MANOGNA TUMMANEPALLY

**Boston, MA** | **Open to Relocate** | +1 (813) 869-1456 | [manogna30@gmail.com](mailto:manogna30@gmail.com) | [LinkedIn](https://www.linkedin.com/in/manogna-tummanepally) | [GitHub](https://github.com/manognat30?tab=repositories) | [Portfolio](https://manognat30.github.io/Manogna-Portfolio/)

**Education**

|  |  |  |
| --- | --- | --- |
| **University of South Florida, Tampa FL** |  | **Aug 2023 - May 2025** |
| *Master of Science, Artificial Intelligence & Business Analytics* |  | *Tampa, Florida* |
| **Jawaharlal Nehru Technological University, Hyderabad** |  | **Jul 2018 - Aug 2022** |
| *Bachelor of Technology, Information Technology* |  | *Hyderabad, India* |
| **Work Experience** | | |

**Aramark |** *Student Lead Analyst**@USF – Tampa, USA*  **Aug 2023 – May 2025**

• Played a key role in managing and cleaning operational data from **12+ dining and facility sites** at USF, using **Excel** and basic **SQL** to produce weekly reports on labor, food costs, and compliance that supported regional decision-making.

• Involved in cross-functional efforts with BI and IT teams to review and validate **Power BI dashboard inputs**, ensuring accurate integration of **23+ KPIs** used in executive reporting and operational reviews.

• Introduced **5 standardized data templates** that streamlined site-level reporting processes, **cutting down repetitive** formatting tasks

by nearly **6 hours per week** and setting the groundwork for future automation.

**DXC Technology |** *Data Scientist I – Bengaluru, India* **Oct 2021 - Jul 2023**

• Engineered and maintained **scalable data pipelines** using **Integrate.io** and **SQL**, processing **1.2M+ healthcare records** (claims, EHRs, service logs) to support **HIPAA-compliant** analytics and reporting.

• Conducted **exploratory data analysis (EDA)** using **Python** (Pandas, Seaborn) to uncover patterns in ER utilization, collaborating with care coordination teams to improve resource allocation and risk stratification.

• Built and fine-tuned **predictive models** for ER readmission, customer churn, and sales forecasting using **XGBoost**, **Random Forest**, and **Logistic Regression**, achieving 12–15% lift in AUC-ROC and recall.

• Deployed real-time models using **AWS Lambda, S3,** and **API Gateway**, enabling scalable, low-latency predictions with logging and monitoring for improved reliability and reduced infrastructure overhead.  
  
**Certifications & Trainings**  
**Microsoft Certified: Azure Data Scientist Associate June 2025 – June 2026**Certification number: BT04D3-D9B54D

**Project Experience**

**Synthetic Voice & Health Data Generation Apr 2025 - May 2025**

* Built a privacy-preserving synthetic dataset using **CTGAN**, **TVAE**, and **Gaussian Copula** on a high-dimensional healthcare and voice dataset (1100+ features, 96% missing data), validated using **KS-Statistic** and **correlation similarity**.
* Enabled **>94%** model accuracy on downstream tasks without using real patient data, earning **1st place** at the **NIH Bridge2AI Voice Hackathon** for innovation in synthetic data generation.

**Hate Speech Detection in Memes Sept 2024 – Nov 2025**

•Developed a multimodal hate speech detection pipeline by integrating **TF-IDF + SVD** text features with **CNN-based image embeddings,** using **Dense Neural Network** to improve performance across precision, recall, and AUC on Facebook Meme dataset.

• Conducted a comparative evaluation of text-only, image-only, and combined models, with the text-based neural network achieving **69.4% accuracy**, and the **multimodal model** demonstrating greater effectiveness in capturing nuanced meme-based hate speech.

**BAT: Deep Learning Methods for IDS in Wireless Networks Feb 2024 – May 2024**

* Developed BAT-MC, a deep learning IDS combining 1D CNNs, BLSTM, and attention layers, leveraged PyTorch and Tensor Board training/monitoring and demonstrated enterprise-grade intrusion classification accuracy (84.25%).
* Preprocessed and normalized over 125,000 network traffic records, applying one-hot encoding and scaling, and transformed byte-level packet data into structured tensors for model training without manual feature engineering.

**Technical Skills**

* **Programming Languages & Libraries:** Python, C++/C, Java, Pyspark, SQL, Java, PyTorch, TensorFlow, OpenCV, Scikit-learn, Pandas, Matplotlib, NumPy, Seaborn, Theano, BeautifulSoup
* **BigData Ecosystem & ETL:** Apache Spark, Airflow, Hadoop, Hive, Kafka, Snowflake, Azure Synapse, BigQuery
* **Tools:** PowerBI, Tableau, Descriptive and Inferential Statistics, Hypothesis Testing, A/B Testing, Time Series, Survival Analysis.
* **LLMs:** LangChain, LangGraph (basic knowledge), HuggingFace Transformers
* **Enterprise AI:** Agentic AI (exploratory), Knowledge Graphs (familiar), Natural Language Query Modeling
* **DataBase:** PostgreSQL, MongoDB, MySQL, OracleDB