Dhananjaya Kumar Venkata Mittapalli

Boston, MA| +1 8578328906 | mittapalli.d@northeastern.edu | linkedin.com | github.com

EDUCATION

Northeastern University, Boston, MA Sept 2023 - Present

Master of Science - Data Science

Related Courses: Data management and processing, Supervised and Unsupervised Machine Learning, Algorithms, MLOps, LLM

PROFESSIONAL EXPERIENCE

Natixis Investment Managers Boston, MA, USA

Data Science Engineer Jan 2025 - Aug 2025

- Developed a **conversion attribution model** to evaluate marketing-driven lead performance, correlating campaign interactions with high value Sales Qualified Leads (SQLs), enabling the Sales team to prioritize profitable client segments.
- Engineered orchestration workflows using **SQL**, **Python**, and **Rundeck** to automate revenue pipeline refreshes, seamlessly integrating trade activity, AUM movements, and CRM insights into Power BI dashboards for real-time executive reporting.
- **IMPACT:** These initiatives contributed to a **15% increase** in **lead conversion rates**, **accelerated daily reporting cycles by 40%**.

SaksOFF5th Bangalore, India

Supply Chain Business Insight Engineer May 2022 - Aug 2023

- Collaborated with cross-functional teams and business stakeholders to develop the supply chain data architecture that captures end-to-end KPIs and performed **statistical analysis** and **data mining** techniques to derive actionable insights.
- Used **SQL**, **Python** and **Airflow** to streamline data procurement processes by implementing automated workflows, effectively prioritizing tasks and managing multiple projects simultaneously, which **reduced processing time by 30%.**
- **IMPACT:** Facilitated the Chief of Supply Chain Officer to launch and **execute multiple business-focused initiatives**. Reduced report generation time from **2 days to 4 hours**, enhancing operational efficiency.

TheMathCompany Bangalore, India

Data Analyst Sep 2020 - May 2022

- Spearheaded a team to **productionalize data migration solutions** with **AWS S3 staging** and automated the entire **ETL data pipeline** via **Airflow hosted on an EC2 instance**. Ensured the solutions to be both high available and scalable.
- Designed and developed Tableau and MicroStrategy dashboards to visualize marketing KPIs, enabling the marketing team to monitor performance effectively and make data-driven decisions.
- **IMPACT:** Achieved **140% increase** in **execution speed** and maintained **less than 2% downtime** for all business critical triggers. Fostered a collaborative environment through clear and concise communication between the internal team and external clients. Expanded team size, secured and managed multiple new projects due to clients' high satisfaction with work.

TECHNICAL SKILLS

Big Data	Python, R, Pandas, Scikit-learn, Apache Spark, Apache Airflow, Fivetran, DBT, Databricks
Data Science	EDA, Hypothesis Testing, Classification and Forecasting, Adobe Analytics, NLP
Microsoft Office	Excel, Powerpoint, Word, Outlook
DBMS	Snowflake, MySQL, Hadoop, MapReduce, Oracle, SaS, NoSQL, MongoDB
Cloud	AWS (EMR, EC2, S3, Lambda, SageMaker), Azure (Synapse Analytics, Data Lake, Cosmos DB)
DevOps	Docker, Kubernetes, Travis CI/CD, GIT, TensorFlow, Pytorch, JIRA
Reporting Tools	Microstrategy, Looker, Tableau, PowerBI

ACADEMIC PROJECTS

Smart City Feb 2024 - Apr 2024

Northeastern University Boston, MA

- Leveraged advanced data analytics and big data technologies (Python, Spark, Pandas) to address urban challenges in Vancouver, focusing on housing affordability, energy consumption, and public transportation.
- Used machine learning models (GBTRegressor, RandomForest, Linear Regression) to predict housing trends and forecast CO2 emissions, aiding strategic planning for housing and energy conservation.
- Created intuitive visualizations using Matplotlib and Folium to convey findings effectively, providing comprehensive reports to guide policy-making and urban development initiatives.

Blog Generation Application using LLaMA 2 and LangChain June 2024 – Present

Northeastern University Boston, MA

- Built a **Streamlit** app that generates personalized blogs using **LLaMA 2 (7B)** model, reducing manual content creation time by over **50%** for researchers and data scientists.
- Achieved **90%** accuracy in generating relevant content within specified word limits (up to **256 words**) through integration of **LangChain** for dynamic prompt engineering and **CTransformers** for efficient model inference.