Venkata Sai Rishitha Seelam

🔗 LinkedIn | ✉ mailtosairishitha@gmail.com | 🌐 Portfolio | 🐱 Github | 🌍 Baltimore, MD

DATA ENIGINEER

Results-driven Data Engineer with expertise in designing and deploying scalable data pipelines, optimizing ETL workflows, and driving cloud-based data solutions. Proficient in AWS (S3, Glue, Lambda, Kinesis, Step Functions), Python, SQL, Snowflake, Informatica and Databricks with a track record of delivering actionable insights, reducing process times, and improving data governance.

PROFESSIONAL EXPERIENCE

Academic Engagement & Contributions
University Of Maryland- Baltimore County

Jan' 2024 – Dec' 2024
Baltimore, Maryland

- Conducted weekly recitations for two Data Management courses, covering SQL, relational databases, big data systems, and data warehousing.
- Collaborated with professors to design and develop an interactive dashboard using Power BI to track student performance and key metrics, enabling data-driven insights for academic improvement.
- Assisted in designing the syllabus and developing hands-on exercises in data analysis, database management, and business intelligence, ensuring alignment with industry best practices and data analytics standards.
- Provided lab support and one-on-one guidance to students on data visualization techniques, SQL query writing, and processing techniques using SQL, Power BI, Python (Matplotlib, Seaborn), enabling them to interpret and present data insights effectively.

AWS Cloud Data Engineer
Cognizant Technology Solutions (Healthcare Industry)

Jul' 2022 – Jan' 2023
Chennai, India

- Designed and implemented Data solution to manage customer rebates based on their purchase patterns.
- Optimized scalable ETL pipelines to process 5M+ records using AWS services, ensuring data integrity and consistency.
- Streamlined data availability using AWS Lambda and Glue, transforming data from four diverse sources into a centralized AWS S3 repository and snowflake, achieving a 25% reduction in data retrieval time.
- Implemented Snowflake-based data warehousing solutions managing over 3TB of data, improving query performance by 40%.

Data Engineer & Analyst
Cognizant Technology Solutions (Healthcare Industry)

Feb' 2021 – Jun' 2022
Chennai, India

- Worked on an end-to-end data project to analyze competitor performance relative to our client, delivering actionable insights and developing a dashboard to identify opportunities for revenue growth.
- Migrated legacy Python ETL workflows to PySpark and Databricks, reducing ETL runtime by 70% for 120GB datasets.
- Engineered and deployed ETL pipelines using Informatica PowerCenter to integrate data from 8+ sources, achieving 95% data accuracy.
- Created over 20 interactive dashboards in Tableau and Power BI, enhancing cross-functional reporting efficiency by 40%.
- Designed scalable data warehouse models in snowflake for storing and analyzing 2TB+ of data, enabling efficient storage and retrieval.
- Improved data preprocessing and reporting efficiency by 30% through automation and optimization, utilizing Python scripts, and Tableau/Power BI tools for raw data acquisition and preparation in a data science task.

EDUCATION

- **Masters in Data Science** | *GPA: 3.9*
  University of Maryland, Baltimore County
  Coursework: *Data Science, Big Data, Data Visualization, Database Systems, Cloud Services, Machine Learning*

  Dec' 2022 – Dec' 2024
  Baltimore, Maryland

- **Bachelors of Technology - Electronics & Communication Engineering** | *GPA: 8.6*
  SASTRA University
  Coursework: *Machine Learning, AI Fundamentals, Python, Java, C, C++, Virtual Reality*

  Jul' 2017 – Jun' 2021
  Thanjavur, India

ACADEMIC PROJECTS

US HOUSEHOLD ENERGY PREDICTION MODEL
Oct' 2024 – Dec' 2024

- * Developed a predictive model for U.S. household energy consumption using regression techniques with seasonal adjustments.
- * Optimized and deployed the model for scalability and real-world application.
- * Designed an interactive web application using HTML, CSS and JavaScript for users to visualize energy consumption predictions.

GENE SEQUENCE CLASSIFICATION
Aug' 2023 – Dec' 2023

- * Contributed to the advancement of Splice Junction gene sequence classification through machine learning techniques.
- * Preprocessed gene sequence data to prepare it for machine learning algorithms.
- * Utilized Python, NumPy, Pandas, and Scikit-Learn for data preprocessing and model development.
- * Applied various machine learning algorithms to classify gene sequences.
- * Evaluated model performance using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC.

FACIAL EMOTION RECOGNITION
Jan' 2023 – May' 2023

- * Developed a highly accurate machine learning model for detecting and classifying human emotions.
- * Utilized advanced machine learning algorithms & Computer Vision to train the model on a diverse dataset of facial expression.
- * Achieved high precision and recall in accurately identifying and classifying various human emotions.
- * Implemented feature engineering techniques to extract meaningful facial features for emotion recognition.

TECHNICAL SKILLS

- **Programming & Databases:** Python, C++, PySpark, R, MSSQL, MySQL, SSMS
- **Cloud & Big Data:** AWS (S3, Glue, Lambda, Redshift, Athena), GCP, Databricks, Apache Spark, Spark SQL, Kafka, Hadoop
- **Data Warehousing & ETL:** Snowflake, Datalake, Informatica PowerCenter, Informatica MDM
- **Data Analysis & Visualization:** Power BI, Tableau, Excel
- **Tools:** GitHub, Docker, Agile/Scrum, Machine Learning & Forecasting, VS Code, PyCharm, Anaconda, Jupyter
- **Libraries:** Pandas, NumPy, Scikit-learn, SciPy, Matplotlib, Seaborn
- **Others:** GitHub, Docker, Agile/Scrum, Machine Learning & Forecasting

CERTIFICATIONS

- AWS Certified Data Engineer - Associate (***Amazon Inc***)
- Microsoft Power BI Desktop (*Maven Analytics | Udemy*)
- Data Analytics Certification (*Microsoft | LinkedIn Learnings*)

ACHIEVEMENTS & LEADERSHIP ROLES

- Received the Innovation Catalyst Award at Cognizant Technology Solutions for optimizing complex workflow and data processing, leading to significant performance improvements.
- Head of DAKSH (Technical fest of SASTRA) overseeing workshops and clusters while coordinating and directing various technical sessions during the event.
- Active member in the coding club of SASTRA – Conducted workshops during the fests.