

MANOGNA TUMMANEPALLY

+1 (813) 869-1456 | manognat@usf.edu | [linkedin.com/manogna-tummanepally](https://www.linkedin.com/in/manogna-tummanepally) | PortFolio

EDUCATION

University of South Florida, Tampa

Aug 2023 - Expected May 2025

Master of Science in Business Analytics and Information Systems (BAIS)

CGPA: 3.9/4

- **Relevant Coursework:** Data Mining, Advanced Database Management Systems, BigData, Data Science Programming, Data Warehousing

EXPERIENCE

DXC Technology

Aug 2022 - July 2023

Jr. Data Scientist

- Built and deployed a predictive model for an airline company using Random Forest and Gradient Boosting, achieving 92% accuracy within 3 months, securing a \$500K project for DXC
- Optimized data pipeline efficiency by reducing data preparation time by 30% through streamlined data acquisition, cleansing, EDA, and feature engineering for airline operations
- Delivered actionable insights with SQL and K-Means Clustering, boosting airline guest satisfaction by 20%, and deployed the ML model as an Azure Function, reducing latency by 25% and ensuring 99.9% uptime for production workloads

Wiley Edge (formally Mthree)

March 2022 - May 2022

Data Engineer Intern

- Completed a 2-month Data Engineer training program (Cohort C233), gaining hands-on experience with Apache Spark, SQL, AWS, and Snowflake for big data processing, ETL development, and data transformation
- Collaborated on team-based projects to design and implement data workflows, enhancing skills in Python programming, data storage, retrieval, and technical troubleshooting

PROJECTS

Detection of Hate Speech in Memes | *TensorFlow, PyTorch, CNN, RNN, LSTM, Autoencoders*

- Developed a multimodal machine learning system by integrating CNNs for image analysis and TF-IDF with Dense Neural Networks for text, resulting in 69.41% accuracy for text-only models and 65.12% for combined models on the Facebook Hate Meme Dataset.
- Optimized model performance through feature extraction techniques like pre-trained image models and dimensionality reduction, improving computational efficiency by 25% and achieving 64.79% precision.
- Processed thousands of annotated memes in the Facebook Hate Meme Dataset to train models for robust hate speech detection, adhering to GDPR standards and ensuring scalable, ethical content moderation.

Retail Sales Insights and Forecasting – Data Warehousing Approach

- Designed and implemented a data warehouse schema with 4 dimension tables and 1 fact table, enabling efficient analysis of 10+ sales metrics, including monthly trends and cumulative sales, reducing query execution time by 30%
- Built interactive Power BI dashboards visualizing 8+ key metrics such as monthly growth, top-performing shops, and category-wise sales, improving stakeholder decision-making efficiency by 40%
- Developed 12 analytical SQL queries and a machine learning model to forecast sales trends, achieving 92% prediction accuracy, helping reduce inventory costs and minimize stockouts by 20% through data-driven inventory management

Scalable Sentiment Analysis of IMDb Movie Reviews Using PySpark

- Developed a scalable PySpark-based sentiment analysis system to process and classify 50,000 IMDb movie reviews, achieving 90.8% AUC-ROC with Logistic Regression and enhancing prediction accuracy for positive and negative sentiment trends in the entertainment industry.
- Engineered advanced NLP features, including TF-IDF vectorization, tokenization, and lemmatization, to analyze audience sentiment patterns, improving data preprocessing efficiency by 30% and providing actionable insights to optimize content strategies

DEMONSTRATED CAPABILITIES

- **Programming Languages:** Python, Java, SQL, C, C, JavaScript, Ruby
- **Database Management:** PostgreSQL, MySQL, SQL Server, MongoDB, Azure SQL Database
- **Data Engineering Tools:** Apache Spark, Hadoop, Hive, Snowflake
- **Data Science & Machine Learning:** Scikit-Learn, PyTorch, Keras, Pandas, Numpy, Matplotlib, TensorFlow, SciPy, Matplotlib, Seaborn
- **DevOps Version Control:** Docker, Kubernetes, Git, GitHub
- **Big Data & Cloud Pipelines:** ETL Pipelines, Data Migration, Workflow Automation, Data Modeling
- **Visualization Tools:** Tableau, Power BI
- **Other Technologies Skills:** Jupyter Notebook, Statistics, Critical Thinking, Problem-Solving, Communication Skills