Ben Athiwaratkun

Email: pa338@cornell.edu Phone: 413-884-3087

EDUCATION

Cornell University
Fourth *Year Phd Candidate in Statistics and Special Masters in Computer Science*
Advisor: Andrew Gordon Wilson

Williams College
Bachelor of Arts in Mathematics and Economics with Honors in Mathematics, June 2012
* Benedict First Prize in Mathematics 2010, Magna Cum Laude, Phi Beta Kappa and Sigma Xi

RESEARCH EXPERIENCE

Improving Stability in Deep Reinforcement Learning with Weight Averaging (UAI workshop, 2018)
* Improving deep reinforcement learning results with weight averaging.

Improving Consistency-Based Semi-Supervised Learning with Weight Averaging (Arxiv, 2018)
* Explore weight averaging methods to improve semi-supervised learning on image classification. Achieved state-of-the-art in all categories on CIFAR-10 and CIFAR-100.

An Exploration of Bayesian Methods for Auto-Encoders (In submission)
* Exploring different formulations of Bayesian approach on variational autoencoders, with application for diverse sentence generation and semi-supervised learning.

Adversarial Deep Averaging Networks for Cross-Lingual Domain Adaptation (TACL 2018)
* Developed a model for domain adaptation between languages using adversarial training
* Experiments on English to Chinese sentiment classification show that our method significantly outperforms baselines such as a machine translation system.

Probabilistic Word Representations with Subword Information (ACL 2018)
* Developed a method that incorporates subword information to enhance the quality of word density representations.
* Our model can handle rare or out-of-vocabulary words, and be applied to other languages without further hyper parameter tuning.

Hierarchical Density Order Embeddings (ICLR, 2018)
* Developed a methodology to train Gaussian representations on hierarchical data.
* Achieved state-of-the-art for WordNet hypernym prediction and HyperLex lexical entailment

Multimodal Word Distributions (ACL, 2017)
* Exploring the use of multimodal distribution as word representation to capture meaning multiplicity of words.

Neural Language Model for Malware Classification (ICASSP 2017)   Summer 2016
* Applied neural language models (LSTMs, GRUs) to sequences of APIs to extract file features.
* Experimented with multiple attention mechanisms on sequences of API features for malware file detection.

RELEVANT WORK EXPERIENCE

Research Intern   **Summer 2017 (AWS AI Lab, Palo Alto)**
- Explored the idea of balancing dictionary-level and word-level representations for word embeddings using group sparsity regularization.
- Pivoted the idea and applied the subword structure to multimodal word distributions.

Research Intern   **Summer 2016 (Microsoft Research, Redmond)**
- Experimented with neural language model and attention mechanism, as well as character level convolutional neural networks for malware detection. Published a paper to ICASSP 2017.

Research Intern   **Summer 2015 (Loop AI Labs, San Francisco)**
- Trained recurrent neural network language models (RNN, GRU, LSTM) on large dataset to improve the semantics of embeddings for relationship extraction tasks
- Comparing word embeddings trained by multiple models such as word2vec, GloVe, WordRep, as well as embeddings evolved by supervised learning such as tree LSTM for relationship extraction

MACHINE LEARNING PROJECTS

Heuristics Optimization for Neural Network Training   **Fall 2014**
- Explored the use of simulated annealing, genetic algorithm, and dynamically dimension search for neural network training.

Book Rating Prediction   **Spring 2014**
- Used a probabilistic matrix factorization algorithm with collaborative constraint to predict book rating.

Movie Revenue Prediction   **Spring 2014**
- Developed a neural network model with tailored cost function to perform movie revenue prediction.

Swingy Monkey - Reinforcement Learning   **Spring 2014**
- Developed a Q-Learning algorithm to train an agent to play a game (swingy monkey). The model successfully trained the agent to become increasingly more competent over time.

Pacman Tournament   **Spring 2014**
- Used support vector machine, Kmeans++, and state machine algorithm for Pacman Final Project Tournament. Won 2nd place of all teams (Harvard CS181)

RELEVANT COURSEWORK
- Machine Learning, Heterogeneous Parallel Programming, Natural Language Processing for Social Interactions, Heuristics Optimization, Advanced Machine Learning, Database, Computer Vision

AWARDS
- International Physics Olympiad 2007, Silver Medal

TOOLBOXES

Deep Learning Tools: Torch, Theano, Pylearn2, Tensorflow, PyTorch
Programming Languages: Python, Lua, R, Matlab/Octave, Java, C++, C, SQL

PROFILE

LinkedIn: www.linkedin.com/in/benathi   **Github**: www.github.com/benathi
Personal Website: benathi.github.io
Google Scholar: www.goo.gl/tuD8H4