Ninad Khargonkar

LinkedIn: /in/ninadkhargonkar | ninad.khargonkar@gmail.com | ninception.github.io | Ph. 413-345-9601
248, Amherst Rd, Sunderland, MA – 01375

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

Generative Models for Document Embeddings
Prof. Benjamin Marlin, UMass Amherst — Oct–Dec 2018

- Working on the problem of learning a generative model for document level embeddings/vectors in an unsupervised manner. Using Edward/Tensorflow (Python) for the coding the probabilistic models.
- Current direction focuses on trying to improve the performance on hamiltonian monte carlo in inference of the document vectors by adopting better negative example sampling techniques.
- The next part includes tweaking the overall inference scheme to improve the performance on downstream nlp task of sentiment analysis in the multi-domain setting.

Chronic Disease Modelling from Sparse Data
Prof. Peter Haas, UMass Amherst — Jul–Oct 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 using the maximum entropy principle for learning.
- Piece-wise likelihood method was employed to partition the feature space graph to approximate the exact inference for the distribution with constraints picked from market basket analysis over the data.
- Constraints given by market-basket were also compared with top ranked feature pairs found out by L-measures (measuring statistical dependence from data) and the pipeline of data loading, feature selection and optimization was implemented in Python using Scipy, Numpy and Pandas.

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

- Developed a part of speech (POS) tagger for Hindi by transferring information using a learning a word level translation between Hindi & English using word embeddings and a parallel corpora between the 2 languages.
- No supervised part of speech information was provided during training and the 300 dimension fastText mono lingual word embeddings were used for learning the transition between words.
- The unsupervised 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.

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

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

Work 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

Language: Python, Java, R, Matlab, SQL | **Libs:** Keras/PyTorch, Scikit-learn, Pandas | **Others:** Git, Vim, LaTeX, AWS.
Coursework: Machine Learning, Neural Networks, Bayesian Learning, Computer Vision, Natural Language Processing.