

## EDUCATIONAL QUALIFICATION

**Integrated Masters in Mathematics and Scientific Computing from IIT Kanpur**

(2006-2011)

## WORK EXPERIENCE

**Data Scientist, Agara Labs**

(Aug'17 - Mar'18)

### *Text Classification*

- Classification of emails across two large sets of labels (multi-label classification), with classifiers based on n-gram CNN, LSTMs, and also multitask models involving shared layers for the two classification problems. Introduced further branches within a multitask model based on segmentation of data.
- Developed and maintained the models and explored scope for expansion over segments.

### *Response Generation*

- Natural Language Generation based on Language Modeling for Response Generation. The performance was better (than ones trained on wiki-data) due to a smaller and closed vocabulary but not relevant.
- K - nearest neighbour approach when considering the embedding vectors of the emails that were extracted from layers of the classification model.
- This work was presented at **WAIU-2017 at IISc Bangalore**.

**Data Scientist, Metripping**

(Jan'17 - Jul'17)

### *Destination Rankings*

- As a part of travel search and inspiration: based on user inputs, destinations are recommended as an ordered list. This involved data collection across sources and data mining
- Used various machine learning algorithms to cluster destinations (using clustering algorithms like DB- Scan, KMeans along with methods from computational geometry). Similar destinations/experiences were determined based on user generated text data.
- Use of image processing over satellite imagery to develop and quantize features pertaining to a destination. Accounting for imagery differences

**R&D Engineer, Synopsys**

(Feb'14 - Oct'16)

- Optimised spatial configuration of design components to achieve compact designs and reduced design violations. Added support for design rules pertaining to fabrication of wiring on microchips.
- Developed an optimisation solver for non-convex problems using geometric programming after approximate formulation of the problems to one involving posynomial constraints and objective(s).

**Analyst, Genpact & IIFL**

(Jun'11 - Dec'13)

- Design, Analysis and Implementation of strategies for high frequency algorithmic trading.
- Predictive modeling: Modeled for bankruptcies/charge-offs across demographics and helped collections team decide their further actions..

## PROJECTS (SELF MOTIVATED)

### *Newsfeed classifier : TOI Vs The Hindu*

(2016)

- Scraped facebook pages of two leading newspaper publications in India.
- Surveyed various classifiers to compare the accuracies. Random Forest scored highest with 84% accuracy over a test sample. Unsupervised learning: clustering into two clusters yielded an accuracy of over 60%.

## TECHNICAL SKILLS

- **Programming:** Python, Keras, Tensorflow
- **Skills:** Deep-Machine Learning, Natural Language Processing, Neural Networks, Data Mining, Optimisation

## RELEVANT COURSES

- Data Structures and Algorithms, Linear Algebra, Probability and Statistics, Machine Learning, Convolutional Neural Networks