DHANYASHREE A

DATA ENGINEER

Phone: 7624813971 Email:dhanyashreea86@gmail.com

Shape

**OBJECTIVES**

With over 2 years of experience in the data engineering field, expertise includes utilizing Python and PySpark on Databricks for advanced data processing and analysis. Specialization in designing and optimizing data workflows and transformations with PySpark, managing and querying data in MySQL, and developing efficient data pipelines and integration solutions. Proficient in leveraging Databricks' scalable infrastructure to enhance ETL processes and ensure high performance in large-scale data environments. Experience also includes implementing data quality measures and collaborating with cross-functional teams to deliver actionable insights and meet business needs.

Shape

**EXPERIENCE SUMMARY**

* Expertise in designing, building, and maintaining scalable data pipelines using Python and PySpark on Databricks for efficient data processing and integration.
* Proficient in managing and querying relational databases like MySQL, with solid knowledge of SQL for data manipulation and analysis.
* Participated in Scrum Meeting to discuss the Module
* Experienced in developing ETL workflows to transform, clean, and load data from various sources into target systems.
* Good hands on with AWS, Amazon S3 bucket.
* Excellent verbal and written communication, interpersonal skills, and strong abilities in decision-making, problem-solving, and customer service.

Shape**EDUCATION**

Vivekananda Collage of Engineering and technology,puttur,MCA

• GPA: 8.83

• Coursework: Computer Application, Machine learning in python, data Structure

Shape

**PROJECT SUMMARY**

**Data Integration and Analysis for Medical Diagnostics**

* **Objective:** Enhance the efficiency and accuracy of medical diagnostic processes by integrating and analyzing disparate data sources using modern data engineering tools.
* **Tools Used**:Python, PySpark, SQL, AWSS3
* **Description**: AWS S3 buckets were utilized to securely store diverse medical data, establishing a scalable storage architecture. PySpark workflows were developed on Databricks to automate ETL processes, transforming raw data into structured formats efficiently. On Databricks, SQL queries were crafted and optimized for in-depth analysis of medical records and diagnostic results. Python scripts automated data integration and synchronization tasks. Performance optimizations for PySpark jobs and SQL queries on Databricks enhanced overall processing efficiency. Collaboration with healthcare professionals ensured that the solutions met their needs and provided ongoing support.

**LICENSE PLATE BASED PENALTY ENFORCEMENT FOR NONHELMET RIDERS**

* **Objective:** Implement an automated system to identify and penalize motorcycle riders who are not wearing helmets, leveraging license plate recognition technology to enforce traffic safety regulations.
* **Tools Used**:Python,Django,YOLOV5
* **Description**: The project aims to enhance road safety by developing a system that automatically detects non-helmeted motorcycle riders using license plate recognition. The system captures images of riders, processes them to identify those not wearing helmets, and issues penalties based on license plate information. The goal is to improve compliance with helmet laws and reduce motorcycle-related injuries.

Shape

**TECHNICAL SKILLS**

Languages: C++, C, Python,Pyspark,AWS,Databricks

Database: MySQL,NoSQLShape

**PERSONAL PROFILE**

Data of Birth: 01-05-2000

Nationality: India

Language: English,Kannada,Hindi

Address: Puttur,manglore,DK

**DECLARATION:**

I,Dhanyashree A, hereby declare that the information contained herein is true and correct to the best of my knowledge and belief.

Date: / / 2024

Place: Bangalore