KSHAMA ADITI LETHAKULA

+(940)-299-1591| lethakula.k@northeastern.edu | LinkedIn

EDUCATION

Master of Science in Information Systems   *May 2024*

Northeastern University, Boston, MA

Coursework: Data Science Engineering Methods and Tools, Application Engineering & Development, Data Management and Database Design, Web Design and User Experience Engineering

Bachelor of Technology in Electronics and Communication Engineering   *August 2021*

G. Narayanamma Institute of Technology and Science, Hyderabad, India

Coursework: data structures, object-oriented design, operating systems, DBMS

TECHNICAL SKILLS

Languages: Python, Java, and C programming.

Big Data Tools: Spark, Hadoop, SQL, ETL, Scala, Kafka, Linux, Shell Scripting

Cloud: Amazon EC2, S3, Oracle, Google App Engine

Databases: MySQL, MS SQL Server, PostgreSQL, Oracle DB, MongoDB, Hive

Web Technologies: Spring Boot, Django, REST APIs, JavaScript, React JS, GitHub

Visualization/Automation Tools: Power BI, Tableau, Jenkins, Selenium

PROFESSIONAL EXPERIENCE

Analyst, *Colruyt Group, India*   *Jul 21 - Aug 22*

- Designed, implemented, and maintained **ETL** pipelines for 2 projects involving fraudulent pharmacies and claims data.
- Processed and transformed the **6-7 Terabytes** of streaming data from Kafka & loading it into **Hive** and **MS SQL DB**.
- Developed code using **PySpark**, and **SQL**. Maintained the workflow through the **Airflow** scheduler and shell scripting.
- Automating the manual auditing reduced the time spent generating audit reports from 3-4 days to **2 hours**.
- Rendered ombud assistance by handling requests using ServiceNow ITSM to provide developers with better database design and analysis of few applications by optimizing database objects, queries.
- Recognized with SPOT award from Colruyt for effective programming and positive impact.

Intern, *Colruyt Group, India*   *Feb 21 – June 21*

- Designed a new physical database model using SAP's power designer and developed those models as schemas in Oracle DB in Cloud.
- Replicated and decommissioned a few databases as a part of the migration of a few services from Object Star to other DB environments.

ACADEMIC PROJECTS

Airbnb Listings Analysis   *Oct 2022 - Dec 2022*

- Web scrapped data to extract information from over 76,000 listings leveraging BeautifulSoup in python.
- Assessed listings to anticipate different factors affecting price by implementing Random Forest, SVM, and XGBoost.
- Applied tokenization to get listings' top 10 amenities; used ggplot2, maps, ggthemes in R to visualize the data and convey results.

Spatial Hotspot Analysis using Apache Spark, *Big Data project*   *Sep 2022 - Nov 2022*

- Developed a distributed application using Spark APIs and Hadoop File System to perform range and distance queries.
- Applied spatial statistics to spatial-temporal data to identify hotspots using Spark SQL.

E-commerce Database Design   *Aug 2020 - Sep 2020*

- Led a team of 3 to design an e-commerce database schema using UML and ER diagrams; further optimized it by 20%.
- Built a database with SQL and executed DML, DDL operations, triggers, and stored procedures; further, crafted back-end in python which interacts with database and integrated with front-end.

OTHERS

Certification: ServiceNow ITSM Fundamentals