Raghuveera Narasimha

+1(680)356-8722 | raghunaras97@gmail.com | www.linkedin.com/in/raghunaras97/

Summary

Data professional with expertise in data engineering, business intelligence, and machine learning. Proven ability to design and implement end-to-end data solutions (Azure), develop and optimize ETL pipelines (ADF, Databricks, Airflow), create impactful visualizations (Tableau, Metabase), and apply ML techniques to solve business problems. Experience in the FinTech industry, focused on improving data quality and supporting data-driven decision-making.

Work Experience

Data Analyst, Fintech, Open Financial Technology Pvt. Ltd, Bengaluru, Karnataka, India **Mar 22 - Aug 23**

- Designed and implemented a Delta Lake on Azure Data Lake Storage (ADLS Gen2), enabling ACID transactions, schema enforcement, and time-travel capabilities for 1TB+ daily Payments data.
- Applied Z-ordering on high-cardinality columns to optimize data skipping, reducing query runtime by 35% for analytical workloads.
- Utilized Unity Catalog for centralized governance, enabling role-based access control for 8+ teams which improved data discovery.
- Built and optimized Delta Lake-compatible ETL pipelines on Databricks, leveraging features like auto-optimize and compaction to reduce processing time by 6 hours.
- Orchestrated data pipelines using Apache Airflow, integrating Delta Lake updates and reducing manual intervention by 40%.
- Maintained a hybrid architecture by syncing curated Delta Lake datasets to Snowflake, optimizing materialized views and star schemas to improve query performance by 20%.
- Optimized the database by creating indexes and partitioning large datasets, thereby reducing querying time for reporting and analysis, which enhanced overall data accessibility.
- Built a clustering analysis model using K-Means to segment customers, improving targeted marketing efforts and increasing engagement by 43%.
- Performed ad-hoc funnel analysis in Tableau, identifying 35% KYC-related checkout drop-offs. Utilized A/B testing to optimize KYC methods, reducing drop-offs by 25%, decreasing churn by 10%, and increasing conversion rates.
- Embedded ARIMA models in Tableau, improving revenue forecasting accuracy by 15% and enabling proactive strategy adjustments.
- Designed and maintained 50+ interactive dashboards in Tableau and Metabase, utilizing LOD calculations, parameter-driven filters, and complex joins to meet dynamic business needs. Implemented row-level security in Tableau, ensuring GDPR compliance and data privacy across 8+ user groups.

Associate Data Analyst, XLDynamics Pvt. Ltd, Mysore, Karnataka, India **Jun 21 - Mar 22**

- Streamlined loan processing and data cleansing to a U.S. mortgage company using advanced Excel functions (LOOKUPs, INDEX MATCH, XNPV), Power Query and Power Pivot reducing report time and enhancing analytical productivity.
- Developed Power BI dashboards for loan management, improving decision-making speed and operational workflow efficiency.
- Led knowledge transfer for two new hires on standardized procedures, fostering team efficiency and maintaining high accuracy in daily reporting and loan review processes.

Technical Skills & Tools

- **Cloud & Data Engineering :** Python, PySpark, Apache Airflow, AWS (Glue, EMR, S3, Redshift, Kinesis, Lambda), Azure (Data Factory, Databricks, Synapse, ADLS Gen2, Blob Storage), Docker, Kubernetes, Git, Git Actions
- **Data Analysis & Business Intelligence:** Python, R, SQL, Pandas, NumPy, Matplotlib, Seaborn, ggplot2, Looker
- **Machine Learning, Deep Learning, and Statistical Analysis:** Scikit-learn, Keras, TensorFlow
- **Databases:** MySQL,PostgreSQL, MongoDB, Cassandra, Redis, Neo4j
- **Other Tools:** Jira, Microsoft Office, Alteryx

Projects

COVID-19 Data Pipeline and Analysis with Azure Data Factory

- Automated ETL workflows using Azure Data Factory (ADF) with Data Flows and Databricks/HDInsight, reducing manual effort by 40% via triggers, parameters, and dynamic variables.
- Secured pipelines with Managed Identities/Key Vault and streamlined deployments via Azure DevOps CI/CD, improving data reliability by 30% with Azure Monitor alerts.

Airfare Prediction and Optimization with PySpark

- Utilized PySpark, MLlib, StringIndexer, OneHotEncoder, and GBTRegressor to build an airfare prediction model, achieving strong performance metrics (RMSE: 12.60, R-squared: 0.96), through rigorous data cleaning, feature engineering and selection.
- Delivered actionable insights through visualizations and reports, highlighting the key factors influencing price changes and enabling data-driven recommendations for pricing

Advance Chest Disease Detection using R-CNN | Syracuse University

- Enhanced Chest Disease detection from chest X-rays by developing a Mask R-CNN model with PyTorch, incorporating strategic data augmentation and balanced sampling for optimized dataset, thereby boosting diagnostic accuracy and model reliability.

Education

SYRACUSE UNIVERSITY, NY, USA – **M.S. in Applied Data Science**, GPA: 3.8 Aug 2023 - May 2025
VTU, NIE SOUTH, Mysore, India – **B.E. in Electrical and Electronics Engineering** Jul 2016 - Aug 2020

Certifications: Azure Data Fundamentals, Azure Data Engineer, IBM Data Science, Tableau Desktop Specialist