Avineshwar Krishna Singh

Uttar Pradesh, India | avineshwarkrishnasingh@gmail.com | +91- 7814762078 | LinkedIn | GitHub

Summary

Highly motivated Data Engineer with over 3+ years of experience and a strong Computer Science background. Proven ability to leverage analytical skills and technical expertise to design and implement data solutions. Collaborative team player with a passion for knowledge sharing and a focus on achieving success through teamwork.

Experience

R1 RCM - Software Engineer ETL | Sep 2022 - Present
- Designed and maintained ETL pipelines (SQL, Python, SSIS, Azure Databricks) for data transformation, ensuring data accuracy, reliability, and accessibility for BI.
- Analyzed complex datasets to identify patterns, trends, and insights, informing business decisions and optimizing data processing/analysis.
- Collaborated with cross-functional teams to maintain data consistency throughout the ETL pipeline, meeting business requirements.
- Proactively identified and resolved data quality issues using various tools and techniques, delivering high-quality data for stakeholders.
- Migrated ETL solutions, slashing workload from 3 FTE to 0.5 FTE while maintaining quality through 90% workflow efficiency gains.

Assimilate Solutions, A SitusAMC Company - Quality Engineer ETL | Jan 2021 - Sep 2022
- Owned data validation for data service scrum team, ensuring data integrity using SQL queries and automated Python scripts on Snowflake.
- Validated Informatica mappings and workflows for accurate data processing and loading.
- Documented test cases and results for efficient data quality tracking and monitoring.
- Collaborated with onshore teams and stakeholders to meet data quality requirements

GlobalLogic - Associate Analyst | Oct 2019 - Jan 2020
- Leveraged crowd-sourced data to train Google Lens for rich shopping experiences.
- Collaborated on training Google Lens models for enhanced shopping experiences.
- Improved team efficiency and quality through data analysis and reporting.

Education

Lovely Professional University - B. Tech Computer Science & Engineering | 2015 - 2019
Sri Chaitanya Jr. College - Senior Secondary | 2013 - 2015
Saint Mary's School & College - High School | 2012 - 2013

Skills

- **Programming languages**: Python, Pyspark, SQL, T-SQL
- **Databases**: Snowflake, PostgreSQL, SQL-Server, MySQL
- **System Design**: Database Design, ETL - Pipeline, Integration Design
- **ETL - Tools**: SSIS, Informatica, Azure Databricks, Azure Data Factory
- **Reporting**: Tableau, Power BI, Excel
- **File Handling**: Txt, Xml, Tab, Csv, Excel, Json
- **Third Party Libraries**: Pandas, SQLAlchemy, Pyodbc, Numpy, ScikitLearn, Tensorflow Dask, Pyspark, Google-Pubsub_v1, datacompy, Py test, Pandera, Pysnooper
- **Other Tools & Technologies**: Bitbucket, Github, Jira, Azure, Agile methodologies, Scrum, Postman, API, Chat GPT, AI Prompt, Microsoft Office, Kafka, Zookeeper, Debezium, Docker

---

Projects

ETL-Pipeline
Leveraged multithreading techniques for parallel processing tasks, along with robust error handling featuring retry logic. Implemented comprehensive logging to facilitate efficient troubleshooting. Proficient in crafting automated alerts for success/failure notifications.

ETL- DataStreaming
Built real-time Python data pipeline leveraging Pub/Sub v1 for efficient message ingestion and parallel processing via multithreading. Implemented robust error handling with retries and automated success/failure alerts for streamlined troubleshooting and proactive issue management.

ML- Model
Developed stock prediction model using historical Yahoo Finance data. Employed technical indicators & probability for signal generation. Implemented feature selection & trained/evaluated multiple ML models. Optimized model performance through retraining & utilized saved models for real-time prediction.

Testing Tool - ETL
Developed Python GUI for ETL testing, automating data comparisons (flat files & databases), SQL queries, downstream checks, and report generation.

Certifications

- [Data science with Python for Data Analysis](#)
- [Snowflake – The Complete Masterclass](#)
- [The Complete Python Bootcamp From Zero to Hero in Python](#)
- [Power BI Essential Training](#)
- [Tableau Essential Training](#)