SRAVAN KUMAR MANGALAGIRI

+1 315-886-7555 • [mangalagiri2407@gmail.com](mailto:mangalagiri2407@gmail.com)• [LinkedIn](http://www.linkedin.com/in/sravan-kumar-mangalagiri) • [Portfolio](https://sravan2418.github.io/sravan-kumar.github.io/)

# WORK EXPERIENCE

**Application Development Analyst,** Accenture Technology Sep 2019 – Jan 2022

* Spearheaded high-performance ETL pipeline development for AT&T, consolidating data from heterogeneous sources into a data warehouse, boosting processing efficiency by 30%, and enhancing KPI tracking and decision-making.
* Led Azure Synapse and Data Factory integration, improving ETL efficiency by 25%, ensuring data integrity, and driving key analyses on Customer Satisfaction, Risk Management, and Financial Performance.
* Optimized CI/CD pipelines on Azure, achieving 40% faster deployments and a 30% higher success rate, ensuring seamless project execution.
* Led efforts to troubleshoot and resolve PVT automation testing issues, implementing targeted fixes that improved system reliability by 20% and streamlined the deployment process.
* Developed and optimized SQL queries to extract, transform, and load data for Power BI dashboards, delivering real-time insights that empowered AT&T leadership to make informed strategic decisions.
* Developed a chatbot using Python and Shell scripting to automate daily report generation and assist with data-driven tasks, saving 4 hours per day and enhancing efficiency in creating marketing materials.
* Collaborated with cross-functional teams to address customer-related payment issues and track KPIs, improving customer satisfaction and optimizing payment processes by leveraging data insights to support AT&T leadership in strategic decision-making.
* Participated in daily team scrum meetings to review progress, plan next steps, conduct deployment reviews, sprint demos, and retrospectives, while addressing roadblocks and tracking end-to-end project status and issues using JIRA.

**Research Assistant, Syracuse University**

*Data Science Duties* Jun 2024 – Present

* Developed and deployed chatbots with Python, Streamlit, OpenAI, and Cohere, enhancing user interaction and satisfaction by 35%.
* Documented Python environments (Codespaces, Streamlit), improving setup efficiency by 40% and reducing deployment time 30%.
* Implemented LLMs and RAG systems for advanced Q&A bots, boosting answer relevance by 40% and cutting query time by 35%.
* Applied prompt engineering and NLP to enhance chatbot accuracy by 40%, leveraging vector databases for 50% faster retrieval.

*Data Engineer Duties* Jan 2023 – May 2023

* Designed data models using Erwin Data Modeler and implemented a data pipeline leveraging AWS S3 for storage, Glue Crawlers for automated schema discovery, and Amazon Redshift for optimized data storage and analysis.
* Developed real-time dashboards in QuickSight, utilizing Amazon Athena for initial data querying, visualizing user engagement and song trends, and improving reporting efficiency by 30%.

# TECHNICAL SKILLS

**Programming Languages:** Python, R, SQL, PySpark, Scala, Go, Unix, Java

**Certifications:** AWS Certified Machine Learning – Specialty, Microsoft Certified Azure AI Fundamentals (AI-900)

**Frameworks & Tools:** MLFlow,Git, Kubernetes, Agile, TensorFlow, Apache Spark, Hadoop, Kubeflow, Seaborn, Matplotlib, FASTAPI, Pandas, NumPy, SKlearn, Hive, Jenkins, Streamlit, Keras, PyTorch, BASH, Docker, GraphQL, Hadoop (MapReduce, YARN, HDFS), Apache Kafka, Apache Drill, Apache Spark, Apache Airflow, Hive

**Database & ETL/ELT:** MySQL, MongoDB, Apache Cassandra, Redis, Neo4j, PostgreSQL, Snowflake, DynamoDB, NoSQL

**Visualization Tools:** Tableau, Power BI, Looker, Kibana, Shiny, Plotly, Seaborn, ggplot2, Elastic search

**Machine Learning:** Linear/Logistic Regression, Support Vector Machines (SVM), Gradient Boosting Machines (GBM), XGBoost, Random Forest, Decision Trees, k-means clustering, CNNs, RNNs, LSTMs, Naive Bayes, KNN, NLP, Bagging, GPT, BERT, Reinforcement Learning, Hyperparameter Tuning (Grid Search, Random Search), Dimensionality Reduction Techniques (PCA, t-SNE)

**Cloud Technologies:** AWS (Redshift, DynamoDB, Glue, Hive, Neptune, Athena, QuickSight, S3, EC2, RDBMS, Lambda, Bedrock, SageMaker, CodeBuild, CodePipeline, CloudFormation, Step Functions, Managed Kafka, DevOps), Azure (Synapse Analytics, Data Lake Storage, Data Factory, Databricks, DevOps, ML Studio, Machine Learning), Google Cloud Platform (GCP), BigQuery, Snowflake, DBT Cloud

# Software Tools: GitHub, GitLab, MS Excel, MS PowerPoint, MS Word, JIRA, Linux

# EDUCATION

|  |  |
| --- | --- |
| **Syracuse University**, School of Information Studies, Syracuse, NY | May 2024 |
| **M.S. Applied Data Science** |  |
| *Relevant Coursework:* Advanced Big Data Management | Data Warehouse | Big Data Analytics | Natural Language Processing | Applied Machine Learning | Text Mining | |

**PROJECTS**

**Data Warehouse and BI Reporting for Publisher Company -** SQL | DBT| Snowflake | Azure blob Storage | Power BI

• Implemented a data warehouse using Kimball methodology, building end-to-end ELT pipelines using OLTP data sources stored in Azure Blob Storage, Minio, and S3 buckets to Snowflake for staging, with data rapidly transformed using DBT Cloud to build dimension, fact, and OBT tables, and creating interactive dashboards in Power BI, resulting in a 50% increase in operational efficiency and a 40% boost in user engagement.

**Lacrosse Game Analytics Box Score –** Pyspark | MySQL | Minio | MongoDB | Apache Drill | Docker

• Engineered the "Lacrosse Game Analytics Box Score" pipeline using PySpark to process and structure real-time game data from Minio, transforming it into a document-based format (JSON) and storing it in MongoDB for efficient retrieval and real-time rendering by Sidearm web developers. Utilized Apache Drill for querying and integrated MySQL for post-game updates to player and team statistics, resulting in a 20% increase in fan engagement.