

Satya Vinod Kantipudi

Dallas, TX | +1(972)836-6556 | vinodkantipudi1510@gmail.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of Texas at Dallas

August 2022 - May 2024

Master of Science in Business Analytics

GPA: 3.95/4.0

- Relevant Coursework: database foundation for business analytics, big data, advanced statistics for data science, business data warehousing, cloud computing

Amrita School of Engineering

Aug 2017 - May 2021

Bachelor of Technology in Mechanical Engineering

GPA: 3.3/4.0

- Relevant Coursework: computer science fundamentals

WORK EXPERIENCE

Data Engineer Intern, Center for Brain Health

Feb 2023 – Present

- Engineered and deployed scalable ETL pipelines leveraging Python and AWS Glue, resulting in a 40% reduction in data processing time and enabling real-time analytics and reporting
- Led the initiative to streamline data collection processes by implementing DMS, resulting in a 60% reduction in data latency and enabling near real-time availability of data for reporting and analysis purposes
- Enhanced data accessibility by 80% for downstream applications by spearheading the implementation of AWS Glue Crawler to crawl and catalog raw data in AWS Data Catalog, enabling seamless data integration and analysis
- Automated AWS Glue jobs using Step Functions, resulting in an impressive 80% reduction in manual intervention

TECHNICAL SKILLS

Programming Languages: Python, R, SQL

Databases: PostgreSQL, Microsoft SQL Server, NoSQL

Tools and Software: Spark, Hadoop, Hive, Kafka, HBase, Cassandra, Airflow, Snowflake, Tableau

Amazon web services: S3, Lambda, Athena, IAM, RDS, Glue, Step Functions, API Gateway, Redshift, Quick Sight

Soft Skills: Problem solving, Communication, Collaboration, Attention to detail

ACADEMIC PROJECTS

AWS ETL Data Pipeline on YouTube Data

- Developed and deployed a highly efficient data pipeline using AWS Lambda, Glue, and S3, enabling seamless ingestion, processing, and storage of YouTube data. This empowered data-driven insights for strategic decision-making
- Implemented cost optimization strategies and performance enhancements for data lake architecture, resulting in saving the 100+ DPU hours per week

Built Scalable Cloud Data Pipeline with DBT, Snowflake and Airflow

- Developed modular ETL architecture using DBT to transform 1TB of data from S3 into analytical models in Snowflake
- Achieved near real-time data availability by scheduling Airflow to trigger DBT models every 5 mins as new data lands in S3
- Enabled robust monitoring via Slack alerts and email notifications using AWS SNS for pipeline runs handling 1TB data daily

Ecommerce Analytics Platform on Cloud

- Designed and built batch and real-time pipelines processing 50000 eCommerce records daily using AWS tools like S3, Glue, Athena, and Kinesis
- Architected a robust and scalable real-time pipeline with Lambda, DynamoDB and SNS to handle 10K events/sec for DDoS and bot attacks

EXTRACURRICULAR & LEADERSHIP ACTIVITIES

AWS Organization at UTD

- Led a team of 8 student volunteers to organize AWS workshops on big data analytics using S3, Glue and Athena for 50+ attendees
- Conducted 20+ interview preparation workshops for AWS roles, helping students build confidence through practice sessions