

# AWS DATA ENGINEER

**SAI SUNANDH RAMAYANAM**

Phone: +1 928 550 2812

Email: [saisunandhramayanam@gmail.com](mailto:saisunandhramayanam@gmail.com)

Portfolio: [sai-sunandh.github.io/](https://sai-sunandh.github.io/)

## PROFESSIONAL SUMMARY:

---

- With over 3+ years of experience in data engineering and AWS cloud services, I have a proven track record of designing, implementing, and optimizing data architectures to support scalable, secure, and high-performance analytics solutions
- Led stakeholder requirements sessions to align on-premises infrastructure with AWS capabilities and conducted detailed meetings with stakeholders to understand existing infrastructure, data sources, and compliance requirements, ensuring a clear roadmap for transitioning to AWS cloud solutions.
- Designed and deployed secure and developed a robust AWS architecture incorporating Amazon S3 for scalable storage, Amazon Redshift for efficient data warehousing, and Amazon QuickSight for dynamic data visualization, optimizing overall data accessibility and migration processes.
- Orchestrated setup of AWS resources and established and configured essential AWS resources including EC2 instances for compute capacity, VPCs for network isolation, and IAM roles for secure access management, creating a secure and isolated environment for data migration and processing.
- Conducted comprehensive data mapping and schema design for Amazon Redshift: Created detailed data mapping and schema designs, defining source-to-destination transformations and designing target schemas for Amazon Redshift to ensure accurate and seamless integration of diverse data sources.
- Utilized AWS Glue and PySpark running on AWS EMR to perform efficient ETL (Extract, Transform, Load) processes, transforming and cleaning data to enhance data quality and integrity.
- Utilized AWS DMS for seamless extraction and migration of on-premises databases while Employing AWS Data Migration Service (DMS) to facilitate the extraction and migration of on-premises databases to AWS cloud storage, ensuring an efficient and reliable data transfer process.
- Optimized data loading into Amazon Redshift and S3 and enhanced the performance of data storage and retrieval by optimizing the data loading processes into Amazon Redshift and S3, ensuring scalable analytics and efficient reporting capabilities.
- Executed rigorous data validation post-migration and conducted extensive data validation and testing activities following data migration, ensuring the accuracy, integrity, and completeness of the data across AWS environments.
- Conducted performance testing and query optimization on Amazon Redshift and Performed thorough performance testing and query optimization on Amazon Redshift, improving data processing efficiency and enabling faster access to critical insights for stakeholders.
- Created user-friendly and interactive dashboards and reports using Amazon QuickSight, providing stakeholders with valuable business insights and facilitating informed decision-making.
- Implemented AWS KMS and IAM policies for data encryption and compliance and Applied AWS Key Management Service (KMS) and IAM policies to implement robust data encryption and access control measures, ensuring data security and adherence to regulatory requirements such as HIPAA.
- Leveraged Python (Pandas, NumPy, Scikit-learn) for data analysis and machine learning and Utilized Python libraries such as Pandas, NumPy, and Scikit-learn to perform advanced data analysis and machine learning tasks, leveraging Python's powerful ecosystem for data-driven solutions.
- Developed Python scripts for data preprocessing while Created custom Python scripts to preprocess and clean large datasets, enhancing data quality and preparing data for accurate modeling and analysis.
- Implemented machine learning models (Gaussian Naïve Bayes, Random Forest) in Python, Developed and deployed machine learning models using algorithms like Gaussian Naïve Bayes and Random Forest, generating predictive insights and supporting data-driven decision-making processes.
- Integrated Java applications with AWS services (S3, Lambda) and Enhanced data processing capabilities by integrating Java applications with AWS services such as Amazon S3 for storage and AWS Lambda for serverless computing, ensuring scalable and efficient backend solutions.
- Developed RESTful APIs and backend services in Java and Designed and developed RESTful APIs and backend services using Java, facilitating seamless integration of data processing workflows with AWS infrastructure and enhancing overall system performance.

