Shubhangi Ghosh | Curriculum Vitae

shubhangisghosh@gmail.com • shubhangighosh.github.io

Education

- **Master's of Science** **Sep 2019 –**
  Data Science – ETH Zurich
- **Bachelor of Technology** **Jul 2015–Jul 2019**
  Electrical Engineering – Indian Institute of Technology Madras
  CGPA: 9.14/10.0
  Research Interests: Machine Learning, Deep Learning, Natural Language Processing
  Relevant Courses: Machine Learning, Deep Learning, Natural Language Processing, Convex Optimisation, Probability and Stochastic Processes, Linear Algebra, Data Structures and Algorithms, Information Theory and Coding Theory

Research Experience

Research Projects

Extreme classification on graph data **Jul 2019 – Present**

- *Indian Institute of Science*
  Guide: Prof. Partha Talukdar
  - Posed the problem of multi-label classification with extreme number of labels (in the order of millions) as a semi-supervised learning problem to support the prediction of rare/missing labels with connected node information. The semi-supervised learning task was modelled using Graph Convolutional Networks on a provided graph over instances.
  - Constructed a label cooccurance graph and learnt label embeddings using DeepWalk to mediate the extreme dimensionality of the label space and also supplement rare labels with cooccuring label information.

Multi-hop question generation **Jan 2019 – May 2019**

- *Indian Institute of Technology Madras*
  Guide: Prof. Mitesh Khapra
  - Proposed and implemented sentence-level coverage mechanism for utilizing multi-hop information (information across sentences) for question generation.
  - Used Graph Convolutional Networks (GCNs) for leveraging knowledge from an entity graph defined on supporting passages in an End-to-end fashion.
  - Evaluated the appropriateness of GCNs for the task of Question Generation. [Report]

Time Series Forecasting **May 2018–Jul 2018**

- *Nanyang Technological University, Singapore*
  Guide: Prof. Suresh Sundaram
  - Developed a **novel architecture** which introduced recurrence in fuzzy neural networks, while using a faster one-shot projection based learning algorithm, for time series forecasting.
  - Tested on synthetic dynamical system identification problems and chaotic time series problems.
  - Achieved **ten times faster** performance than previously known algorithms, while also achieving similar or **better accuracy**. This work has been accepted by the **IEEE SSCI conference 2018**. [conference paper]

Automatic Speech Recognition (ASR) for Indian Languages **Dec 2017–May 2018**

- *Indian Institute of Technology Madras*
  Guide: Prof. S.Umesh
  - Developed **full-scale ASR systems** spanning both conventional machine learning models and the more recent deep learning methods.
  - Collected, processed and standardized native language text to create a **corpus with over 15 million words**. This was used to train a RNNLM (Recurrent Neural Network Language Model) for the languages Tamil, Telugu and Gujarati. [report]

Extraction of Definitional Sentences **Jan 2018–Apr 2018**

- *Indian Institute of Technology Madras*
  Guide: Prof. Sutanu Chakraborti
  - Proposed a hybrid model for Definition Extraction from text. The strength of lexicosyntactic pattern-matching approaches as well as the generalization capability of the Word Class Lattice approach were leveraged for better performance.
  - Achieved **improved F-measure (by 10%)** on definition extraction tasks. [report]

Talks and Presentations

- **Recurrence in Fuzzy Neural Networks and a faster training algorithm for Time Series forecasting** **Apr 2018**
  IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2018), Bangalore
  Presented a novel approach of introducing recurrence in fuzzy neural networks, while using a faster one-shot projection based learning algorithm, for time series forecasting. [SLIDES]
- **Page Rank for Word Sense Disambiguation** **Apr 2018**
  Indian Institute of Technology Madras
  Presented the idea of exploiting the capability of the Page Rank algorithm to handle the circularity of word definitions for an application in word sense disambiguation. [SLIDES]

Industrial Experience

Defence Research and Development Organization Bangalore **May 2017–Jul 2017**

- *Manager: Regu Kumar*
  Mentor: Alka Soni *Bangalore, India*
  The project was aimed at building a Real Time Executive for an Avionics system. It involved developing optimized code for task scheduling, interrupt handling and context switching for PowerPC P1024RM processor, e500v2 core. The developed application will be used in production of an airborne platform application. [REPORT]

DrumUP Bangalore **Dec 2016–Jan 2017**

- *Manager: Vishal Dutta*
  Mentor: Raghavendra Kumar *Bangalore, India*
  The internship involved developing Javascript applications to edit images and add captions online, and download analytics reports as CSV files from the DrumUp web frontend. The developed application **went into production**. [GITHUB]

Awards and Fellowships

- **NTU-India Connect Research Fellowship, *2018***: The NTU-India Connect Research Fellowship is offered to **meritorious undergraduate or graduate students** from Indian universities to pursue their research at Nanyang Technological University, Singapore (NTU) for a period of two to six months.
- **Among the top 4 in India** in the entrance exam conducted by **Indian Statistical Institute** in *2015*.
- Placed **third** in the Amazon AWS Deep Learning Hackathon at Shaastra, *2018*.
- **INSPIRE Fellowship, *2014***: The INSPIRE Fellowship is offered by the **Govt. Of India** to students who feature in the national **top 1 percent** of students among those who clear the CBSE board Class XII Exam.
- **Certificate of Merit** received from the **HRD Ministry, Govt. of INDIA** for excellent performance in the CBSE class X Examination, *2012*.

Skills and Tools

- **Programming languages**: C, Python, C++, Bash, TEX
- **Libraries**: TensorFlow, Scikit-learn, NLTK, SpaCy
- **Applications and Tools**: Embedded C, Scientifc Python, MATLAB, Octave, LTSpice
- **Operating Systems**: Linux, Windows

College Activities

- **The Fifth Estate, Correspondent**: The Fifth Estate is the student media body of IIT Madras. Apr 2017–Present
  - Initiated and authored magazine segments, reported campus matters.
  - Interviewed top researchers and other guest speakers.
- **Shaastra Web-Operations, Coordinator** Shaastra is the annual technical festival of IIT Madras. May 2016–Jan 2017
  - Handled the Shaastra participant-details database and web-signup forms.

Extra-Curricular Activities

- **Sports**: Jul 2015 - May 2016
  Was a part of the National Sports Organisation trained swimming team in my first year of college.
- **Music**
  Certified to have completed Grade 3 Keyboard Examination by Trinity College, London.
  Active participant of several events organised by Music Club, IIT Madras.