Jay Kamleshkumar Madhu

Email: jay.madhu101@live.com | **Phone:** 812-360-0858 | **LinkedIn:** linkedin.com/in/jay-madhu/ | **GitHub:** github.com/jaym096

EDUCATION

Master of Science in Data Science | Indiana University Bloomington **|** GPA: 3.81 / 4.0 **Aug 2019 – May 2021**
Relevant Coursework: Machine Learning, Big Data, Statistics, Deep Learning, Exploratory Data Analysis.

Bachelor of Technology in Computer Engineering | Ganpat University, India **|** GPA: 8.36 / 10 **Aug 2014 – May 2018**

Online Certifications (MOOCs)
Statistical Learning from **Stanford University** [link] **January 2020**
Computer Vision Nanodegree from **Udacity** [link] **August 2020**

PROFESSIONAL EXPERIENCES

Graduate Student Researcher – Kelley School of Business, IU - Bloomington, IN **Dec 2020 – Present**
- Actively working under Prof. Sagar Samtani on a project related to Cyber Threat Intelligence which involves performing temporal analysis to investigate how vulnerabilities propagate in open-source code available on GitHub, and devise a technique using Deep Learning to detect such vulnerabilities at an early stage and prevent exploitation of the code.

Graduate Teaching Assistant - Indiana University - Bloomington, IN **Aug 2020 – Dec 2020**
- Mentored 45 undergraduate students for the course of **Object-Oriented Software methods.**
- Conducted office hours and doubt sessions to reinforce learning concepts and improve their performance throughout the coursework.
- Graded their assignments and guided them in the final course project.

Research Intern – Space Applications Center, ISRO – Ahmedabad, India **Jan 2018 – May 2018**
- Independently worked on implementing a python-based pipeline to generate thermodynamic diagrams using multi-dimensional satellite data, integrating it with a web-based data visualization platform, providing real-time accessibility of the diagrams to the scientists.
- Collaborated with other engineers to contribute to MOPy, an in-house pythonic library developed for meteorological and Oceanographic data, by writing code routines to manipulate, analyze, and visualize satellite data.
- Performed exploratory data analysis to find the trends in the atmospheric stability indices over the Indian region and build a time series (AR) model to predict the values of the indices and studied their feasibility to forecast severe weather occurrence.

Software Development Intern – Zenn Systems – Ahmedabad, India **July 2017 – Oct 2017**
- Developed a desktop application in C# for industrial automation to handle serial data from 6 packaging machines, to log the data, and simultaneously upload it on the client's application server, generating local reports in excel for accounting and inventory purposes.
- Improved the performance of the application by using multithreading to provide near real-time updates of the inventory, eliminating the laborious process of manually logging details of over 50,000 packaged products per day.

TECHNICAL SKILLS

Languages: C++, Java, Python, R, SQL, HTML, CSS.
Domain Knowledge: Statistics, Machine Learning, Deep Learning, Graphical Models, Computer Vision, NLP, Android Development.
Libraries: Scikit-Learn, NumPy, Pandas, OpenCV, NLTK, PyTorch, matplotlib, Seaborn.
Tools and Frameworks: AWS EC2, GCP, MongoDB, MySQL, SQLite, Git, Flask, Jupyter Notebook, Linux, Apache Spark, Android Studio.

KEY PROJECTS

Missing Value Imputation *[Python, PyTorch, matplotlib, GCP]*
- Implemented a conditional GAN to predict surrounding pixels of an input image from the MNIST dataset, given the center patch of the image. Applied regularization to control the prediction of the border pixels only.
- Trained 3 type of architectures: Baseline, regularization with λ = 0.1 and λ = 10 and compared their performance.

Network Compression Using SVD *[Python, PyTorch, GCP]*
- Compressed model parameters of a deep neural network by more than 70% using SVD without compromising on the results.
- The method significantly improved the space-complexity of the model stored in the memory leading to fast computation.
- Compared model performance with respect to the no. of parameters required by the network to understand the effect of compression.

Automatic Image Captioning *[Python, PyTorch, OpenCV, NLTK]*
- Used a pre-trained ResNet50 architecture as the encoder and RNN-LSTM as a decoder, to recognize the context of an input image and generate human-readable description of the content of the image. Train the model on the MS-COCO dataset.

Audio Denoising *[Python, PyTorch, librosa]*
- Implemented various neural network architecture such as Fully Connected Network, 1D CNN and RNN-LSTM to learn the mask to remove noises from the audio files.

Parts of Speech Tagging *[Python]*
- Predicted POS of sentences by using 3 types of Bayesian Nets and solving them using HMM and MCMC achieving ~94% word accuracy.

Self-Supervised Learning Using Pretext Tasks *[Python, PyTorch, GCP]*
- Trained a CNN model on CIFAR-10 dataset using self-supervised learning by first learning visual representations and then using those representations to train the model using transfer learning. Compared the performance of the model with a baseline model.

POSITIONS OF RESPONSIBILITY

- **Head Coordinator:** Led a team of 40 students to organize techno-cultural events at Convergence 2017, Technical Fest of UVPCE.
- **President and Co-founder:** Computer Science club at Ganpat University (2015-2017)
- **Captain**: Computer Engineering Department Football team (2014-2015).