Aryaman Rao

Chicago, Illinois(Open to Relocate)
312-678-1614 aryamanrao18@gmail.com My GitHub Repository My Portfolio LinkedIn

Education

University of Illinois at Chicago **Aug 2022 – May 2024**
Masters of Science in Computer Science *Chicago, Illinois*

Experience

University Of Illinois at Chicago **Sep 2022 – Present**
Senior Data Engineer *Chicago, Illinois*

- Orchestrated a complete overhaul of the university's data warehouse solutions, deploying advanced data solutions like Amazon Redshift, Snowflake, Google BigQuery, and Oracle within a robust AWS and Google Cloud Platform (GCP) infrastructure. This initiative led to a 70% increase in query performance and a significant expansion of the university's analytics capabilities, integrating Apigee for enhanced API management and adding compliance measures. Leveraged ElastiCache for improved caching capabilities.
- Spearheaded the redesign of ETL workflows, integrating Apache Airflow, AWS Glue, Databricks, and utilizing PL/SQL to cut processing times by over 60%. Implemented comprehensive data integrity and quality checks using Deequ, enhancing the reliability of data models and analytics outputs. Added automation tools for streamlined operations, employing Bash and managing data ingestion processes.
- Developed analytical solutions using Tableau and Microsoft Power BI, coupled with predictive analytics models in Python and R, to facilitate data-driven decision-making and strategic insights across various academic departments, effectively employing HTML, CSS, and JavaScript to enhance frontend data visualizations. Implemented configuration management processes and utilized Apex for streamlined programming.
- Optimized SQL and BigQuery operations, employing AWS cost-management practices to halve query times and substantially reduce operational costs. Facilitated quicker, more economical access to analytics insights, employing predictive analytics and data mining to streamline decision-making processes. Utilized DynamoDB for efficient NoSQL data modeling and MFA for enhanced security, and applied OLAP techniques for multi-dimensional analytical queries.
- Established stringent data governance protocols to ensure GDPR and CCPA compliance, utilizing Apache Atlas for metadata management and AWS Lake Formation to secure data lakes, significantly enhancing data privacy and security measures. Integrated SSO for user authentication. Enhanced communication skills across multidisciplinary teams, incorporating statistics to better analyze data trends and insights.
- Utilized Kubernetes for efficient container orchestration and Terraform for infrastructure management, which collectively reduced operational costs by 40% and improved scalability and resilience of cloud resources. Also integrated Splunk for monitoring and analyzing machine-generated big data. Emphasized DevOps methodologies, including Kanban, to streamline workflows and utilized Unix for system management. Explored functional programming to improve data processing efficiency.
- Led the adoption of a cloud-first approach, promoting the use of AWS SageMaker and Azure Machine Learning for research purposes, thereby fostering a culture of innovation and making significant contributions to academic projects and studies. Integrated PowerShell for scripting and automation, and implemented Salesforce for CRM functionalities.
- Configured and maintained CI/CD pipelines using Jenkins and CircleCI, alongside AWS CodePipeline, to ensure high reliability and continuous integration of data services, enhancing deployment frequency and system stability. Highlighted Agile methodologies to optimize project management. Employed REST for API management and utilized Kotlin on the JVM for backend developments.
- Mobilized and led cross-functional teams to align data initiatives with institutional business objectives, achieving a 15% increase in operational efficiency and revenue. Utilized Confluence for streamlined project documentation and Jira for efficient task management, enhancing project delivery and team collaboration. Focused on interpersonal skills to enhance team interactions.
- Curated and conducted advanced workshops and training sessions on best practices in data engineering and cloud analytics for faculty and staff, promoting a culture of innovation, collaboration, and continuous learning. Implemented Infrastructure as a Service (IaaS) solutions to support educational platforms. Highlighted the use of Java and C# in developing tailored educational tools. Introduced essential concepts of Hadoop, Hive, and HBase to enhance big data processing capabilities.

Nekthan Gaming **April 2021 – July 2022**
Lead Data Engineer *Hyderabad, India*

- Spearheaded the design and implementation of a cutting-edge, scalable data architecture on Microsoft Azure, utilizing Azure Data Lake, Azure Databricks, and Azure Synapse Analytics. This initiative was central to driving a comprehensive analytics transformation across the gaming industry, enabling deeper insights into player behavior and game performance through product analytics. Established robust data structures for optimizing data storage and retrieval.
- Led the strategic migration of gaming analytics systems to cloud-based platforms, achieving a 40% increase in data processing efficiency. This pivotal move significantly reduced operational costs while enhancing the agility of the data analytics framework to emerging gaming trends.
- Developed and fine-tuned real-time data processing pipelines using Apache Kafka, Apache Spark, and Apache Flink. This advancement was key in managing billions of gaming events per day with minimal latency, dramatically improving the real-time analytics of in-game actions and player interactions. Successfully implemented MySQL for efficient data management and storage solutions.
- Implemented advanced data analytics and machine learning models in close collaboration with the data science team, which were instrumental in providing predictive insights. These insights led to significantly improved game design decision-making processes, directly impacting player engagement and retention.
- Directed the adoption of CI/CD practices for data pipelines using Jenkins and GitLab, markedly improving the deployment frequency of game analytics features. This initiative also ensured higher data quality and reliability, directly contributing to a more robust gaming experience.
- Championed data governance and compliance initiatives within the gaming sector, ensuring strict adherence to GDPR and CCPA regulations. Employed data cataloging tools like Apache Atlas and Collibra to maintain high standards of data privacy and security, reinforcing player trust.
- Played a pivotal role in fostering a data-driven culture within Nekthan Gaming. This was achieved by advising clients and conducting workshops on best practices in business intelligence and analytics, emphasizing the unique requirements of the gaming industry.
- Conducted comprehensive evaluations of new data technologies and tools, leading the integration of innovative solutions such as Airflow for workflow automation and Docker for containerization. These efforts created a more agile and efficient data engineering environment, crucial for rapidly evolving game development cycles. Designed and developed prototypes to test the efficacy of new technologies before full-scale implementation.
- Expertly integrated Cassandra and MongoDB for handling large volumes of unstructured game data, along with PostgreSQL for structured data management. These database technologies were essential in supporting the complex data needs of multiplayer online games, enhancing the scalability and performance of the gaming platform. Emphasized the use of JSON for data interchange and storage.

