Venkata Sowmya Kopparapu
Data Engineer

✉ sowmya@jobsearchmail.net 📞 2067515415 📍 Redmond, WA in Linkedin

Profile

As a Data Engineer with 4+ years of experience, I have a proven track record in designing, implementing, and maintaining robust data infrastructure and pipelines. Proficient in Python, SQL, and Scala, I specialize in ETL processes, leveraging tools like Apache Spark and Hadoop for distributed data processing. Collaborating with cross-functional teams, I have developed data models, ensured data integrity, and implemented data security measures. With a strong analytical mindset, I continuously identify areas for improvement, delivering efficient and reliable data solutions.

Skills

Programming Languages
Python, Scala, SQL, NoSQL

Frameworks & Tools
Flask, Django, MS SQL, MySQL, PostgreSQL, MongoDB, Git, GitHub, Docker, CI/CD pipelines, AWS, Azure DataBricks, Jenkins, Statistics, Power BI, HTML, CSS, & Agile

Data Processing & Streaming
PySpark, Apache Kafka, Pandas, NumPy, Airflow, Snowflake, ETL

Big Data Tools
Spark, Hive, Presto, Apache Hadoop, Airflow, Snowflake, Kubernetes

Professional Experience

Data Engineer, *KPMG* — 12/2022 – present, Remote, USA
- Designing, creating, deploying, and managing the company's data architecture.
- Developing, constructing, testing, and maintaining databases and large-scale data processing systems.
- Building, maintaining, and optimizing ETL processes and data pipelines to improve data flow and collection efficiency.
- Ensuring data quality and accessibility for business intelligence and data analytics teams.
- Collaborating with data scientists, business intelligence analysts, IT team, and management to understand and deliver data requirements.
- Ensuring adherence to data privacy regulations and best practices. Implementing proper security measures in data handling and storage.
- Developing data models for database structures and building data warehouses to store and retrieve vast amounts of data efficiently

Data Engineer, *Capgemini* — 12/2017 – 06/2021, Bengaluru, India
- Developed and implemented strategies to optimize code and improve scalability of data processing pipelines, resulting in a 30% reduction in processing time and improved system performance.
- Conducted thorough analysis of existing code and data processing workflows to identify areas for optimization and improvement, resulting in a 25% improvement in processing time.
- Conducted a comprehensive analysis of customer data, identifying key drivers of customer satisfaction and loyalty, leading to a 15% increase in customer retention rates.
- Improved customer satisfaction by 20% through analysis of customer feedback and implementing data-driven improvements to products and services.
- Built Spark based application to detect data quality issues and alarm on call, which helped reduce ETL jobs failure and improve data processing.
- Improved system performance in data processing pipelines by handling edge cases where data is not available.
- Regularly handle production issues during on call by providing necessary hot fixes to mitigate issues and provide a plan for long term solutions.
- Detect patterns during on call and provide necessary documentation, scripts & commands to ease handling of issues during on call and reduce turnaround times.
- Performed in-depth analysis of historical claims data to identify emerging risk patterns and inform risk mitigation strategies.
- Analyzed large datasets to identify trends, patterns, and anomalies, providing valuable insights to inform business decisions.

Education

Bachelor of Engineering, *BVRIT College of Engineering* — 09/2013 – 05/2017, Hyderabad, India
Information Technology