**SKILL MATRIX**

Category	Skills
Cloud Services	AWS (EC2, Lambda, Glue, Redshift, RDS, S3, SNS, SQS, CloudWatch, CloudTrail, IAM, AWS Elastic Load Balancing (ELB), AWS Auto Scaling, AWS CloudFormation, CloudFront, Elastic Beanstalk, AWS Machine Learning Services), Microsoft Azure (Azure Data Factory, Azure Synapse Analytics), Google Cloud Platform (BigQuery, Dataflow)
Big Data Technologies	Apache Hadoop, Apache Spark, Apache Kafka,
Database Technologies	MySQL, SQL, Amazon RDS, Amazon Redshift, Data Warehousing
DevOps	AWS Development Operations (DevOps), Jenkins, Docker, Heroku, Kubernetes
Data Analytics	Pandas, NumPy, SciPy, Data Analytics, Data Science, Data Structures & Algorithms, Business Analytics, R, Excel/Numbers/Sheets
Machine Learning	Scikit, Machine Learning, AWS Machine Learning Services
Programming Languages	Python 3.12.0 , C/C++, Java
Web Development	HTML/CSS, Django, Web Development

## WORK EXPERIENCE

### AWS DATA ENGINEER || ACCENTURE

September 2021 – December 2022

- Led requirements gathering sessions with stakeholders to understand current on-premises infrastructure, data sources, and compliance needs, ensuring alignment with project objectives and AWS capabilities.
- Developed a scalable and secure AWS architecture using Amazon S3 for storage, Amazon Redshift for data warehousing, and Amazon Quick Sight for visualization, ensuring seamless data migration and accessibility.
- Set up AWS resources including EC2 instances, VPCs, and IAM roles, ensuring secure and isolated environments for data migration and processing.
- Conducted comprehensive data mapping and schema design, outlining source-to-destination transformations and target schema for Amazon Redshift.
- Implemented AWS Glue for ETL processes, leveraging PySpark on AWS EMR to transform and clean data, ensuring high data quality and integrity.
- Used AWS Data Migration Service (DMS) to extract data from on-premises databases, ensuring efficient and reliable data transfer to AWS cloud storage.
- Loaded transformed data into Amazon Redshift and Amazon S3, optimizing data storage and retrieval for high performance and scalability.
- Performed rigorous data validation and testing post-migration to ensure accuracy, integrity, and completeness of the migrated data.
- Conducted performance testing and query optimization on Amazon Redshift, ensuring efficient data retrieval and processing.
- Created interactive dashboards and reports using Amazon QuickSight, providing stakeholders with valuable business insights and data visualization.
- Configured IAM roles and policies, AWS KMS for encryption, and CloudTrail for auditing, ensuring robust security and compliance with HIPAA.
- Reduced turnaround time by 30% through streamlined grading and feedback processes for Data Structures, Large Scale Data Structures, and Python courses at Northern Arizona University.
- Maintained accurate grade records and facilitated effective communication between students and faculty, enhancing the academic experience and support system.
- Coordinated project logistics, including resource allocation, timelines, and risk management, ensuring smooth execution and delivery.
- Delivered comprehensive training sessions for Fresenius Kabi's staff on the new AWS-based infrastructure, ensuring seamless adoption and usage.
- Developed detailed documentation of the new architecture, processes, and operational guidelines, ensuring clear knowledge transfer and support.
- Implemented data encryption and compliance measures, ensuring all data migration activities adhered to industry standards and regulations such as HIPAA.
- Utilized CloudWatch for real-time monitoring and logging of AWS resources and applications, ensuring optimal performance and quick issue resolution.
- Executed the cut-over plan and provided post-migration support, ensuring a smooth transition and minimal disruption to operations.
- Mentored new team members through knowledge transfer sessions, fostering a collaborative and skilled project team environment.
- Interacted with clients and project stakeholders to gather feedback, provide updates, and ensure project deliverables met quality and timeliness standards.

### TEACHING ASSISTANT || NORTHERN ARIZONA UNIVERSITY

January 2023 – May 2024

- Mentored students in Data Structures, Large Scale Data Structures, and Principles of Programming Languages (Python), providing comprehensive tutoring and grading support for approximately 150 students per course, fostering a deep understanding of data engineering principles.
- Developed personalized tutoring sessions, significantly enhancing student comprehension and performance in complex topics such as algorithm analysis, data structure optimization, and data processing techniques.
- Streamlined grading and feedback processes for assignments and exams in Data Structures, Large Scale Data Structures, and Python courses, reducing turnaround time by 30% and ensuring timely, constructive feedback, analogous to optimizing data workflows in ETL processes.
- Utilized data engineering skills to maintain accurate grade records, employing efficient data management techniques to ensure data integrity and accessibility.