Markor Technologies **May 2019 – March 2021**
Data Engineer- 2 *Hyderabad, India*

- Led strategic AWS deployments, enhancing data processing efficiency with S3, Redshift, and Glue. Applied IaC, auto-scaling, and load balancing to support a 25% market reach growth for an e-commerce leader reducing operational costs while employing C for backend development. Focused on creating robust data models for scalable and efficient data processing.
- Revolutionized data pipelines with Apache Airflow, improving efficiency. Implemented vibrant DAGs and advanced monitoring, slashing analytics turnaround by 50%, which significantly enhanced business decision-making.
- Developed a real-time analytics platform using Amazon Kinesis and Tableau on AWS, boosting customer engagement by 30% and conversion rates. Employed scalable microservices and real-time processing, leveraging machine learning for predictive insights.
- Integrated AWS IAM and Amazon Macie, establishing robust data security that ensured 100% GDPR and CCPA compliance. Reduced compliance inquiries, enhancing customer trust with automated monitoring and reporting. Applied strong time management skills to handle multiple projects and meet tight deadlines.
- Spearheaded legacy system migration to AWS, halving processing times and enhancing decision-making speed by 35%. Exercised AWS Migration Hub and Snowball for efficient transfers, fostering an agile organizational culture through cloud-native adoption. Upgraded the data serialization process and incorporated Parquet format for optimized storage efficiency. Managed version control and collaboration using Bitbucket.
- Integrated SOAP web services for seamless communication between heterogeneous systems, enhancing the interoperability and data exchange capabilities of the organization's IT infrastructure.

Projects

Real-time Fraud Detection and Transaction Processing using Advanced Machine Learning Techniques | **January 2023**

- Orchestrated the deployment of an avant-garde system leveraging Apache Kafka for seamless live data streaming, dramatically amplifying fraud detection efficiency within the financial ecosystem.
- Innovatively crafted and integrated tailor-made ML algorithms, achieving an extraordinary 95% success rate in pinpointing fraudulent transactions, substantially mitigating false positives.
- Masterful navigated the management of over 100 million transactions daily via Apache Spark, securing ultra-low latency in fraud detection while upholding stellar system performance.
- Deployed a comprehensive array of datasets to empirically demonstrate a machine learning model's capacity to diminish fraud by 60%, thereby potentially safeguarding assets in excess of $500 million, exemplifying the project's scalability.

Scalable Data Pipeline Construction and Optimization for Real-time Analytics | **November 2022**

- Unified EHRs, wearables, and social determinants using Apache NiFi and Talend, incorporating real-time analytics and blockchain for secure, interoperable data sharing, enhancing data integrity and responsiveness.
- Implemented HL7 FHIR protocols with smart contracts to automate and ensure consistent data exchanges, streamlining healthcare operations and enhancing ecosystem efficiency.
- Engineered predictive models with TensorFlow, augmented by federated learning for decentralized, privacy-preserving analytics, personalizing healthcare management while safeguarding patient data.
- Crafted a user-friendly interface with React.js and D3.js, integrating advanced data visualization like interactive 3D models for accessible, engaging health data insights and improved patient outcomes.

Serverless BMI Calculator Web App | **October 2020**

- Deployed a serverless BMI calculator web application utilizing HTML, CSS, and JavaScript for the front end, coupled with AWS Lambda functions and API Gateway for the back end.
- Executed deployment on AWS, deploying the application on Amazon S3, providing tailored diet plans contingent on BMI categories.
- Hosted the application on Amazon S3, link: http://serverless-bmi-calculator.s3-website-us-east-1.amazonaws.com/

Technical Skills

Languages: Python, Java, C, c++, HTML/CSS, JavaScript, SQL
Developer Tools: Visual Studio Code, GIT, TensorFlow, PyTorch, Pandas, Numpy, Apache Spark, NLTK, Bootstrap, Django, Flask, Jquery, Angular, React, Django, React, Express.js, Apache spark, Pyspark, Node.js, D3.js, Devops, Eclipse, Google Cloud Platform, Android Studio
Technologies/Frameworks: Linux, Jenkins, Agile methodology, AWS, Azure, Rest API, Web API, MongoDB, PostgreSQL, GitHub, JUnit, WordPress, Kubernetes, Helm, ArgoCD, Grafana, Prometheus, Datadog, Development Life cycle, ServiceNow
Certification: Azure Fundamentals, AWS Certified Cloud Practitioner, Introduction to HTML, CSS, Javascript
Publications :link: http://doi.one/10.1729/Journal.28537