YUKTHA LAKSHMI A B

(315) 952-7379 · yukthalakshmiab@gmail.com · LinkedIn · GitHub . Portfolio

EDUCATION

Syracuse University, *Syracuse, NY* — May 2024
Degree: Master of Science in Applied Data Science — GPA: 3.8 / 4.0
Coursework: Data Warehouse, Big Data Analytics, Applied Machine Learning, Cloud Management, Quant Analytics

Dayananda Sagar College of Engineering, *Bengaluru, KA* — May 2020
Degree: Bachelor of Technology in Information Science and Engineering — GPA: 9.3 / 10

WORK EXPERIENCE

ProHealth Group Inc., Birmingham, AL — June 2023– Present
Data Analyst Intern
- Pre-process raw data from SQL Server, utilize dbt to create data model & apply transformations based on business logic
- Visualize aggregated data in PowerBI, apply business formulas using DAX and created 6 reports live in production
- Identify key metrics that affect the operation and create measures using DAX, building KPI's for finance dashboard
- Monitor trend changes in finance dashboard and trigger messages to business users as alerts
  Advanced Skills: SQL Server Management Studio, PowerBI , Power Query, DBT

iConsult Collaborative, Syracuse, NY — Aug 2022 – May 2023
Data Engineer Intern
- Built pipelines to integrate data from shared folders, performed data wrangling using DBT leveraging python and SQL
- Constructed a data model using DBT, utilized macros and user defined functions to write business formulas
- Schedule and maintain the deployment using GitHub actions to load data in snowflake
- Designed dashboards in tableau to visualize the university student data based on requirements gathering
  Advanced Skills: Python, Snowflake, Tableau, ELT, Databricks

Larsen & Toubro Infotech, Bengaluru, KA — Aug 2020 – Aug 2022
Data Engineer
- Developed 10+ reports using both Power BI and Spotfire for analyzing the oil & gas production in Chevron
- Achieved an efficiency rate of less than 20% wastage in refiners through data-driven insights from these reports
- Implemented an application using PowerApps by connecting to CDS pipeline and APIs by testing in Postman tool
- Migrated database from Teradata to SQL Server, transformed data using SQL and loaded to Spotfire for visualizing
  Advanced Skills: Power BI, Spotfire, Teradata, Power Apps, SSMS, Scikit, ETL

PROJECTS

Music Recommendation System – GitHub **-** (*Python, Machine Learning, Principal component analysis, Seaborn*)
- Cleaned raw data, convert attributes from numerical to categorical using one-hot encoding method
- Applied association rules (EDA) to understand dependency of variables, conducted dimensionality reduction
- Obtained optimized recommendation system by K-means, DBSCAN, Agglomerative Clustering

Road Safety Analysis (*PySpark, Linear Regression, Random Forest, SVM, Clustering, Jupyter notebook*)
- Enhanced the clustering model for identifying collision patterns by **96.57%** using Sum of Squared Error(SSE)
- Improved silhouette scores, utilizing Kmeans, PCA and Elbow method with **23,000+** accidents with PySpark
- Detected 350 anomalies, measured by squared Euclidean distance with 0.413 threshold, build clustering model of anomaly detection employing PySpark

Bitcoin Trend Analysis – GitHub - (*Python, Forecasting, ARIMA, SARIMA, LSTM*)
- Perform data exploration to understand correlation in the data set , leveraged one-hot encoding method
- **Forecast** weighted price using linear regression, gradient boost trees and performed time series analysis
- Predict weighted price for the **next 10 days** based on root mean square error values

TECH SKILLS

Data Science:	NumPy, Pandas, Scikit-learn, Matplotlib, TensorFlow, Keras, R Programming
Data Engineering:	MySQL, Teradata, Snowflake, Azure Data Factory, Microsoft Azure, Visual Studio, SSMS, RDBMS
Other Tools:	Atlassian, Advanced Microsoft Excel, Power Apps, PowerBI, Power Flow, Spotfire, Tableau

CERTIFICATIONS

- Snowflake – Hands on Official- Data Warehouse — Sep 2023
- Databricks Lakehouse Fundamentals — Nov 2023
- Agile Developer — Jan 2022