

Ankur Rajput

Experienced Data Engineer adept with more than 7 years of experience in crafting scalable data pipelines and infrastructure. Skilled in ETL processes, diverse programming languages, and optimizing data systems for efficiency. Proficient in data modeling, warehousing, and visualization, facilitating actionable insights. Effective collaborator with a track record of driving data-driven decisions and innovation in collaborative environments.



[linkedin.com/in/ankur-rajput56/](https://www.linkedin.com/in/ankur-rajput56/)



rajputankur56@gmail.com



+91-9045239426



rajputankur56.github.io/

WORK EXPERIENCE

Senior Data Engineer

EXL Services

Apr 2024 - present

- Developed and maintained data ingestion pipelines utilizing **AWS Lambda, CloudFormation, Glue**, and **Redshift**, among other technologies, to onboard client data from multiple sources.
- Implemented business logic in **Snowflake** on stored proc, task, stream etc, Created role hierarchy for better access control and governance.
- Implemented features on AI/ML features like cortex analyst, Document AI etc. in Snowflake.
- Performed Disaster recovery for **Snowflake & AWS** for prod.
- Explored and completed a POC on MS-fabric.

Software Development Specialist

Amdocs Development Center India

June 2021 - Mar 2024

- Developed and integrate a **PySpark**-based next Gen Catalina report generation tool, integrating **Glue, Snowflake, Kafka**, and **RDS** to process large volumes of batch data. Streamlined data processing resulted in a 30% reduction in report generation time and significantly improved data insights.
- Performed **Spark** optimisations to decrease overall pipeline runtime on large datasets using concepts such as Partitioning, Bucketing, Broadcast Joins, Repartitioning and utilising different formats such as Parquet and AVRO to increase read-write throughput.
- Implemented Python/Spark workflows on AWS to efficiently load and analyze daily batch data updates, ensuring the timely updating of dashboards with accurate insights.
- Migrated Monolith architecture to microservices architecture for BIMS
- Worked on feature functionalities like 3DS, Chargeback, Push Notification and project patented Retain Technology.
- Conducted multiple knowledge transfer session to team.

System Engineer

Tata Consultancy Services Limited

July 2017 - June 2021

- Successfully identified anomalous vendor ratings given by SMEs by modelling an isolation forest algorithm from 2-year-old financial data.
- Worked on **Boto3** to connect to AWS services like S3, Step Function and Lambda to automate Data Migration tasks.
- Spearheaded the development of a log parsing and analytics tool using **Django, Python**, boosting client satisfaction by 28% for report generation.
- Completed project using **CNN U-Net AI** model for Brain Tumor detection.

SKILLS

- Python, SQL
- Flask, Django
- Postgres, Database Modelling
- PySpark, Snowflake, Pandas, Numpy
- AWS - EC2, S3, Lambda, Glue, RDS, MSK, ECR, Redshift, CloudFormation
- AWS - Code Commit, Code Deploy
- boto3, scikit-learn, Tensorflow
- Docker, ElasticSearch, Kafka
- Git, BitBucket

CERTIFICATION

- Apache Spark 3 - Spark Programming AWS Essentials
- Certified Scrum Master
- CutShort Certified Python - Advanced Machine Learning A-Z
- Python A-Z - Python For Data Science Deep Learning A-Z - Hands on Artificial Intelligence Neural Network
- Professional Java

PERSONAL PROJECT



Stock Analysis Tool

Tool to download historical data for tickers and provides analytics abilities to user to perform multiple kind of indicator analysis as well as self built analysis. Tool is able to provide performance of tickers and user's investment through a interactive charts. Tools is built using python, dash, plotly, and yahoo finance apis.



LinkStore

<http://linkstore.pythonanywhere.com>

User-friendly utility to save and fetch link/urls on the go, utility is designed in django using powerful tools like ORM, Pagination and javascript, Google O-Auth Authentication etc. Users can add and modify links and can directly access them from .application.

Inditeck Blog



<https://inditeck.wordpress.com>

Blog to share information related to new technology, new big things, technology news etc on wordpress. I like to explore more on technology and write about it, Inditeck is one way to do it.

Stock Json Parsing



<https://github.com/rajputankur56/parseStockData>
Json parsing utility to get stock and ticker data from multiple json files, parse and process and dump it into Database, completely configurable setup to update source and target system.

VOLUNTEER WORK, AWARDS AND INTERESTS

- Excellence Innovation Award by Amdocs
- Awarded for delivering In-house Microservice architecture session for the Amdocs unit.
- 2 times Active Learner by TCS BFSI unit
- Awarded 3 times with On-the-Spot Award
- Star Of The Month
- Awarded for delivering Python Sessions
- Promising trainee in Morgan Stanley(TCS)
- Earned Gold Badge in Python & SQL on HackerRank
- Local Guide for Google Maps

EDUCATION



Dr. Bhimrao Ambedkar University Agra

Bachelor Of Computer Application

June 2014 - June 2017

Coursework - Operating System(Linux), Data Structure, Databases, C/C++, Java, Networking etc.