

Jayantha Nanduri

+1 (617) 992 4852 [Email](#) [LinkedIn](#) [Portfolio](#) [Github](#)

PROFESSIONAL SUMMARY

Data Scientist with **2+ years** of expertise in the **Banking and Financial Technology** sectors. Skilled in developing scalable, efficient cloud-based data pipelines that propel data-driven projects using **Python, PySpark, SQL, and AWS**.

EDUCATION

Northeastern University | Master of Science in Computer Science Jan 2023 - May 2025
Courses - Data Mining Techniques, Large Scale Parallel Data Processing, DBMS, Machine Learning

TECHNICAL SKILLS

Languages	Python, Java, R, C++
Frameworks	Spark, Django, PyTorch, Airflows, Statistics
Development & Database	PostgreSQL, MySQL, MongoDB
Cloud & DevOps	Docker, Git, AWS, RabbitMQ

EXPERIENCE

Khoury College of Computer Science | Graduate Teaching Assistant Aug 2023 - Present
Python, Java Boston, MA

- Leveraged expertise in algorithms and data structures, specializing in Java and Python, to enhance student understanding and academic performance in the Algorithms course.

Jocata Financial Advisory & Technology | Data Scientist Oct 2020 - Dec 2022
Python, Airflows, SQL, AWS Hyderabad, India

- Automated data pipelines by implementing event-triggered workflows for efficient data transformation using **Airflows and AWS**, leading to a **10% increase** in pipeline efficiency and automation.
- Conducted Exploratory Data Analysis (EDA) on 10TB of historical tax data using **Pandas and AWS Athena** to create a comprehensive set of proprietary indices, evaluating the financial health and risk profiles of entities.
- Trained an ensemble of **Random Forest, XGBoost, and LightGBM**, improving default prediction **accuracy by 40%** through aggregate risk scoring, offering in-depth insights into financial risk and stability.
- Boosted **deployment efficiency by 25%** through SonarQube integration and Git for code management, optimizing AWS EC2 deployments.

Jocata Financial Advisory & Technology | Data Science Intern Jan 2020 - Oct 2020
Python, PyTorch, OpenCV, OCR, AWS Hyderabad, India

- Reduced **Storage Footprint by 40%** through Efficient Image Preprocessing using AWS Lambda and AWS Batch to perform real-time image resizing, compression, and format conversion.
- Developed and integrated a custom Faster R-CNN model into the image preprocessing pipeline, accurately identifying and classifying languages in images and improving text extraction **accuracy by 20%**.
- Integrated the extracted insights into interactive dashboards and business applications using Amazon QuickSight and Tableau, empowering data-driven decision-making.

PROJECTS

Early Sepsis Detection | *Python, Docker, MLFlow, PyTorch, AWS* | [Github](#)

- Developed LSTM model using MLFlows for the early prediction of sepsis, leveraging clinical data from over 40,000 patients, achieving early detection with high accuracy - **94%**.
- Implemented a robust MLOps framework to streamline the deployment, monitoring, and management of machine learning models in a clinical setting.

Distributed Graph Analysis | *PageRank, Python, Java, PySpark, Map Reduce, AWS* | [Github](#)

- Implemented PageRank in both MapReduce and Spark for analyzing extensive web graphs, **achieving a 40% reduction in processing time** and improving both scalability and fault tolerance.
- Employed lineage information to enhance debugging, performance optimization, and result validation, improving **accuracy of outputs by 20%**.