Vineetha Bhavya Rayalla

Data Engineer

vbrayalla8@gmail.com |+1 470-485-4569 | TX

SUMMARY

Data Engineer with over 4 years of experience developing efficient ETL pipelines, enhancing data analysis, and implementing secure cloud storage solutions. Proven track record at BlackRock and Max Healthcare, where I improved data processing speeds by 30% and the accuracy of financial forecasts by 25%. Expert in SQL, Python, AWS, Azure, and machine learning algorithms, contributing to better financial and healthcare outcomes. Demonstrated success in project leadership, team collaboration, and improving data management practices. Committed to applying technical skills to optimize data-driven decisions and operational efficiencies.

SKILLS

Methodology: SDLC, Agile, Waterfall
Programming Language: Scala, R, Python, SQL
IDE's: PyCharm, Jupiter Notebook
Big Data Ecosystem: Hadoop, MapReduce, Hive, Apache Spark, Pig
ETL Tools: SSIS
Cloud Technologies: AWS, Azure, Google Cloud Platform
Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow, Kafka
Reporting Tools: Tableau, Power BI, SSRS
Database: MS SQL Server, Snowflake, MySQL
Other Tools: Git, MS Office
Operating Systems: Windows, Linux

EDUCATION

M.S. Computer Science | University of Alabama at Birmingham, Alabama

EXPERIENCE

BlackRock, TX | Aug 2022 – Current | Data Engineer

- Engineered and managed ETL pipelines using SSIS and Azure Data Factory, achieving seamless integration of diverse data sources, which enhanced the accessibility and reliability of data for financial analytics, and significantly improved the accuracy of financial reporting within the company.
- Deployed the Hadoop ecosystem, including MapReduce, Hive, and Apache Spark, to enhance data processing capabilities, achieving a 30% reduction in processing times and expediting market analysis and risk assessments.
- Developed predictive financial models using machine learning algorithms like Linear & Logistic Regression, Random Forest, and CNNs in Python, increasing the company's financial forecasting accuracy by up to 25%.
- Orchestrated the development of cloud-based data architectures on AWS, Azure, and Snowflake, integrating services like AWS Glue, S3, Athena, and Lambda to refine data analysis and reporting frameworks, streamlining financial data analysis.
- Used Case class in Scala to convert RDD's into Data Frame in spark
- Implemented Spark using Scala and performed cleansing of data by applying Transformation and Actions.
- Created dynamic reports and dashboards with Tableau and Power BI, transforming complex datasets into clear, actionable insights for stakeholders, and led the optimization of SQL Server and MySQL databases to elevate data integrity and expedite query resolutions.
- Drove the adoption of Agile methodologies, leveraging tools like Jira and Jenkins for streamlined project delivery, and championed version control best practices with Git, promoting a culture of continuous improvement and excellence in data management operations.

Max Healthcare, India | Jan 2019 - July 2021 | Data Engineer

- Orchestrated data workflows to merge, clean, and transform healthcare information using SQL, Python, and R, significantly improving system integration and data quality through strategic use of ETL tools like SSIS and Azure Data Factory.
- Developed and applied machine learning algorithms to predict patient health outcomes, empowering healthcare professionals to create more effective treatment plans.
- Implemented robust, secure cloud-based storage and computing solutions with AWS, Azure, and Snowflake, enhancing the agility and security of healthcare data analytics.
- Analyzed large-scale healthcare datasets in real-time using Hadoop and Apache Spark, providing immediate insights for clinical decision-making and operational improvements.
- Designed and disseminated dynamic, interactive dashboards with Tableau and Power BI, providing crucial data-driven insights to medical staff for improved patient care and operational efficiency.
- Led multidisciplinary teams to deliver projects on schedule and within quality standards, employing rigorous project management practices and innovative data management technologies like Docker and Airflow to enhance system robustness and efficiency.