- Facilitated effective communication between students and faculty, leveraging skills in data visualization and reporting to provide clear and actionable insights into student performance and progress.
- Applied principles of data analysis and data quality management to develop robust assessment methods, ensuring fair and consistent evaluation of student work.
- Implemented automated solutions for repetitive administrative tasks, akin to automating data pipelines, to improve efficiency and accuracy in grade record maintenance.
- Guided students in practical applications of data structures and algorithms, preparing them for real-world data engineering challenges and enhancing their problem-solving skills.
- Conducted workshops on advanced Python programming and data processing techniques, equipping students with essential skills for data engineering roles.
- Collaborated with faculty to design and update curriculum content, integrating the latest data engineering concepts and practices to keep the coursework relevant and challenging.

## **PROJECTS**

### **DATA WAREHOUSING**

- Designed and implemented a data warehouse using dimensional modeling principles, incorporating AWS Redshift and S3, resulting in a 30% performance improvement over traditional RDBMS solutions.
- Utilized AWS Glue and PySpark for ETL processes, enabling the seamless integration of diverse data sources into the data warehouse and ensuring data consistency and reliability.
- Applied advanced analytical queries extensively in Microsoft Access, MySQL, and Amazon Redshift to extract valuable insights and facilitate data-driven decision-making for business stakeholders.
- Implemented indexing and query optimization techniques in MySQL and Redshift to enhance data retrieval speed and efficiency within the data warehouse environment.
- Developed robust reporting mechanisms using Amazon QuickSight, enabling effective presentation of analytical findings and aiding stakeholders in making informed business decisions.
- Employed data engineering best practices to ensure data quality and integrity, using data validation and cleansing techniques within the ETL pipeline.
- Leveraged AWS Data Pipeline and Step Functions to automate data workflows, improving the efficiency and reliability of data processing and integration tasks.
- Created detailed documentation of data warehouse architecture, ETL processes, and query optimization strategies, ensuring clear knowledge transfer and maintainability.
- Collaborated with data scientists and analysts to design and implement scalable data solutions, supporting advanced analytics and machine learning projects.
- Conducted performance tuning and monitoring using AWS CloudWatch, ensuring optimal performance and quick issue resolution within the data warehouse environment.

### **DISEASE PREDICTION MODEL USING MACHINE LEARNING**

- Pioneered a disease prediction model utilizing machine learning techniques, leveraging NumPy, Pandas, and Scikit-learn in Python to process and analyze large datasets effectively.
- Achieved an accuracy rate of 90% using the Gaussian Naïve Bayes algorithm, demonstrating the model's effectiveness and potential for real-world applications in the medical field.
- Conducted extensive data preprocessing and feature engineering, improving model accuracy and reliability by transforming raw data into meaningful features for the prediction model.
- Collaborated with a multidisciplinary team, integrating domain expertise into the model development process to ensure the model's relevance and applicability to medical data.
- Validated and tested the model using cross-validation techniques, ensuring robustness and generalizability by preventing overfitting and assessing performance on unseen data.
- Documented the entire model development process, including code, methodology, and results, to facilitate reproducibility and knowledge sharing among team members and stakeholders.
- Employed AWS SageMaker for scalable model training and deployment, ensuring efficient use of computational resources and easy integration with other AWS services.
- Utilized Amazon S3 for data storage and management, ensuring secure and scalable handling of large datasets used in the model development process.
- Implemented monitoring and logging using AWS CloudWatch, enabling real-time tracking of model performance and quick identification of any issues or anomalies.

- Developed visualization tools and dashboards using Amazon QuickSight to present model predictions and insights effectively, aiding stakeholders in understanding and utilizing the model's outputs for decision-making.

## **DEEP STATE TESTING**

- Implemented the Deep State framework in C++ to enhance software assurance testing, focusing on robustness and reliability.
- Successfully achieved a 40% decrease in production defects through rigorous testing methodologies and proactive bug detection strategies.
- Elevated overall software quality by 30% through systematic testing and implementation of quality improvement initiatives.
- Attained a code coverage of 41.43% by conducting thorough coverage analysis, ensuring extensive testing of critical code paths.
- Streamlined testing processes and methodologies to enhance efficiency and effectiveness, contributing to smoother software development lifecycle outcomes.

