ANUPRITA DHAMALE

Boston, MA 02120 | (857)-318-6365 | dhamale.a@northeastern.edu | linkedin.com/in/dhamaleanuprita

EDUCATION

Northeastern University, Boston, MA   **Expected graduation - Dec 2023**
Master of Science in Information Systems
Relevant Courses: Designing Advanced Data Architectures for Business Intelligence, Application Engineering and Development, Database Management and Database Design, Business Analysis and Information Engineering

All India Shri Shivaji Memorial Society, College of Engineering (Pune University), Pune   Jul 2016 - May 2020
Bachelor's in Computer Engineering
Relevant Courses: Discrete Mathematics, Data Structures & Algorithms, Database Management Systems, Data Warehousing & Mining, Machine Learning, Information Systems and Engineering Economics, Cloud Computing, Project Management

SKILLS

Programming Languages:	Python (Scikit, NumPy, Seaborn, Matplotlib, Nltk, Pandas), R, HTML, C++, Java
Database:	MySQL server, Microsoft SSMS, DBeaver, MongoDB, PostgreSQL, Google BigQuery, Oracle SQL developer, AWS Redshift, AWS S3
Big Data:	Hadoop Ecosystem, HDFS, Spark, Impala, Hive
Tools:	Alteryx, Talend, ER Studio, Lucidchart, RStudio, Jypter Notebook, Microsoft Office (Excel, Word, PowerPoint), ServiceNow, Service First, Tableau, PowerBI, Microsoft Visio, Informatica
ML Algorithms:	(Linear, Logistics, Lasso & Ridge) Regression, KNN, Naïve Bayes, K-Means clustering, Decision Tree, Random Forest, SVM, Gradient Boosting, Time Series Analysis

PROFESSIONAL EXPERIENCE

iASYS Technology Solutions, Pune   Aug 2020 - Mar 2021
Data Analyst Intern
- Translated business **KPI** (gross profit, inventory management) and proposed cost-effective solutions using **Tableau** to automate day-to-day operations to optimize human and capital resources of **11 projects** which improved efficiency by **22%**
- Implemented SQL database schemas and scripted **230+** SQL queries using **Oracle SQL Developer** for various clients by collaborating with application integration team to derive statistics related to the overall assets on **AWS redshift**
- Migrated over **6GB** of vehicle testing data, performed **ETL** operation and onboard data into **Tableau** to develop **20+ dashboards** and visualize data to bolster key business decisions in accordance with KPIs from client industries
- Accelerated publication of vehicle performance metrics and improved **asset tagging** by up to **80%**, by developing, analyzing, and reporting trends in dashboards and ad-hoc reports which further improved the accuracy of operational decision
- Developed test cases and assisted project manager to create feasible schedules using **MS Project** for communication and cooperation with other teams using **Agile methodology** to discuss and bring about optimal solutions

ACADEMIC PROJECTS & LEADERSHIP EXPERIENCE

Business Intelligence and Data Warehousing - AdventureWorks2017 (Power BI, Tableau, Talend)   Nov 2021- Dec 2021
- Designed a Data Warehouse from four different geographic locations and various data sources
- Implemented an access frontend allowing connections to machines facilitating seamless transfer of data
- Developed strategic BI Dashboards in **Power BI** and **Tableau** based on product store sales, online sales and rejects analysis
- Pipelined large scale data of a business from multiple database sources and integrated into SCD using **Talend**

IMDB and MovieLens data analysis (MySQL, Alteryx, Tableau, Talend)   Oct 2021 - Nov 2021
- Utilized **Alteryx** to import data for a multi-star schema model to database and accomplished further analysis on the data
- Imported 1.5M data into **SQL server**, executed preprocessing of data (EDA), and load staging data employing **Talend**
- Visualized data for top 1000 movies (by revenue) in **Tableau** to support key business decisions and audited reports
- Created dimensional model by building ETL processes through Talend to integrate facts and dimensions of current dataset

NYC bike sharing (MySQL, Tableau, Oracle ADW, Azure SQL)   Dec 2019 - Jun 2020
- Used Tableau data prep to import denormalized data to **MySQL** database and accomplished further analysis on data
- Researched 10 years of dataset and performed data cleaning in **Tableau** data prep and acquired error-free data
- Implemented data normalization for the last 10 years of data and eliminated discrepancies employing **Oracle ADW**
- Executed **SQL queries** in Oracle ADW to analyze bike trips, stations, and system data to identify bike usage trends

Twitter sentimental Analysis (Python, Excel)   Dec 2019 – Jun 2020
- Examined, scraped & classified raw dataset to estimate the toxicity of a tweet using **Nlkt** for stemming and tokenization
- Explored and visualized data in the cleaned tweets using **word cloud** plot to gain insights
- Analyzed preprocessed data using Bag of Words, TF-IDF, and Word Embeddings techniques to convert into features-sets
- Applying machine learning techniques including **logistic regression**, **decision tree**, **support vector**, and **random forest** to this dataset achieving an F1 score of **0.7** for different models and feature sets

CERTIFICATIONS

- Big data analytics with Hadoop and Apache sparks   Dec 2021
- The complete Python Bootcamp   Sep 2020
- The complete SQL bootcamp course   May 2020