Kumar Priyanshu

Data Engineer

+917004123035 ◇ kr.priyanshu0309@gmail.com ◇ Bengaluru, Karnataka, India ◇ LinkedIn ◇ krpriyanshu.app

SUMMARY

Analytical and detail-oriented Data Engineer with 3+ years of experience in designing, building, and optimizing scalable data pipelines using technologies like Apache Spark, Python, SQL, BigQuery, Dataproc and Hive. Proven track record in reducing data processing time by over 40% through process optimization and end-to-end ETL pipeline development.

SKILLS

Data Technologies   PySpark, Apache Hive, HDFS, Databricks, ETL, Data Warehouse, OLAP, BigQuery, Dataproc
Cloud Platforms   GCP, Databricks
Programming   Python, Shell Scripting, SQL (MySQL, Oracle)
Tools/Frameworks   Unix, Windows, Autosys, Jira, BitBucket, Confluence, VS Code, Intellij, Agile, Scrum, Jenkins, Git

PROFESSIONAL EXPERIENCE

Data Engineer   Aug '22 — Present
LTIMindtree   Bengaluru, India
- Developed and optimized data pipelines to process multi-terabyte datasets at a multinational investment bank, leveraging PySpark, Python and Hive to generate actionable insights, resulting in over 40% faster data processing.
- Utilised Spark SQL and PySpark functions to perform complex data transformations, joins, and data cleaning on 1000+ datasets, ensuring high-quality data for analysis and reporting.
- Loaded terabytes of high-quality, partitioned data into Hive tables in Parquet format, utilizing source file formats such as Avro, ORC, and CSV, reducing storage space up to 75% and enabling advanced analytics.
- Enhanced ETL workflows by implementing automated shell scripts, reducing manual work and testing efforts by 60%.
- Utilized Autosys for job scheduling and led the migration of jobs to Control Center, finishing 2 weeks ahead of schedule.
- Collaborated with stakeholders and Agile teams to refine requirements and implement scalable solutions, contributing proactively in Scrum events.

MERN Intern   Feb '22 — May '22
Mindtree   Bengaluru, India
- Designed and implemented 5+ dynamic user interfaces using React.js, enhancing user engagement by 25%.
- Contributed to web application development, improving project completion rates by 15%.
- Acquired hands-on experience with the MERN stack (MongoDB, Express.js, React.js, Node.js), advancing full-stack development capabilities.

EDUCATION

Bachelor in Information Science and Engineering, Sapthagiri College of Engineering   Jan '18 — Dec '22
                                                                                                    Bangalore

PROJECTS

Retail Analytics Data Pipelines   Jan '21 — Dec '21
- Developed data pipelines using PySpark and SQL for a retail analytics project, integrating data from multiple sources and generating insights for product performance and customer behavior.
- Enhanced the scalability of the data pipeline, enabling processing of over 5TB of data, leading to faster insights and better operational decision-making.

CERTIFICATIONS

Databricks Certified Associate Developer for Apache Spark 3.0

Python for Data Science, Coursera

SQL for Data Science, Coursera

AWARDS

Spot Award
Awarded with Spot award for exceptional dedication and noteworthy contribution to work