## **COVID SUPPORT SYSTEM – DJANGO WEB APPLICATION**

- Developed a web application using the Django framework to support COVID-related initiatives, leveraging its robust features for rapid and secure web development.
- Utilized Python for backend development, ensuring robust functionality, seamless data management, and efficient handling of COVID-related information.
- Implemented HTML and CSS for frontend design, focusing on creating user-friendly interfaces and ensuring accessibility for all users.
- Deployed and managed web applications on AWS S3 buckets, ensuring high availability, scalability, and security for seamless user access and performance.
- Integrated various features such as user authentication, data storage with Amazon RDS, and interactive components to enhance usability and provide a comprehensive user experience.
- Ensured scalability and performance optimization through efficient coding practices, deploying the application on AWS Elastic Beanstalk for seamless scaling and management.
- Utilized AWS Lambda for serverless computing, handling specific tasks and reducing the server load, improving the application's responsiveness.
- Employed AWS S3 for data storage and backup, ensuring secure and scalable storage solutions for user data and application assets.
- Implemented monitoring and logging using AWS CloudWatch, providing real-time insights into application performance and enabling quick resolution of any issues.
- Ensured data security and compliance with IAM roles and policies, implementing best practices for user data protection and privacy.
- Documented the entire development process, including code, features, and deployment strategies, to facilitate knowledge sharing and future maintenance.

## **EDUCATION**

---

### **NORTHERN ARIZONA UNIVERSITY || MASTER'S, COMPUTER SCIENCE**

JANUARY 2023 – MAY 2024

- Proficient with strong coding skills in Python, C++, and C, capable of developing efficient, scalable, and robust software solutions across various applications and systems.
- Advanced knowledge of data structures, modern regression, Expertise in designing and implementing complex data structures, performing regression analysis, and leveraging data analysis tools such as R and SQL to extract actionable insights and drive data-driven decisions.
- Skills in data warehousing and network security: Experienced in designing and managing data warehouses for optimal data storage and retrieval, along with implementing network security measures to ensure data protection and compliance with industry standards.

## **AWARDS & ACHIEVEMENTS**

- 
- Secured a full-time opportunity by achieving an All-India Rank of 3834 out of 350,000 participants in TCS Code Vita Season 9, demonstrating exceptional problem-solving and coding skills.
  - Recognized as the top performer within the European project group at Accenture, contributing significantly to project success and team performance.
  - Earned a Certificate of Achievement for completing "Step into Robotic Process Automation" during GUVI's RPA SKILL-ATHON 2020, showcasing proficiency in automation technologies.

- Received the College Level BEST NSS VOLUNTEER Award for outstanding service during 2018-2019, highlighting dedication to community service and leadership abilities.
- Awarded a Certificate of Merit for a notable poster presentation on ENGINEER'S DAY, reflecting strong communication skills and the ability to convey complex technical concepts effectively.

## **CERTIFICATIONS**

---

- AWS cloud Foundations
- DevOps
- Programming for Everybody (Getting Started with Python), Python (basic)
- Data Science Foundations, Data Analysis using Python, Machine learning using Python
- Power Platform Fundamentals
- SQL
- Jenkins
- Web Development Using Django

## **VOLUNTEERING**

---

- Coordinated and led volunteer events and activities at Accenture: Organized and participated in various events, such as blood donation camps and fundraising initiatives, demonstrating a commitment to corporate social responsibility and community engagement.
- Active participant in One Step Towards Society (environment wellness plantation NGO): Engaged in tree planting and environmental wellness activities, promoting sustainable practices and contributing to environmental conservation.
- Organized and supported crisis and tragedy fundraising events: Led and participated in fundraising activities to support communities affected by crises and tragic events, showcasing strong leadership and compassion.
- Student Ambassador, NAU: Provided support to 500 new students during campus events, facilitating their seamless integration into university life through effective communication and resource coordination.
- Class Representative, Pragati Engineering College: Enhanced student-faculty communication and organizational efficiency by spearheading events, fostering a collaborative academic environment, and advocating for student interests.
- Student Coordinator, Technical Fest, Pragati Engineering College: Orchestrated the planning and logistics of technical events, ensuring their successful execution through meticulous coordination, leadership, and attention to detail.