Md Mustafizur Rahman

https://www.ischool.utexas.edu/~nahid/

Email : nahid@utexas.edu
Mobile : +1-434-227-3832

Education

University of Texas at Austin — Austin, TX
PhD in Information Science; GPA: 3.94 (Expected August 2021) — *Aug. 2016 - Present*

University of Virginia — Charlottesville, VA
Masters in Computer Science; GPA: 3.72 — *Aug. 2014 – May 2016*

Bangladesh University of Engineering and Technology — Dhaka, Bangladesh
M.Sc. in Computer Science and Engineering; GPA: 3.83 — *Mar. 2011 – Aug. 2013*

Bangladesh University of Engineering and Technology — Dhaka, Bangladesh
B.Sc. in Computer Science and Engineering; GPA: 3.95 — *Jan. 2006 – Feb. 2011*

Professional Experience

Research Assistant
University of Texas at Austin, Austin, TX — *Aug. 2016 - Present*

Answer Selection in Non-factoid Question Answering using Deep Learning - Developed a Convolutional Neural Network (CNN) based approach for non-factoid question answering. **Skills**: Python, Keras.

Researcher Intern
Bing Core Relevance, Microsoft, Sunnyvale, CA — *June 2019 – August 2019*

Question Answering using BERT and Multi-tasking - Worked on Bing's Tail question answering module using BERT via Multi-tasking. **Skills**: Python, TensorFlow.

Research Intern
Samsung Research America, Mountain View, CA — *May 2018 – Aug. 2018*

Open Domain Question Answering - Developed an open domain question answering system for Samsung's *Bixby* using a combination of a deep machine reading comprehension model and a deep learning to rank model. **Skills**: Python, Keras.

Applied Machine Learning Intern
Los Alamos National Laboratory, Los Alamos, NM — *June 2017 – Aug. 2017*

Semi-supervised Deep Learning for NLP - Developed a semi-supervised graph-based regularized deep learning model for cancer pathology reports. **Skills**: Python, Keras.

Graduate Research Assistant
University of Virginia, Charlottesville, VA — *May 2015 - Aug. 2015*

Hidden Topic Sentiment Model. (Rahman and Wang, WWW 2016). **Skills**: Java, Apache OpenNLP.

Selected Publications [Complete List]

1. **Md Mustafizur Rahman**, Mucahid Kutlu, and Matthew Lease, "Understanding and Predicting the Characteristics of Test Collections," *in preparation*

2. **Md Mustafizur Rahman**, Mucahid Kutlu, and Matthew Lease, "Efficient Test Collection Construction via Active Learning," *10th International Conference on The Theory of Information Retrieval (ICTIR 2020)*, Stavanger, Norway, 2020.

3. **Md Mustafizur Rahman**, Mucahid Kutlu, and Matthew Lease, "Constructing Test Collections using Multi-armed Bandits and Active Learning," *28th International World Wide Web Conference (WWW 2019)*, San Francisco, USA, 2019. (*Acceptance Rate: 20%.*)

4. Kezban Dilek Onal, Ye Zhang, Ismail Sengor Altingovde, **Md Mustafizur Rahman**, and others, "Neural Information Retrieval: At the End of the Early Years," *Information Retrieval Journal*, Springer, 2018.

5. Malay Bhattacharyya, Yoshihiko Suhara, **Md Mustafizur Rahman**, and Markus Krause, "Possible Confounds in Word-based Semantic Similarity Test Data," *20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, Portland, OR, USA, 2017. Blog: https://humancomputation.com/blog/?p=9492

6. **Md Mustafizur Rahman** and Hongning Wang, "Hidden Topic Sentiment Model," *25th International World Wide Web Conference (WWW 2016)*, Montreal, Canada, 2016. (*Acceptance Rate: 16%.*)

7. **Md Mustafizur Rahman**, Md. Monirul Islam, Kaziyuki Murase and Xin Yao, "Layered Ensemble Architecture for Time Series Forecasting," *IEEE transaction on Systems, Man and Cybernetics*, 46(1): 270–283, 2016.

Skills

Programming language: Java, Python, C, C++, LaTeX, TeX

Deep Learning: CNN, LSTM. Framework - Keras, Scikit-Learn, TensorFlow

Machine Learning: Naïve Bayes, Logistic Regression, Linear Regression, Bagging, Boosting, Random Forest, SVM, K-Nearest Neighbour, K-means, Expectation Maximization (EM), Hidden Markov Model (HMM), Active Learning. Co-training, Self-training, Semi-supervised Learning, Graph Regularization and Reinforcement learning – Multi-Armed Bandit algorithms, A/B Testing

Text mining: Probabilistic Latent Semantic Analysis (pLSA)

Operating Systems: Windows, Linux