410 Riverside Ct
Santa Clara CA - 95054
Phone No: +1-857-399-9024

SAI NIKHIL THIRANDAS

saint.math.163@gmail.com
GitHub | Coursera | Brilliant.org

A mathematician and a Software Engineer, who has passion for innovation and loves solving complex challenges. I am interested in applications of Machine Learning/Deep Learning in fields of Computer Vision/NLP/Recommendation Systems.

EDUCATION

- M.S. in Applied Mathematics (Machine Learning); Northeastern University, Boston; Fall 2020 – Present; GPA: **3.91/4.0**
- MOOC: MLOps, Deep Learning, Mathematics for Data Science, Data Structures, Algorithms, Object Oriented Design.
- Coursework: Machine Learning, Computer Vision, Applied Statistics, Linear Algebra, Probability, Mathematical Modeling.

TECHNICAL SKILLS

- Python, Java, R, C/C++, MATLAB, Mathematica, SQL, PHP, Perl, HTML, CSS, TypeScript, XML, JSON, Visual Basic.
- PyTorch, TensorFlow, OpenCV, NumPy, pandas, Matplotlib, scikit-learn, SymPy, Spark, Hadoop, Kafka, Hive, Zookeeper.
- Git, Jupyter Notebook, Linux, Docker, PyCharm, IntelliJ IDEA, AWS (SageMaker), GCP, Elasticsearch, Angular, Spring, Junit.
- Regression, Classification, Ranking, Recommendation Systems, Clustering, Dimensionality Reduction, Bagging, Boosting, Feature Engineering, Neural Networks, Deep Learning, Computer Vision, Natural Language Processing, Optical Character Recognition.

EMPLOYMENT

Machine Learning Engineer — **Waterline Data** — **Dec 2018 – Aug 2020**

- Built a first ever unstructured data processor in Lumada Data Catalog (LDC) to extract text from images using OCR & image processing techniques and achieved over 85 % accuracy. Removed data drift and improved accuracy to 95 %.
- Developed a customer service AI chatbot using Elasticsearch and NLP to resolve user queries by redirecting to FAQs. Reduced number of tickets logged by 60 % and improved retention rate of people using LDC by 50 %.
- Optimized duplicate row detection algorithm using probabilistic approach; reduced time complexity from $O(n^2)$ to $O(n)$.
- Implemented an asynchronous Spark job that helps to sync new/purge ghost content metadata in LDC.

Teaching Assistant/Data Club Leader — **Northeastern University** — **Sep 2020 – Present**

- Courses: Calculus 2 (Multivariate and Vector Calculus), Matrix Methods in Data Analysis and Machine Learning.
- Led the Data Club – Spring 2022 at Northeastern. Taught around 100 common interview problems in all levels of difficulty.

PROJECTS

- Brain CT Hemorrhage Classification & Segmentation Performed binary classification using Xception Net to classify brain CT scan slices achieved an F-Score of 0.76. Used class weighting to account for imbalance and improved F-score to 0.82. Applied Bayesian Hyperparameter Optimization to reduce training time by 70 %. Performed semantic segmentation using U-Net and achieved an IoU of 0.66. Leveraged multiple shades of CT scans and 3D convolutions to improve IoU to 0.71.
- Auto Colorization of Grayscale Images Implemented Zhang et al., 2016 paper for automatic colorization of grayscale images using CNN and Deep Learning techniques. Modified the complexity of CNN architecture to achieve similar colorizations but with 80 % less training time. Achieved 77 % accuracy using a nearest neighbor based approximate image similarity measure.
- Matrix Factorization for User Rating Predictions Derived update rules and implemented Weighted Alternating Least Squares for predicting missing user ratings of MovieLens data. Improved MSE by 62 % compared to baseline (mean predicting) model.
- Data Modeling using Markov Chain Performed Time Series Analysis of average runs of opening batters in baseball from 1871 – 2015 with a Markov Chain. Performed autocorrelation and GoF test at 5 % significance level to determine valid states of chain.
- Image classifier for the SVHN dataset Built a CNN classifier model with 3 convolutional layers and 2 fully connected layers for digit recognition on street view house number images. Applied MaxPooling, BatchNormalization, Dropout and Early Stopping callback techniques to increase the validation accuracy on baseline from 55 % to 89.55 %.
- Debiasing Word Vectors Used 50-dimensional GloVe vectors to represent words. Performed Word Analogy task and implemented equalization algorithm presented in Boliukbasi et al., 2016 to remove gender bias.
- Video to PDF Developed a productivity web-application to covert video to a PDF with embedded subtitles that can help save user time up to 90 % time when revising lecture concepts. Utilized the image processing techniques like gray scaling, binarization, SIFT descriptor to compare similarity between successive frames and PIL library to embed subtitles.
- Northeastern News Updater Developed a Google Chrome extension to get instant notification updates from NEWS @ Northeastern portal using JavaScript, AJAX, HTML, and CSS. Was awarded a merit scholarship of $ 25,000.

EXTRA ACADEMIC ACTIVITIES

- Ranked in the top 10 in a CodeSprint (an algorithm competition) on **HackerRank** and won a 1 TB HDD.
- Contributed to an open-source organization named **SymPy** during Google Summer of Code application process.