ANKITA ASHOK INDI

📞 425-866-5218   ✉ indi.a@northeastern.edu   in linkedin.com/in/ankita-indi/   ○ github.com/ankitaindi

PROFESSIONAL EXPERIENCE

iSpecimen   **May 2022 – Jan 2023**
Data Engineer Co op

- Developed a solid proficiency in data analysis, data mining, reporting, and dashboard creation, transforming complex data into actionable insights
- Reduced overall run time of **ETL** pipelines by **40%** by critical migrations of metadata, old staging environment, views and reports from PostgreSQL to Snowflake
- Collaborated with the team to build business cases for initiatives aimed at improving business operations
- Created compelling visualizations and dashboards in Power BI and Tableau that presented intricate data findings in a clear and concise manner, aiding decision-making processes

Infosys Limited   **Jun 2018 – July 2021**
Data Engineer

- Engineered a data driven platform using **Informatica** ETL for warehousing 30M orders data from Salesforce CRM
- Developed a rule base alert system to notify Data Engineering and Data Science teams about critical 100+ data sources
- Led the team of 3 associates to develop and deploy 100 jobs using **CI/CD** Jenkins, Bitbucket and Git within 3 months
- Performed Query optimization and tuning,improving DB performance and decreasing DB resource consumption by **37%**
- Tuned data reconciliation process to magnify data accuracy by 25 per cent and cut down **ETL** workflow run time by 1 hour
- Collaborated with a cross-functional team of software engineers and product managers to develop a web-based application using Java contributing to the company's client management system
- Reduced infrastructure costs by **15%** by right-sizing resources and implementing auto-scaling policies
- Actively participated in scrum ceremonies with distributed **Agile** teams following software development life cycle SDLC

TECHNICAL SKILLS

Programming languages: Java with DataStructures, Python, SQL, PL/SQL, PySpark, Shell Scripting
Databases and Operating System: MySQL, Oracle, Postgres, SQL Server, Windows, Linux
ETL & BI: Informatica PowerCenter, Alteryx, Talend, Tableau, Power BI, Data Profiling, Data Warehousing
Cloud Services and Tools: AWS, VPC, EBS, ELB, Auto Scaling, IAM, EC2, S3, RDS, Lambda, Cloudwatch, SQS, Git/GitHub
Tools/CI/CD: Snowflake, SDLC, CI/CD, Distributed Storage Systems, Docker, Kubernetes, Terraform

EDUCATION

Northeastern University   **Sep 2021 – Dec 2023**
Master of Science in Information Systems

KLE Technological University   **June 2014 – May 2018**
Bachelor of Engineering – Computer Science

PROJECTS

RESTful services using Cloud | *Spring Boot, MySQL, Terraform, S3, CloudWatch*
- * Implemented a 3-tier RESTful service on AWS using Java Spring Boot, MySQL, and Terraform for provisioning the VPC, compute environment, database, and storage. Utilized Terraform to ensure infrastructure consistency across multiple deployments
- * Developed a robust CI/CD pipeline using GitHub actions to automate the deployment processes. Used CloudWatch to monitor and alert on system metrics, such as CPU utilization and memory usage

Designing Data Warehouse and Business Intelligence for IMDB | *SCD2, MYSQL, Tableau, PowerBI*
- * Integrated a huge dataset of 80 million records from multiple sources like MySQL, Oracle, SQL Server, PostgreSQL, BigQuery, and flat files to staging and integration databases by building ETL data pipelines and performed source to target mapping in Talend Studio
- * Executed data architectures, data pipelines, data modeling and price history tracking with **SCD 2** for technical analysis
- * Performed data cleaning, profiling, and data preparation using Alteryx, implemented error handling using reject tables, performance tuning, and load statistics leading to a faster load of transformed data from 6 hours to 20 minutes
- * Built interactive data visualization dashboards and reports to gain business insights from IMDB data using Tableau and Power BI