**Jayantha Nanduri**
Boston, MA | +1 617-992-4852 | nanduri.j@northeastern.edu | LinkedIn | GitHub

## EDUCATION

Northeastern University, Khoury College of Computer Sciences, Boston, MA. Expected May 2025
*Master of Science in Computer Science* GPA: 4.0
Teaching Assistant for Discrete Mathematics: CS 1800
Related Courses: Large Scale Data Processing, ML, DBMS, Data Mining, Algorithms, Web Development, Design Patterns

## EXPERIENCE

**Jocata Financial Advisory & Technology, Hyderabad, India** Oct 2020 - Dec 2022
*Machine Learning Engineer | Python, PostgreSQL, Pandas, Apache AirFlow, TensorFlow, Django, Docker, AWS*

- Responsible for Designing and Developing AI-driven solutions to address complex challenges in the financial sector.
- Analyzed Tax filing data of 5000 companies by performing Exploratory Data Analysis (EDA) using BigQuery and Pandas to assess the creditworthiness of bank customers.
- Fine-tuned computer vision algorithms (CV) for identity verification, and employed optical character recognition (OCR) and natural language processing (NLP) to streamline information extraction processes.
- Deployed REST APIs for ML algorithms using Docker, enhancing migration efficiency and achieving a **40%** increase in operational productivity, significantly speeding up project scalability and streamline updates.

**Jocata Financial Advisory & Technology, Hyderabad, India** Jan 2020 - Sep 2022
*Machine Learning Intern | Python, API, Docker, TensorFlow, Image Preprocessing, Tesseract, Two-shot Detectors*

- Employed a four-point transformation to rectify the perspective distortion, making the textual content more legible and easier to process for analysis and object detection.
- Used TensorFlow's object detection API to refine a Faster R-CNN model, focusing on accurate English text detection and utilizing Named Entity Recognition (NER) to extract customer names.
- Resulting in an increase in accuracy from **80% to 95%**, substantially advancing the efficiency of the KYC verification process.

## PROJECTS

**FaceMap** | GitHub Jan 2023 - May 2023
*Apache Airflow, Gaussian Blurring, Object Detection, Nearest Neighbors, Tree-Based algorithm, GANs*

- Created a dataset using the public LFW celebrity dataset and then generated synthetic images using a pre-trained GAN to mimic the vast customer database of a Bank.
- Engineered a continuous integration (CI) pipeline for efficient image preprocessing, incorporating the Sobel filter to detect sharp pixel intensity changes.
- Developed a tree-based algorithm for identifying nearest neighbours within the database, facilitating the detection of potential duplicate accounts.
- Achieved a significant enhancement in query performance, reducing search times by **70%**.

GSTDNA Nov 2021 - Dec 2022
*Python, Statistics, Credit scoring, Pandas, Regression, Sales Forecasting, Risk assessment*

- Implemented data-driven strategies and continuous monitoring to manage loan exposures and collections, using historical tax data to enhance creditworthiness assessments of bank loan applicants.
- Formulated detailed statistical indices for assessing customer behaviour, designed to ensure that recent positive data does not disproportionately influence the overall credit profile assessment of an entity.
- Reduced false positive defaults by 30%, significantly improving the bank's ability to precisely detect and manage default risks.

## SKILLS

**Machine Learning Algorithms** - LLM, Collaborative Filtering, CNN, GAN, Transformers, SVM, Gradient Boosting.
**Programming Languages**: Python, Java, JavaScript, C, C++, HTML, CSS, MatLab
**Technologies**: Scikit-learn, TensorFlow, PyTorch, TensorRT, Flask, Django, Docker, Git, AirFlows, Kafka, Spark, Hadoop
**Databases**: SQL, MySQL, PostgreSQL, MongoDB, Hive