

Education

New Jersey Institute of Technology, Newark

Sep 2023 – Dec 2024

Master of Science in Data Science

- ° Rel. Coursework: Deep Learning, Reinforcement Learning, AI, ML, Big Data, Data Mgmt. Systems Design, Data Analytics, Statistics.

Gandhi Institute of Technology and Management, Hyderabad

Jun 2019 – May 2023

Bachelor of Technology in Computer Science and Engineering

- ° Rel. Coursework: Big Data, AI, Data Warehousing and Mining, Robotics and Automation, Cloud Computing, Database Mgt Systems.

Projects

Face Emotion Detection Using CNNs

- Built and compared **CNN architectures** (VGG-16, ResNet-50, DenseNet121, MobileNetV2) for **binary facial emotion classification**, achieving a **99.89% test accuracy**.
- Tackled **class imbalance** using **SMOTE** and pre-processed real-world data from **Google Images** via **resizing**, **normalization** and **augmentation** (random rotations, flips, cropping) to improve model generalization and reduce overfitting.
- **Tech Stack**: Python, TensorFlow, Keras ImageGenerator, OpenCV, SMOTE, Scikit-learn

Casual Inference in Healthcare

- Built and implemented an **end-to-end** casual inference pipeline to estimate **individual treatment effects** for patients using the MIMIC-III clinical databases.
- Achieved a **34% improvement** in treatment effect estimation accuracy using doubly robust methods with **PSM**, **IPW** and **CF**.
- Deployed interpretability modules (DoWhy, SHAP) for transparent decision support in medical and results against known randomized controlled trials.
- **Tech Stack**: Python, DoWhy, EconML, Scikit-learn, Pandas, Seaborn, PostgreSQL (MIMIC-III)

Predictive Modeling for Optimizing Bank Marketing Campaigns Using Machine Learning

- Built and optimized **machine learning models** (Logistic regression, SVM, XGBoost, Neural Networks) to predict **client behavior** and improve banking campaign performance.
- Achieved **85%+ AUC-ROC**, **98% accuracy**, **82% recall** on **XGBoost**, using **SHAP** and **permutation importance** for model interpretation.
- **Tech Stack**: Python, Scikit-learn, Ensemble Learning, Model Interpretability.

Stock Prices of Apple Inc. (AAPL) Using Time Series Analysis

- Implemented **Exponential Smoothing** and **LSTM models** to predict **AAPL stock prices**, achieving high forecasting accuracy.
- Predicted the **AAPL closing price** on December 12, 2023 as **\$178.869**, providing key insights for **investment strategies**.
- Utilized SQL for efficient data retrieval from large financial datasets.
- **Tech Stack**: Python, TensorFlow, Keras, Pandas, NumPy, Matplotlib, Seaborn, Statsmodels, SQL (for data retrieval)

DDoS Attack Detection Using Machine Learning

- Developed a **DDoS detection** tool using Gaussian Naive Bayes and Adaboost, **reducing computational load by 40%** and enhancing attack detection with Python, Sklearn, and PyQt.

Sentiment Analysis of Uber & Ola

- Built a multi-layer sentiment analysis model using **DFFNN**, **CNN**, and **Google Word2Vec** to **classify 3000+ tweets** about Uber and Ola, achieving a **30% improvement in accuracy** with TensorFlow/Keras.

Internship Experience

Phoenix Global, Data Analytics Engineer

May 2022 – Jun 2022

- **Led** a team of interns in driving a **20% improvement in data-driven decision-making** by implementing and optimizing **ETL processes**, enhancing **data quality** and **integration** for analytics.
- **Spearheaded** the development of **interactive dashboards** in **Tableau** and **Power BI**, improving **data visualization efficiency** by **25%**.
- Oversaw the end-to-end creation of a **live analytics dashboard** project, aligning team efforts to deliver **actionable business insights** that supported **strategic decision-making**.
- Managed **team collaboration**, ensuring smooth project execution and timely delivery of insights, while fostering a culture of **continuous learning and improvement**.

Technical Skills and Certifications

Programming Languages: Python, R, C/C++, Java, Bash; **DBMS**: MySQL, PostgreSQL, No-SQL;

OS: Linux (Ubuntu, CentOS), macOS, Windows; **Tools**: Streamlit, PowerBI, Tableau, Looker, Git, GitHub, JIRA, Trello.

Frameworks: PyTorch, TensorFlow, AutoML, Apache Spark, Hadoop, MapReduce, Seaborn, Plotly, Bokeh.

Libraries: Pandas, NumPy, PySpark, stats models, Scikit-learn, Matplotlib, Seaborn, NLTK, Keras.

Cloud Services: AWS (EC2, S3, Lambda), GCP, Azure

Activities

CYSEC GITAM, Technical Team Member

Jan 2021 – May 2023

- Contributed to **cybersecurity** initiative, applying data science and machine learning for **threat detection** and **OSINT analysis**, and **organized CTF events** to enhance skills in **penetration testing** and **threat hunting** within cybersecurity.