

Jay Kamleshkumar Madhu

Email: jay.madhu101@live.com | **Phone:** 812-360-0858 | **LinkedIn:** linkedin.com/in/jay-madhu/ | **Website:** jaym096.github.io
GitHub: github.com/jaym096 | **Location:** Bloomington, IN

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**
Statistical Learning from **Stanford University** [\[link\]](#) **January 2020**

PROFESSIONAL EXPERIENCES

- Graduate Teaching Assistant - Indiana University - Bloomington, IN** **Aug 2020 – Dec 2020**
 - Course: Object-Oriented Software Methods
 - Connected with students to help and motivate them throughout the coursework and grade their assignments. Conducted office hours and doubt sessions to reinforce learning concepts and improve their performance.
- 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 the data from INSAT-3D satellite, 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.
 - Adopted best coding practices to improve application performance 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

Programming Languages: C#, C++, Java, Python, R.

Machine Learning: Regression techniques, K-means, SVM, naïve-bayes, Tree-based methods, Neural Network, Gradient Boosting, LDA, PCA, Gibbs Sampling, Metropolis-Hasting, HMM, Deep Learning.

Statistics: A/B test, Chi-Squared test, T-test, ANOVA, Time Series Analysis, Gradient Descent, Newton's Method.

Libraries: Scikit-Learn, NumPy, Pandas, OpenCV, NLTK, PyTorch, matplotlib, Seaborn, tidyverse, ggplot.

Tools and Frameworks: AWS EC2, MongoDB, SparkSQL, PySpark, MySQL, SQLite, Git, Flask, Tableau, Jupyter Notebook.

KEY PROJECTS

- Credit Fraud Detection using Machine Learning** *[Python, Scikit-Learn, matplotlib]*
 - Applied Machine Learning algorithms like Random Forest classifier, Logistic Regression and Decision Tree to detect fraudulent behavior on 284,807 transactions. Achieved f1- Score of 0.90, 0.74 and 0.81 for the respective models.
- LDA for topic modelling** *[Python]*
 - Used Latent Dirichlet Allocation with Gibbs Sampling for inferencing topic distribution within the corpus and word distribution per topic
 - Evaluate the dimensionality reduction capability of LDA to support document classification and compared it with the bag of words representation. Achieved performance score of ~0.95.
- NYC Taxi Analysis** *[Python, Pandas, SQLite, Jupyter Notebook, XGBoost]*
 - Predicted the total taxi fare by analyzing over 1 million past trips using machine learning technique of XGBoost.
 - Implemented an ETL which extracts data from CSV files using pandas, performs data preprocessing and load into SQLite Database.
 - Performed feature engineering tasks to determine the key factors involved in predicting the fare.
 - Evaluated and assessed the model on the test data using RMSE as a performance metric, achieving an RMSE of ~0.5 with 95% accuracy.
- Application of AI on Real World Problems** *[Python]*
 - Predicted POS of sentences by creating 3 Bayesian Nets and solving them using HMM and MCMC achieving ~94% word accuracy.
 - Applied Viterbi algorithm to find horizons from input images of scenic landscapes.
 - Implemented text decryption using Metropolis-Hasting algorithm.
- Image Orientation Classification** *[Python]*
 - Implemented machine learning techniques like KNN and Neural Nets from scratch to predict Image Orientation (0°, 90°, 180°, 270°) on dataset of 40,000 images and compared their performances. Achieved an accuracy of 88% for the Neural Network.

POSITIONS OF RESPONSIBILITY

- Head Coordinator:** Lead a team of 40 students to organize techno-cultural events at Convergence 2017, Technical Fest of UVPCE.
- Microsoft Student Partner (2015-2017).
- Captain of the Computer Engineering Department Football team (2014-2015).