal trends to enhance forecast accuracy.

ACHIEVEMENTS / EXTRA-CURRICULARS

2024 – President, Data Science Club: Responsible for mentoring students, inviting industry experts and organizing events.
2023 – Pat on the Back: Performance Award from the manager for my exemplary contributions to deliver the project 'MFGOG'.
2018 – Academic Merit: University Honors Award for the best academic performance among all freshmen.