Ninad Khargonkar

linkedin.com/in/ninadkhargonkar | ninad.khargonkar@gmail.com | ninception.github.io | Ph: (413) 345-9601

Education

University of Massachusetts, Amherst — Amherst, MA
M.S in Computer Science — GPA: 4.0/4.0 — Sep '17 – May '19

Indian Institute of Technology, Kanpur — Kanpur, India
B.S in Mathematics and Scientific Computing — Jul '13 – Jun '17

Projects

Approximate Distribution for Sparse Data — Amherst, MA
Prof. Peter Haas, UMass Amherst — Jun–Aug 2018

- Worked on the problem of learning a distribution from sparse data set of disease prevalence and also modelling the transition of disease states across a year for 9 diseases.
- The principle of maximum entropy was used to infer the distribution with the marginals for individual diseases along with some top (ranked) feature pairs acting as the constraints to the optimization problem.
- The pipeline of data loading, feature selection and optimization for the input data set was implemented in Python.

Hindi Part of Speech tagging — Amherst, MA
Prof. Brendan O'Connor, UMass Amherst — Nov–Dec 2017

- Developed a part of speech (POS) tagger for Hindi by learning a word level translation between Hindi and English using word embeddings and a parallel corpora between the 2 languages.
- No supervised part of speech information was provided while training for the Hindi words and the 300-dimensional fastText mono-lingual word embeddings were used for training.
- The tagger also improved the performance by 12% on the proxy task of domain classification of documents in Hindi when the pos labels were added as an additional feature to an LSTM based model.

Modelling Uncertainty in Deep Learning — Amherst, MA
Prof. Erik Learned-Miller, UMass Amherst — Nov–Dec 2017

- Analyzed the problem of modelling uncertainty in the results from the outputs/scores of a deep neural network.
- Coded a neural network in Keras (Python) for the uncertainty value calculation on cifar-10 and mnist datasets.
- The uncertainty estimates also improved the performance of a linear classifier by around 2% in detecting out-of-training set images (proxy for anomaly detection) when added as additional features.

Topic Modelling of Documents – LDA — Kanpur, India
Prof. Piyush Rai, IIT-Kanpur — Jan–Apr 2017

- Implemented a document-topic categorization model on the Reuters news agency data set in Python (gensim-lib).
- Latent Dirichlet Allocation (LDA) statistical model was used for categorization and Variational Inference along with its stochastic extension was used for scaling up the inference process for the 7700 documents in training set.

Experience

Programmer – BioStats Lab, UMass Amherst — Sep–Nov 2017
Worked with Prof. Alkema's lab on a project on statistical indicators for family planning in developing countries and assisted in coding the simulation exercises and modelling the experiments in R programming language.

Globalink Research Internship — May–Jul 2016
Implemented scale down sampling on graphs using different random walks and analyzed their performance. Statistical graph models were used to test significance of network substructures by simulations. Worked on a side project on simulating team performance in Uefa Euro16 and submitted report to workshop program in Statistics (JSM 2017).

Technical Skills

Languages: Python, Java, C/C++, R, Matlab SQL | **Frameworks/Libs:** Keras/PyTorch, Scikit-learn, Pandas, Spark
Coursework: Machine Learning, Neural Networks, Bayesian Learning, Computer Vision, Natural Language Processing.

Awards & Activities

- Recipient of Inspire scholarship awarded by the Govt. of India for consistent academic performance in undergrad.
- Awarded the Mitacs Globalink scholarship for a research internship in Canada in the summer of 2016.
- Mentored a group of 6 freshmen and helped organize the orientation in the Counselling Service of IIT-Kanpur.