Manasa Addagatla

Email: maddagat@asu.edu| Phone: 6028842621 | Address: Plano, TX | LinkedIn: https://www.linkedin.com/in/manasa-addagatla/

TECHNICAL SKILLS AND CERTIFICATIONS

Programming Languages: Python, Java, SQL, C, C++, Shell scripting.
Big Data Technologies: Hadoop, AWS (S3, EC2, EMR, Redshift, Lambda, Sagemaker, Glue), HDFS, Spark, Hive, Snowflake, Databricks.
Web Development / Machine learning Frameworks: Django, PySpark, Keras, Pytorch, TensorFlow, Scikit-learn, MLlib, HTML, CSS
DBMS: MySQL, PostgreSQL, MongoDB, NoSQL
Data Analytics and Reporting Tools: PowerBI, Tableau, SAS, Seaborn, Excel (Advanced Macros).
Version Control/ DevOps: GIT, Jenkins (CI/CD), Kubernetes, Docker.
Certifications: AWS Certified Solutions Architect Associate.

WORK EXPERIENCE

SaaS Technical Writer **Jun 2022 – Dec 2023**

Enterprise Technology, Arizona State University

- Achieved 25% reduction in support ticket volume while driving self-service adoption through documentation of complex technical features by collaborating with engineering teams, enabling non-technical stakeholders to grasp product capabilities.
- Led a team of 4 technical writers in the creation of user guides, API documentation and reference guides, in-app help resources for multiple SaaS products, contributing to 75% increase in user onboarding efficiency and 30% boost in third-party partnerships.
- Contributed to sprint planning, backlog grooming, and task prioritization in alignment with Agile principles.

Technologies Used: Salesforce, ServiceNow, Dropbox, Microsoft 365, Adobe Creative Cloud, Git, Jira, Confluence, Trello, Excel, Agile.

Analytics Engineer **Dec 2020 - Nov 2021**

Amazon Development Center, Hyderabad, India

- Implemented statistical modeling and data visualization techniques in **Python for counterfeit product detection** using **SQL** for data extraction and cleaning, processed terabytes of real-time and batch data from diverse sources (CSV, Multi DSV, Parquet, Binary).
- Developed a **SQL** and Python Script-based footprint dashboard for tracking 60K+ sellers across the world, identified fraudulent sellers and counterfeit products, leading to swift auto-banning of fraudulent sellers on the Amazon platform.
- Designed and implemented a robust **data warehousing solution** using **AWS Redshift**, leveraging **ETL processes** with **AWS Glue** to extract data from source systems, transforming and loading it into Redshift, resulting in a **30% reduction in data processing time.**
- Utilized **data modeling concepts** (ER Modeling, Normalization, Snowflake Schema, Dimensional Modeling) to structure and optimize the Redshift database, ensuring efficient storage and retrieval of data for analytical purposes.
- Created client-facing dashboards in **PowerBI** using Python to pull data monitoring key metrics in near real-time, like rate of successful auto-banning of fraudulent sellers, number of identified counterfeit products, to enable timely decision-making.
- Collaborated with experienced cross-functional teams to implement a highly efficient machine learning pipeline on AWS infrastructure, ensuring data cleaning, preprocessing, and feature engineering on a diverse dataset optimized for the model.

Technologies Used: Amazon S3, EMR, AWS(Redshift, Lambda, Glue), Hadoop, PowerBI, Python, PL/SQL, Fast API, Shell script, HDFS.

Data Analyst (Supply Chain) **Apr 2018 – Nov 2020**

Virtusa Life Spaces LLC, Nellore, India

- Leveraged **Python, MySQL, Apache Spark** for real-time data analysis in the supply chain domain, utilized SQL querying techniques (subqueries, window functions) to continuously monitor, analyze data streams related to inventory, procurement, and logistics.
- Developed dynamic financial **Tableau** dashboards for real-time tracking of metrics, showcasing live inventory levels, procurement costs, logistics efficiency, and KPIs, implemented robust reporting and documentation facilitating immediate decision-making.
- Utilized **Apache Spark** for real-time processing of supply chain data, identifying and mitigating potential delays in procurement, production, and distribution processes, leading to a **12% reduction in project lead times.**
- Collaborated with project managers, employing root cause analysis and data analytics techniques to address inefficiencies, cost overruns, and disruptions in the supply chain, resulting in an **8% reduction and substantial annual cost savings**, enhanced project timelines, and improved client satisfaction.

Technologies Used: Python, MySQL, Apache Spark, Matplotlib, Tableau, Excel, Pandas, and Data Storage/Database Management.

ACADEMIC PROJECTS

Financial Performance Dashboard:

- Developed and implemented a data storage solution using **Snowflake** to house diverse financial datasets, including income statements, balance sheets, and cash flow statements sourced from multiple **IBRD Statement datasets**.
- Used **Excel advanced macros** to create visualizations, including a comprehensive dashboard showcasing a **15% improvement in profit margins,** a **20% increase in revenue trends**, and a detailed expense breakdown, facilitating data-driven decision-making.
- Conducted granular analysis of historical financial data, identifying trends and patterns, presenting findings with a potential **10% optimization in budgeting**, enhanced expense control strategies, resulting in actionable insights for strategic financial planning.

Advanced Drug Slang Detection and Sentiment Analysis:

- Processed 100,000+ drug slang terms with **Python and Regular Expressions, then built a Google BERT-based Sentiment Analysis** classification model for over 90% accuracy in drug slang sentiment detection.
- Ingested and organized diverse data sources, including transaction logs and user behavior data, into **Snowflake**.
- Created insightful data visualizations, leveraging **Matplotlib, Seaborn**, and **Plotly**, to present analysis findings.

Spatial Cluster Analysis for Regulatory Decision-Making on NYC Taxi Trip datasets:

- Identified 50 of the most statistically significant spatial clusters and outliers in spatial data (hot spots) of a 2009–2012 **NYC Taxi Trip dataset** using **advanced SQL queries** such as sub-queries & joins, Apache Spark, **Hadoop,** and spatial statistics.
- Utilized data collected from 150000 records to create a scalable distributed algorithm using Getis-Ord statistic and **Apache Spark.**

Real-Time Collaborative Image Analysis and Classification on Mobile Clients:

- Developed an Android mobile application using **Python** and **Flask** server to capture images of handwritten digits, divide images into equal parts using **OpenCV,** and classify the identified digits implementing **CNN Classification** using the **TensorFlow framework**, which trains the **deep-learning framework** using the **MNIST dataset.**

EDUCATION

Master of Science, Computer Engineering (Computer Systems) **Jan 2022 - Dec 2023**
Arizona State University, Tempe, Arizona GPA: 3.64/4.00

Bachelor of Technology, **Electronics and Communication Engineering** **Aug 2016 - Sep 2020**
Jawaharlal Nehru Technological University, Hyderabad, India CGPA: 7.59/10.00

LEADERSHIP AND VOLUNTEER EXPERIENCE

Street Cause MGIT, ***President*** **Jul 2019 - Apr 2020**

- Managed an NGO of 120 volunteers, planned and organized multiple fund-raising activities, and coordinated several events for providing help to orphans, senior citizens in old age homes and other less fortunate.

Hyderabad Youth Assembly, ***Vice-Chairperson*** **Jul 2017 - Feb 2018**

- Lead a team of 30 volunteer students from different colleges to provide solar energy solutions to residents of various villages without access to electricity, across Telangana.