SHRADDHA THAKKAR

Boston, MA | (857)-891-7223 | shraddha.thakkar0796@gmail.com | linkedin.com/in/shraddha-thakkar8/ | shraddhathakkar.me/

A graduate student with a specialization in Analytics and data-driven mindset. Passionate to learn and build strategies for challenging problems implementing Data Analytics, Business Intelligence, Data Engineering, and Machine Learning concepts

EDUCATION:

Northeastern University, Boston, MA — September 2018-April 2020
Master's in Informatics (Concentration in Analytics)

Gujarat Technological University, Ahmedabad, India — July 2014-May 2018
Bachelor of Engineering in Information Technology

TECHNICAL SKILLS:

Programming Languages:	Python (Numpy, Pandas, Matplotlib, Sklearn, Seaborn), R (ggplot, dplyr, plotly), SQL, HTML, CSS
Databases:	MySQL, MS SQL Server, Oracle, PostgreSQL
Data Integration/ETL:	Talend, SQL Server Integration Services (SSIS), Alteryx, ER/Studio
Data Visualization:	Tableau, Power BI, Microsoft Excel (Pivot Tables, Pivot Charts, V lookups)
Cloud:	Amazon Web Services (IAM, S3, DynamoDB, Lambda)
Machine Learning:	Regression, Classification, Decision Trees, Random Forest, KNN, K-means clustering, PCA

WORK EXPERIENCE:

Auxi Inc, Cambridge, MA **| Machine Learning Engineer Intern** — May 2020-Present

- Working on AI productivity suite, an AI bot that transforms handwritten sketches into fully formatted PowerPoint or Google Slides presentations and beautifying slides with just a press of a button
- Using ASP.NET C# with Microsoft PowerPoint APIs and Open XML SDK to maintain formats and Google Vision API to recognize text, flowcharts, and frameworks thereby achieving an accuracy of around 90%

Northeastern University, **HRM Department,** Boston, MA **| Business Analyst** — January 2020-March 2020

- Communicated with the stakeholders and tech teams to understand requirements and build iterative solutions
- Conducted users' interviews, gathered and analyzed their data to create an empathy map and personas through research on behavior pattern of the focus group
- Built metrics and business cases to improve customer experience by creating 10+ prototypes, came up with the solution to increase accessibility and usability of the webpage; resulting in a 10% decrease in the bounce rate

Chainaim Pharmaceuticals, **Northeastern University,** Boston, MA **| Data Analyst** — April 2019-June 2019

- Studied the data to segregate the functionalities and tackling problems of reading/cleaning messy data
- Analyzed pharmaceutical data operations using SQL, Microsoft Excel (Pivot Tables, V lookups), and Python to study criteria for finding drugs details, GS1 data standards, and nomenclature for GTN-14
- Identified shortcomings and offered recommendations to help drive client business development resulting in eliminating manual efforts of more than 15 hours/week

PROJECTS:

github.com/shraddha0796

Retail Data Warehouse (Talend, SQL, Tableau, Power BI, ER/Studio, Alteryx)

- Designed a data warehouse (Snowflake schema), pipelined and integrated large dataset of 40 million records from multiple sources (CSV, MySQL, MS SQL Server, Oracle) using ETL techniques on Talend
- Used OLAP, Facts and Dimensions tables and executed SCD to maintain historical records and Error Handling
- Devised various performance tuning techniques, thereby resulting in a total integration time of 22 minutes
- Built interactive dashboards to analyze retail sales and business insight using Tableau and Power BI

Predicting Readmission of a Diabetic Patient (Python, Machine Learning Algorithms)

- Performed Data Cleaning, Data Exploration, and Preparation, Feature Engineering, and Exploratory Data Analysis (EDA) to predict and develop statistical models (Logistic Regression, Decision Trees, Random Forest)
- Implemented modeling and manipulated data to improve accuracy from 53.78% to approx. 90%

Boston- Is it safe? (Data Analysis, Data Visualization, RStudio-ggplot, plotly, dplyr, tidyverse)

- Loaded, cleaned, and converted data into actionable insights of Boston crimes (2015-2018)
- Conducted exploratory data analysis (EDA), implemented hypothesis testing and predicted crime rate for 2019 using time-series analysis (ARIMA model), stated results with 99% confidence interval

AIRBNB Booking and Rental Database (MySQL Workbench, Microsoft SQL Server)

- Modeled Entity Relationship for AIRBNB, Normalized database to 3NF and executed features like Views, Stored Procedures, Functions, Triggers, User Access and Privileges
- Implemented updates on database and encryption of critical data such as passwords and payment details using MD5

Serverless Cloud Application (Python, AWS- s3, Lambda, DynamoDB)

- Configured Lambda function to be triggered on s3 events and publish logs to CloudWatch
- Executed python lambda handler to store all s3 events metadata in DynamoDB