Adarsh Kumar

College Station, TX | +1-959-326-8692 | adarsh0801@tamu.edu | LinkedIn | GitHub | Portfolio

EDUCATION

Texas A&M University   **College Station, TX**
Master of Science in Data Science (MSDS)   Aug 2024 - Present

KIIT Deemed to be University   **Bhubaneswar, India**
B.Tech. in Computer Science and Engineering   Jul 2018 - May 2022

SKILLS

Programming Languages: Java, Python, HTML/CSS, JavaScript, SQL, MongoDB
Developer Tools: Jupyter Notebook, Git, Postman, Beautiful Soup, Tableau, Azure, Databricks
Technologies / Libraries: Pandas, Numpy, SciPy, Scikit-learn, PyTorch, TensorFlow, PySpark, Snowflake, Hadoop, Docker, Transformers, Generative-AI, HuggingFace
Coursework: Data Analysis, Database Management Systems, Data Mining, AI, Natural Language Processing, LLM

WORK EXPERIENCE

Optum   **Gurgaon, India**
Machine Learning Engineer   **May 2022 - Jul 2024**

- Developed and implemented a machine learning algorithm to predict member enrolment in MassHealth programs, **achieving 80% accuracy** and increasing program churn rate by 15%.
- Collaborated with cross-functional teams to **integrate NLP model** into the existing products, reducing **manual effort by 95%** and improving overall **system efficiency by 30%.**
- Built a robust data pipeline solution to seamlessly **transfer 5TB of Salesforce backup data to Snowflake** using Azure Data Factory, ensuring business continuity during a critical server downtime period.

TDP ML Engineer   **May 2022 - Apr 2023**

- Conducted **A/B testing** to evaluate the effectiveness of redesigned internal and provider portals for MassHealth, resulting in a 45% increase in user engagement and involvement.
- Created a high-performance ETL framework with Microsoft SSIS and SQL Server, achieving **daily processing of over 100GB data**, integrated real-time analytics capabilities, enhancing response time for critical business decisions.
- Implemented complex ETL job flows to read and write from flat files and databases, resulting in a **20% increase in data accuracy**, saving $8000 annually through automation.

HighRadius Technologies   **Bhubaneswar, India**
Machine Learning Engineer Intern   **Jun 2021 - May 2022**

- Enhanced scalability and **accuracy of forecasting models by 10%,** enabling more accurate long-term predictions for business planning.
- Customized and optimized machine learning algorithms to **reduce tuning time** by **11% daily and 62% weekly**, resulting in more efficient data analysis processes.
- Researched and integrated machine learning models into web app, improving **accuracy by 3% over previous models**.

PROJECT EXPERIENCE

Rhythm Flow
Tech Stack Used: HuggingFace, Snowflake, Kafka, Apache Airflow, PySpark, Tableau, Streamlit, Generative AI
- Implemented a scalable end-to-end ETL pipeline leveraging Apache Airflow for orchestration to extract Spotify playlist data and load it into Snowflake for analytical processing.
- Built a user-friendly interface using Streamlit to visualize Spotify track data, including metadata, fine-tuned Microsoft Phi-3.5-Instruct model to generate playlists, recommend songs and summarize songs from the lyrics.

German Traffic Sign Recognition
Tech Stack Used: Python, Deep Learning, PyTorch, Self-Attention Modeling, CNN, Streamlit
- Designed and implemented a self-attention CNN model using PyTorch for the German Traffic Sign Recognition Benchmark (GTSRB) dataset achieving 98% accuracy in classifying traffic signs.

Invoice Management Web Application
Tech Stack Used: Python, Machine Learning, Java, JDBC, HTML/CSS, MySQL, Flask
- Developed a full-stack web application to predict invoice clearance dates with a 5% margin of error, incorporating role-based functionalities and secure access control for analysts and managers.