ANSHUMAN DEY KIRTY

📞 312-900-4299 ✉ anshuman.dey.kirty@gmail.com in linkedin.com/in/anshuman-dey-kirty ⌽ github.com/Anshumank399

EXPERIENCE

McKinsey & Company [QuantumBlack AI]   June 2023
Senior Data Engineer | **Data Engineer**   New York, NY
- Led **data modelling & orchestration** for an oil & gas drilling project, **designing a robust pipeline to produce a model feature table** from live 1-second sensor data streams for use by an ensemble ML model. Implemented a production pipeline to minimize Non-Productive Time (NPT) days, optimizing $45M/year operations.
- **Revamped data pipeline architecture** with active learning, enhancing knowledge validation for insurance catalogs. Implemented 'Active Learning' for automated tagging of 100% concept matches, **reducing human intervention by 80%** and resolving project bottlenecks. **Managed kubernetes configurations**, debugging P0 logs, for RCA analysis.
- Transformed business logic into an efficient **data product**, focusing on financial **reporting and analytical automation** using Snowflake based DBT workflow. Significantly reduced manual calculations, reducing the process from 5 days to 15 minutes of data pipeline execution.
- Utilized Spark to design and implement data pipeline for a leading automobile client to be leveraged by the supply chain optimizing ML model, individual parts to vehicle order prioritization, handling data with over 500 million records
- Developed **API** to expose tree-based modeling for margin expansion in price discounting, enhancing profitability by 1% for an electronics company.

Data Engineering Intern   June 2022-August 2022
- Designed an algorithm to handle overlaps in **geospatial** data **improving model accuracy** by 18% and setup sustainable data pipelines for the Metal Mining client company.
- Enforced data quality checks throughout the data ETL for quick data issue resolution.

Societe Generale   January 2019-July 2021
Senior Data Engineer | **Data Engineer** | **Intern**   Bangalore, India
- Developed data pipelines to monitor bank's infrastructure logs with machine learning capabilities reducing critical incidents by 30% yoy.
- Researched potential outage using Isolation Forest based **anomaly detection**; avoided 20+ outage scenarios in 6 months by informing related skill teams about root-cause & impact before an incident.
- Collaborated with CoE to build microservice: **AI as a Service**; initiative to make data based decision accessible.
- Established **automated analytical reporting** for KPIs of risk, performance, obsolescence and vulnerabilities using **statistical analysis** for executive management; saving 10+ hours a week in report preparations.

TECHNICAL SKILLS

Languages: Python, SQL, Scala, Java, R
Specialized Experience: Data Architecture and Modeling, Distributed Data Systems, Spark, Visualization
Databases, Formats & Storage: PostgreSQL, MongoDB, Flink, Presto, Snowflake, Delta Lake, Iceberg, Hive, HDFS, S3
Cloud/Technologies/Frameworks: AWS, GCP, DBT, Docker, Kubernetes, Databricks, Airflow, Dagster

EDUCATION

University of Southern California   May 2023
Master of Science in Computer Science | GPA: 3.8/4   Los Angeles, CA

Vellore Institute of Technology   May 2019
Bachelor of Technology in Computer Science | GPA: 9/10   Vellore, India

PROJECTS

Merchant Transaction Analysis | *Python*   September 2022
- Exploratory and churn rate analysis for merchants based on **time series** transactional data.
- Understand payment activity of the merchants. Tag merchants with attributes based on the data available.
- Identifying and predicting churn.

Yelp Dataset Exploration | *Python, Scala*   January 2022
- Built ETL pipelines for Yelp dataset, analyzing 500K+ reviews for sentiment trends.

HONORS & AWARDS

Published research on Predictive Modeling for Energy Optimization in Elsevier: | *Elsevier Publications*
- Chapter 16: The trend of Energy Consumption in Developing Nations for last two decades: A case study from a statistical perspective.

Golden Ticket Innovation Winner | *Societe Generale Innovation Worldwide Contest*

Contributor list: AutoNLP @ HuggingFace | *Open-Source*