PRAVIN YADAV

☎ +91-8108937558 ✉ Pravin Yadav 🔗 Pravin Yadav 🐙 Pravin Yadav </> Pravin Yadav Ⓗ Pravin Yadav

PROFILE

A data engineer with hands on expertise in big data tools, cloud technologies, analytics and software development techniques.

EDUCATION

Bachelor of Engineering in Electronics And Telecommunication (2017 - 2021) %
University of Mumbai 80.21

SKILLS

Coding Langauges	SQL, Python, Java, NoSQL, DAX, TeX
Frameworks	Spark (Pyspark & Snowpark), Hadoop Eco-System, Pandas, Collections, Matplotlib, Kafka, SQLalchemy, Spring Boot
Tools	Databricks, Postgres, Power BI, Azure, Jupyter Notebook, VS-Code, Git & Git-Hub, Atlassian Project Management Tools, Linux
Technical Skills	Data Engineering, OOP, Data Modeling, Data Analysis, Data Management and Governance, Rest-API, Design Patterns, Test Driven Development
Formats	Delta, Hudi, JSON, CSV, Parquet, XML
Soft Skills	Time management, Listening, Problem Solving, Chat-GPT/Googling, Document Writing

EXPERIENCE

LTIMindtree (Data Engineer) June 2021 - Present

- Worked on multiple end to end ETL pipelines (Apache Spark/ DLT) and dashboarding (Power BI) projects having TB/PB of data for our client (One of the World's Largest Cosmetic Company) including USA, LATAM and CANADA which provided a broader prospective of data driven strategies across regions.
- Worked on multiple Data Design Patterns & Architectures including Dimensional Modelling, Medallion and CDC.
- Architected and implemented an Batch CDC Framework which reduced the execution time for backfilling data by 70%.
- Implemented streaming use case on top of data which were ingested from event hub to ADLS Gen2 using Audoloader and DLT
- Part of an agile project following CI/CD process which provided immense learning by handling and implementing end to end design changes in data model by resolving and enhancing multiple SR's and CR's.
- Optimized multiple ETL Jobs using spark optimizations, reducing repeatable table scan and using table logging technique which removed external dependencies, that led to reduction in execution time by 30-50% on a daily refresh schedule.
- Worked on some major data quality issues which required quicker fixes in heavy load time.
- Got exceptional in communication skills, being provided multiple chance to interact with clients for explaining business level impact analysis, modeling and scalability.

Achievements   Received more than 5 awards or appreciation certificate from managers and client for contributions, optimization and delivery, 100% in Internal SQL Challenge, Winner in Automation Challenge.

Crio.Do (Student Developer Intern) Oct 2021 - Oct 2022 🔗 Projects

- Worked on multiple projects which provided an exposure to industry level backend systems.
- Followed test driven development and scientific debugging procedure to write clean and bug free code.
- Understood the procedure of creating LLD in backend systems using different design patterns and object oriented programming resulting in scalable application design.
- Enhanced the knowledge of Collections, Data structures & Algorithms and SQL, NoSQL Databases.

Achievements   Received DSA champ notebook, 70% in Geektrust Backend Challenge

CERTIFICATION

Databricks Certified Data Engineer Associate (Dec 23-25)

Databricks Certified Lakehouse Fundamentals (Feb 22)