|  |  |  |
| --- | --- | --- |
| 2595 Washington St  Boston MA - 02119  Phone No: [+1-617-818-4646](tel://+1-617-818-4646/) | **Sai Nikhil Thirandas** | [saint.math.1729@gmail.com](mailto:saint.math.1729@gmail.com)  [Portfolio](https://saint1729.me/) | [LinkedIn](https://www.linkedin.com/in/saint1729/) | [Coursera](https://www.coursera.org/user/cec18b9a1998670a3cc03fcd51a1be8a)  [GitHub](https://github.com/saint1729) | [HackerRank](https://www.hackerrank.com/saint1729) | [Brilliant](https://brilliant.org/profile/sai-nikhil-d213k0/) |

A mathematician and a Software Engineer with 7.5 years of work experience, who has passion for innovation and loves solving complex challenges. My career goal is to work as a Machine Learning Engineer.

**Education**

* M.S. in Applied Mathematics (Machine Learning); Northeastern University; Fall 2020 – Present; GPA: **3.91/4.0**
* Indian Institute of Technology, Kharagpur – B.Tech. (Hons.) in Civil Engineering; Fall 2009 – Spring 2013.
* [MOOC](https://saint1729.me/blog/coursera-certifications/): MLOps, Deep Learning, Mathematics for Data Science, Data Structures, Algorithms, Object Oriented Design.
* [Coursework](https://saint1729.me/transcript.pdf): Machine Learning, Computer Vision, Applied Statistics, Linear Algebra, Probability, Mathematical Modeling.

**Employment**

|  |  |  |
| --- | --- | --- |
| **Senior Software Engineer – Data** | **Waterline Data Science** | **Dec 2018 – Aug 2020** |
| * Built a first ever unstructured Data Processor to extract text from images in Lumada Data Catalog (LDC). Developed a novel way to collect training data for improving performance of existing OCR engine by 10 %; acquired at least 2 new clients. * Developed an NLP model to resolve user queries by redirecting to FAQs; reducing no. of tickets logged by 60 %. This improved user retention rate of LDC by 150 %; measured with an efficient user visit logging system. * Integrated the charting framework ngx-charts into LDC and designed multiple UI components useful for DataOps. * Fixed several security issues around field tag association, job and role management of LDC. * Implemented an asynchronous Spark job that helps to sync new/purge ghost content metadata in LDC. | | |
|  | | |
| **Software Engineer 2 – Full Stack** | **Oracle** | **Aug 2015 – Nov 2018** |
| * Developed RESTful web services for File-Based Data Import (FBDI Oracle) using Java, Spring, Angular, Oracle DB, Docker. * Optimized duplicate row detection algorithm using probabilistic approach; reduced time complexity from O(n2) to O(n). * Designed a mobile application to scan OMR sheets of candidates during hiring; reduced time spent in recruitment by 80 %. * Expertise in Oracle Business Intelligence Enterprise Edition (OBIEE), BI Publisher (XML Publisher), Accessibility Evaluation. | | |
|  | | |
| **Software Engineer – R & D** | **Altair Engineering** | **May 2013 – Aug 2015** |
| * Developed a GUI Automation Software by cloning Sikuli. Implemented standard image processing algorithms including Laplace Edge Detection, Pyramid Template Matching, Alpha blending, SIFT descriptor. * Adapted Tesseract OCR’s code, to increase accuracy in text-recognition for screen fonts from 50 % to 95 %. | | |

**Projects**

|  |
| --- |
| * [Matrix Factorization for User Rating Predictions](https://saint1729.me/CS6140_assignment4.pdf#page=27) 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](https://saint1729.me/MATH7241.pdf) Performed Time Series Analysis of average runs of opening batters in baseball from 1871 – 2015 with a Markov Chain. Calculated autocorrelation between original time series and a simulated time series. Performed GoF test at 5 % significance level to determine valid states of Markov Chain in a two-step transition matrix. * [FaceNet – Face Recognition](https://github.com/saint1729/coursera/blob/master/deep-learning/convoutional-neural-networks/week4/Face_Recognition.ipynb) Encoded face image into 128-dimension feature vector (one-shot learning) using FaceNet. Implemented Triplet Loss function to compare Anchor, Positive, and Negative images in training data. Performed face verification and face recognition using the above encodings. * [Debiasing Word Vectors](https://github.com/saint1729/coursera/blob/master/deep-learning/nlp-sequence-models/week2/Operations_on_word_vectors_v2a.ipynb) Used 50-dimensional GloVe vectors to represent words. Performed Word Analogy task and implemented equalization algorithm presented in [Boliukbasi et al., 2016](https://arxiv.org/abs/1607.06520) to remove gender bias. * [Northeastern News Updater](https://github.com/saint1729/news-at-northeastern) 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. |

**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, Angular, Spring, JUnit, Mockito * Git, Jupyter Notebook, Linux, Docker, Kafka, Hadoop, Hive, Zookeeper, Elasticsearch, PyCharm, IntelliJ IDEA, Oracle BI Publisher * Regression, Classification, Ranking, Clustering, Dimensionality Reduction, Bagging, Boosting, Feature Engineering, Neural Networks, Deep Learning, Computer Vision, Natural Language Processing, Optical Character Recognition, Template Matching. |

**Extra Academic Activities**

* Led the Data Club – Spring 2022 at Northeastern. Taught around 100 common interview problems in all levels of difficulty.
* Teaching assistant for Matrix Methods in Data Analysis and Machine Learning (Fall 2021), Calculus 2 (Spring 2021).
* Contributed to an open-source organization named **SymPy** during GSoC-2012 application.
* Ranked in the top 10 in a CodeSprint (an algorithm competition) on **HackerRank** and won a 1 TB HDD.