MANOGNA TUMMANEPALLY

+1 (813) 869-1456 | manognat@usf.edu | linkedin.com/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 for Business, Data Science Programming, Data Warehousing, Adv System Analysis and Design, R Programming

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 collaborated with cross-functional teams to enhance compliance and data quality across enterprise-wide operations
- Provided insights with SQL and K-Means Clustering, increasing airline guest satisfaction by 20%, and deployed an ML model as an Azure Function, reducing latency by 25% with 99.9% uptime

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 with engineers to implement scalable data workflows using Python and Apache Spark, ensuring efficient ML pipeline development and deployment

PROJECTS

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

- Developed a multimodal ML system for hate speech detection, combining CNNs for images and TF-IDF with DNNs for text, achieving 69.41% accuracy for text-only models and 65.12% for combined models, with 25% improved efficiency and 64.79% precision
- 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 anomaly 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, 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%
- Created 12 SQL queries and a machine learning model to forecast sales trends with 92% accuracy, reducing inventory costs and stockouts by 20%

Scalable Sentiment Analysis of IMDb Movie Reviews Using PySpark

- Built a scalable PySpark sentiment analysis pipeline for 50,000 IMDb reviews, achieving 90.8% AUC-ROC with Logistic Regression, improving sentiment trend predictions in the entertainment industry
- Engineered NLP features like TF-IDF, tokenization, and lemmatization to analyze sentiment patterns, boosting preprocessing efficiency by 30% and delivering 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, Looker Data Studio
- **Other Skills & Technologies:** Data Entry, Data Manipulation, Data Quering, Data Storage and handling, Data Transformation & Pattern Identification, Data Documentation and Reporting, Jupyter Notebook, Statistics, Critical Thinking, Problem-solving