Sai Koushik Gandikota

Boston, MA | (669) 222-9587| gandikota.sa@northeastern.edu | Github | Linked In | Portfolio

EDUCATION

NORTHEASTERN UNIVERSITY — Boston, MA
Master of Science in Data Analytics — Sep 2022 - May 2024
Relevant Coursework: Data Management, Data Mining, Computational Visualization — **GPA: 3.75**

VELLORE INSTITUTE OF TECHNOLOGY — Amaravati, AP
Bachelor of Technology in Computer Science Minor in Data Analytics — Aug 2018 - Jul 2022
Relevant Coursework: Objected Oriented Programming, Data Structures, Problem-Solving using Java — **GPA: 3.80**

PROFESSIONAL EXPERIENCE

RK INFO SYSTEMS | Data Engineer Intern — Jan 2022 - Jun 2022

- Developed and optimized a **highly efficient data pipeline** for customer personality analysis using **Python, DataBricks, and Azure** resulting in a 50% reduction in data processing time and allowing for real-time analysis and decision-making
- Leveraged **ad-hoc techniques using SQL** to identify key customer segments and develop highly targeted marketing campaigns, resulting in a **25% increase in Revenue** and a **40% increase in Profitability**
- Engineered **4 Interactive dashboards** and **13 KPIs** using Power BI, resulting in a 20% increase in data-driven decision-making
- Fostered a culture of data-driven decision-making by effectively communicating findings and recommendations to the stakeholders, leading to a 15% increase in successful project outcomes

TECHNICAL SKILLS

Development and Databases: MySQL, PostgreSQL, Oracle Database, MongoDB, Neo4j
Warehousing and ETL/ELT Tools: Snowflake, BigQuery, Apache Airflow, Mage, Databricks
BigData Tools: Apache Spark (PySpark, SparkSQL, MLlib)
Programming Languages: Python (Pandas, Numpy, Scikit-learn, Tensor-flow, Seaborn), R Programming
Cloud Services: Amazon Web Services(AWS), Google Cloud Platform(GCP), Azure
Data Visualization Tools: Microsoft Power-BI, Tableau, Flourish, Google Data Studio, Microsoft PowerPoint, Looker

ACADEMIC PROJECTS

Transportation Analytics: Unveiling Insights from Uber and Lyft | GCP| SQL | Looker | Mage — May 2023 - Jun 2023

- Automated data extraction, transformation, and loading processes **ETL** using **Python scripting and Mage Pipeline**, reducing manual effort by 70% and improving overall data accuracy and reliability
- Conducted **15 Ad hoc analyses** using SQL and Python, generating analysis on rider behavior, resulting in easy pattern detection
- Enhanced stakeholder access by designing and developing **3 Interactive Looker Dashboards**, presenting **13 Key Metrics**, visualizations, and drill-down capabilities for effortless data exploration and decision-making

Optimized E-Commerce Supply Chain Database | SQL | NoSQL | Northeastern University — Jan 2023 - Apr 2023

- Developed a **Normalized schema** consisting of 6 tables, each containing several fields to store relevant data
- Populated database with a **large volume of data**, including thousands of records, providing a comprehensive view of supply chain
- Conducted ad hoc analysis by using complex SQL queries to extract insights and answer specific business questions
- Implemented **4 machine learning algorithms** with an **accuracy of 98%** to predict future demand and optimize inventory levels, significantly reducing stockouts and overstocks and improving overall supply chain efficiency

Tokyo Olympics: Data Extraction and Analysis | Snowflake | Azure | Data Bricks — Jan 2023 - Apr 2023

- Designed and constructed data warehouse using Snowflake to centralize and optimize the storage and analysis of Olympic data
- Scraped Tokyo Olympics data using **AutoScrapper and Data Factory** including athlete profiles, event results, and historical data from multiple sources, to a diverse database of over **10,000 records**
- Automated data cleaning and data Transformation to structured format using Python and Pandas in DataBricks Notebook
- Conducted **10 ad-hoc** analyses leveraging Databricks and Spark SQL, enabling exploration and extraction of critical insights

Covid Classification using Chest X-Rays | Computer Vision | Deep Learning | Tensor Flow — Nov 2021 - Feb 2022

- Implemented the **classification** of COVID-19 Chest X-ray4 images as input, leveraging advanced deep-learning techniques
- Achieved remarkable **accuracy rates of 96%** by training and evaluating state-of-the-art models, including **Resnet, Mobile-Net, and Xception**, on a large dataset of over 3000 images
- Conducted thorough **data wrangling** and pre-processing to ensure high-quality input to the deep learning models
- Successfully deployed the optimized models on a **Raspberry Pi platform**, resulting in an **accuracy rate of 95%**, demonstrating the robustness and real-world applicability of the developed solution

Extra Curricular: Graduate Teaching and Research Assitant [Assisted **5 Professors** and **645 Students**], Youtuber [**Data Science**]

Behavioral Skills: Collaborative, Team Player, Problem Solver, Communication, Leadership skills, Adaptive