Amaan Mansuri

New York, NY +1 (562) 209-6787 amaan.mansuri@nyu.edu amaanmansuri Amaan165

Education

New York University – Center for Data Science New York, USA
Master of Science in Data Science *Expected May 2026*

Nirma University – Institute of Technology Ahmedabad, India
Bachelor of Technology in Computer Science and Engineering *Sept 2020 - Jun 2024*

Experience

Research Assistant New York, USA
Hartley Lab, Department of Psychology, New York University *Mar 2025 – Present*

- Developed a web scraping pipeline to collect and process over **15,000** metadata entries from books and media sources, enhancing stimuli selection efficiency by **60%**.
- Designed and implemented **knowledge graph-based** models to simulate semantic relationships across age groups, boosting prediction accuracy in reasoning tasks by **25%**.
- Created visualizations and result summaries for IRB reports and manuscript drafts, improving clarity and comprehension for non-technical reviewers by **30%** (based on internal lab evaluation).

Research and Development Intern Kadi, India
Johnson Controls-Hitachi *Jan 2024 - Jun 2024*

- Developed a reliable Bluetooth communication infrastructure for Hitachi Air Conditioners, leveraging UART protocol; increased connection stability by **33%**, enabling non-line-of-sight (NLOS) communication.
- Created a user-friendly mobile application using Android Studio that served as a digital remote-control interface, mirroring traditional IR remote functionality and ensuring faster communication between devices by **40%**.
- Increased the communication range from **7m to 55m**, enhancing the functionality of the digital remote.

Research Assistant Ahmedabad, India
Centre of Excellence in Data Science, Nirma University *May 2023 - Jul 2023*

- Designed a system to segment blister packets for QR code printing, increasing information retention to **100%** by preserving details like expiration date and batch number.
- Executed a segmentation solution using **Mask RCNN** and **VIA**, resulting in a **50% decrease** in manual inspection time, utilizing hyperparameter tuning; the new system streamlined operations for the quality assurance team.

Projects

Multi-Horizon Stock Price Prediction (LSTM + FinBERT)) *Github Report*

- Built a deep learning pipeline using **LSTM networks** to forecast 1-day and 3-day prices for **26 major stocks**, integrating **FinBERT-based sentiment scores** to boost accuracy by up to **18.57%** and directional accuracy by **12%**.
- Evaluated four feature sets (price, technicals, sentiment, combined) using **RMSE**, **MAE**, **MAPE**, and $\mathbf{R^2}$; achieved best-case $\mathbf{R^2 = 0.969}$ (AAPL, 1-day); optimized training with `ReduceLROnPlateau`, `EarlyStopping`, and `TensorBoard`.

Stock Pattern Analysis (MiNeD Hackathon Winner) *Github*

- Led the development of pattern analysis system using python, leveraging **Spearman's rank coefficient** to reduce computation time by **90%** compared to the CNN models used by other teams.
- Achieved **first prize** and received accolades for outperforming all participants in the track.

Humor Detection (Machine Learning + NLP) *Github*

- Deployed an NLP-based humor classification system (BoW, TF-IDF, Word2Vec) achieving **88%** accuracy on a **200k** humor text dataset; reduced misclassification rate by **20%** after feature optimization.
- Explored and assessed NLP methods, achieving approximately **88%** accuracy, supported by detailed analysis including a confusion matrix and classification report.

Multi-Lingual Handwritten Digit Classification (Deep Learning + Image Classification) *Paper*

- Redesigned a digit recognition system that introduced an innovative multiplexer model, utilizing an acquired dataset of over **100k** images across **5** regional languages.
- Refined machine learning model training pipeline by incorporating hyperparameter optimization techniques, lifting digit recognition accuracy to **98.88%**, surpassing present benchmarks.

Skills

Languages: Python, SQL, Java, C++, C, Bash, HTML, CSS, JavaScript, Solidity.

Frameworks & Libraries: NumPy, Pandas, scikit-learn, Matplotlib, XGBoost, TensorFlow, Keras, PyTorch, Transformers, NLTK, spaCy, LangChain, Hugging Face, Django, Flask, Hadoop, Spark, Dask, Git.

Tools & Platforms: Tableau, Power BI, Docker, TensorBoard, Airflow, AWS, Unix/Linux.

Coursework

Machine Learning, Deep Learning, Big Data, Programming for Scientific Computing, Data Mining, Linear Algebra, Probability and Statistics, Database Management Systems, Computer Architecture, Operating Systems.