**AISHWARYA VANTIPULI**

 vantipuli.a@northeastern.edu |  | **** | [Portfolio](https://aishuvenkat09.github.io/)

# EDUCATION

 **Northeastern University**, **Boston, MA** Jan 2019 – May 2021

*Master of Science in Data Science* **GPA -** 3.6

* **Related Courses:** Supervised & Unsupervised Machine Learning, Natural Language Processing, Algorithms.

 **Jawaharlal Nehru Technological University, Hyderabad, India** Sep 2013 – May 2017

*Bachelor of Technology, Information Technology* **GPA -** 3.6

* **Related Courses:** Data Structures, Artificial Intelligence, Information Retrieval Systems, Databases.

# TECHNICAL KNOWLEDGE

**Languages:** Python, C, Java, HTML, CSS, XML

**IDE/Tools:** Alteryx, Airflow, Tableau, AWS SageMaker, EC2, S3, Glue.

**Libraries/Packages:** NumPy, Pandas, Scikit-learn, Matplotlib, TensorFlow, PyTorch

**Databases:** MYSQL, SQL Server, Oracle, MongoDB, Google BigQuery

# WORK EXPERIENCE

 **Cincinnati Children’s Medical Research Center, Ohio**  Jan 2022 – Present

*AI/ML Research Scientist*

* Developed an automated bone age interpretation tool that predicts bone age on more than 20,000 trauma hand radiographs and achieved **MAE** of **±6** months using **NVIDIA** **MONAI** deep learning framework.
* Streamlined **Multi-GPU** training pipelines to build different experiments for finding the optimal network architecture and hyperparameters using data versioning tools like **DVC**, **Weights and Biases** & **Pachyderm**.
* Published and presented a poster on A Microservices based open-source **MLOPs** Stack for Medical Imaging Research at conference **SIIM 2022**-Society for Imaging Informatics in Medicine.

**SAP America, Pennsylvania ** Sep 2020 – Jan 2021

*Intelligent Data Management, Intern*

* Leveraged predictive forecasting using **ARIMA** models for predicting future travel expenses of SAP employees over the next quarter which led to better budget planning and **30%** cost savings.
* Conducted Data blending, Data preparation using **Alteryx** and **SQL** for publishing data sources to **Tableau**.
* Utilized **Apache Spark** with Python to develop Big Data Analytics and Machine learning applications.

**Younify, Hyderabad, India ** May 2017 – Dec 2018

*Data Scientist*

* Architected and developed scalable **A/B** testing solutions for regression models that run on **AWS**. Optimized the predictive capability of modelswhich led to substantial increase in profits up to **60%**
* Collaborated with data engineering teams to design and optimize data pipelines, ensuring efficient data extraction, transformation, and loading (**ETL**) processes.
* Integrated **Apache Airflow** with AWS to monitor multi-stage ML workflows to automate data ingestion pipeline for multiple data sources.

# ACADEMIC PROJECTS

## Text to Image Generation

### Skillset : Computer Vision, Natural Language Processing (NLP), Generative adversarial networks (GANs).

* Developed a two-stage generative adversarial network for generating photo realistic images from textual descriptions on (CUB) Dataset by training Generator & Discriminator in a Min-Max game.
* Measured performance of the model using Inception score, KL divergence loss along with human evaluation. Honored with best capstone project in class award among 20 groups.

## Plagiarism Detection

### Skillset : NLP, Text Analytics, K-Means Clustering, FP-Growth, PCA, TSNE.

* Examined 100K documents which contains textual answers given by students with different levels of plagiarism. Implemented Clustering & FP-Growth Algorithm to detect target students with 95% precision.
* Formulated a new strategy to increase the model efficiency by using Dimensionality Reduction Methods like PCA, TSNE and calculated containment and longest common subsequence (LCS) scores.

## Diabetic Retinopathy Detection

### Skillset : Image Classification, Convolutional Neural Networks, Multi-Layer Perceptron, Transfer Learning.

* Devised an automatic DR grading system capable of classifying images based on disease pathologies from four severity levels using Image Classification.
* Compared performances of Logistic Regression, CNN, KNN, MLP and Inception v3 on High Resolution Retina Images and achieved an accuracy of 93% on CNN.