JAYKUMAR PATEL
Data Engineer

+1 (682) 392-4146 | jaykumarp362@gmail.com | Arlington, 76010 | LinkedIn

Experienced Data Engineer with 3+ years of expertise in designing and implementing data pipelines, ETL processes, and data solutions. Proficient in various programming languages, cloud technologies, big data ecosystems, and data engineering frameworks including Hadoop, MapReduce, Hive, Apache Spark, HDFS, Kafka, Airflow, and Snowflake. Seeking opportunities to contribute my data engineering skills to drive data-driven decision-making.

EXPERIENCE

Data Engineer | PNC Financial Services Group, GA | Jan 2024 – Current

- Designed and developed data pipelines and ETL processes for data ingestion, transformation, and loading.
- Worked with big data ecosystems, including Hadoop, MapReduce, Hive, Apache Spark, and HDFS for big data processing and analysis and enhanced existing ETL processes, reducing data loading times by 40% and increasing data accuracy.
- Leveraged cloud technologies (AWS, GCP, Azure) for data storage, compute, and analytics.
- Deployed Tableau dashboards and reports to Tableau Server on Azure.
- Automated data ingestion from various sources, reducing manual efforts by 60% and increasing data processing efficiency.
- Reduced the cost of data storage and processing by 20% by migrating to a cloud-based data platform.
- Developed real-time data streaming solutions using Kafka, enabling instant access to critical data for real-time analytics.

Data Engineer | Accenture, India | Feb 2019 – Dec 2022

- Utilized Python, and SQL to extract, transform, and load data from diverse sources into data lakes and data warehouses and automated testing for data pipelines to validate data transformations and maintain data quality.
- Established robust data monitoring solutions, reducing system downtime by 25% and ensuring data availability.
- Implemented data encryption and access control measures, ensuring compliance with data security standards and reducing security incidents by 20%.
- Collaborated with cross-functional teams, using project management tools like Jira, Confluence, Jenkins, and Git.
- Maintained documentation, trained team members, and provided support for data-related tasks.
- Built and maintained data engineering frameworks such as Kafka, Airflow, and Snowflake to support data operations.
- Conducted data analysis and machine learning using libraries like NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, and TensorFlow.
- Implemented ETL solutions using SSIS and SSAS, automating data workflows and enhancing data quality.
- Developed data visualizations and reports using, Power BI, and SSRS to aid data-driven decision-making.
- Managed and optimized databases, including MS SQL Server, PostgreSQL, and MySQL.
- Assisted in data-driven decision-making by providing actionable insights and recommendations.
- Implemented data quality checks and automated data cleansing processes, resulting in a 25% reduction in data errors and inconsistencies.

SKILLS

Methodologies:	SDLC, Agile, Waterfall
Programming Language:	Python, SQL, R
IDE's:	PyCharm, Jupyter Notebook, Visual Studio Code
Big Data Eco system:	Hadoop, MapReduce, Hive, Apache Spark, Pig, HDFS
ETL Tools:	SSIS, SSAS
Cloud Technologies:	AWS, GCP, Azure
Frameworks:	Kafka, Snowflake
Packages:	NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow
Reporting Tools:	Tableau, Power BI, SSRS
Database:	MS SQL Server, PostgreSQL, MySQL
Other Tools:	CMS, Jira, Confluence, Jenkins, Git, MS Office, ERP, CRM
Soft Skills:	Data Cleaning, Data Wrangling, Critical Thinking, Communication Skills, Presentation Skills, Problem- Solving
Operating System:	Windows, Linux

EDUCATION

Master of Science in Information systems | University of Texas at Arlington, Arlington

B tech in information and communication Technology | Pandit Deendayal Energy University, India