Sandeep Chanda

Kansas City, MO | P: +1 (816) 304-3539 | sandeepchanda24@gmail.com | https://www.linkedin.com/in/sandeep-chanda-de8/

WORK EXPERIENCE

Cognizant. (client: Sanofi) — Hyderabad, India
MDM Developer — Jun 2021 – Nov 2022
- Implemented advanced data manipulation techniques, including **match and merge, filtering, data validation and enrichment**
- Data manipulation resulting in a 21% reduction in data errors and improved data accuracy and the ultimate creation of a golden copy
- Enhanced data processing efficiency by 34% through **Semarchy xDM(MDM Tool)** and **IICS** (**ETL** tool)
- Oracle database for querying, storing and validating data between layers like OD, SD, MD, GD

Cognizant — Coimbatore, India
MDM and ETL Intern — Feb 2021 – Jun 2021
- Excelled in extracting data from multiple sources, achieving a 19% acceleration in loading into the databases.
- Applied robust filtering and transformations in **PowerCenter**, resulting in a 17% enhancement in data quality.
- Streamlined the workflow to Informatica MDM, contributing to the reduction in processing time through efficient transformations and integration techniques.

EDUCATION

UNIVERSITY OF MISSOURI KANSAS CITY — Kansas City, MO
Masters — Expected Dec 2024
Major in Computer Science; Minors in Mathematics and Psychology
Cumulative GPA: 3.52/4.0
Relevant Coursework: Big Data, Data Science, Statistical Learning, Cloud Computing, Advance Operating Systems

KARUNYA INSTITUTE Of TECHNOLOGY And SCIENCES — Coimbatore, India
Bachelor of electronics and communication Engineering — Jul 2017 - May 2021
Cumulative GPA: 3.33/4.0

PROJECTS

Drowsiness and Yawning Detection using Facial Landmarks: — Nov 2024
- Developed a real-time drowsiness and yawning detection system using Python, OpenCV, and dlib, leveraging facial analysis with Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) to detect driver fatigue.
- Implemented face detection using Histogram of Oriented Gradients (HOG) and alarm triggers based on threshold criteria to enhance driver safety through continuous monitoring.
- Shortlisted in first 5 teams out of 100 teams for hack-A-Roo (Hackathon)

Analysis of Relation Between Food Diets, Health and COVID-19 using Power BI and Azure: — Jul 2023
- In This project, we collected data from **Kaggle** and uploaded it to the **Azure SQL database**
- created a **Power BI dashboard** by pulling data from the **Azure database server** and uploaded this dashboard to the web

Goodreads ETL Pipeline using PySpark: — Apr 2023
- The main aim of the project is to capture data in real-time from good reads API using a Python wrapper to store collected data in a local disk and then move to the landing bucket of **AWS S3**
- All the ETL jobs are written **PySpark** and Once ETL job execution is done the **Airflow** DAG runs the data quality check

TECHNICAL SKILLS

Programming Languages: Python, SQL, SemQL,
Tools: Tableau, Power BI, PySpark, Semarchy xDM, Informatica MDM, Informatica Power Center, Azure Databricks, MS Excel, Hadoop, Apache Sqoop, Hive, Apache Spark

Cloud-Based Technologies: Azure**,** GCP, AWS
Certifications & Training: Completed certification course on Data Warehouse –course by Nikolai Schuler