

RESEARCH AREAS

Information Retrieval (IR), Natural Language Processing (NLP), and Deep Learning

EDUCATION

University of Texas at Austin <i>Ph.D. in Information Science (in progress); GPA: 3.86</i>	Austin, TX <i>Aug. 2016 – Present</i>
University of Virginia <i>Masters in Computer Science; GPA: 3.72</i>	Charlottesville, VA <i>Aug. 2014 – May 2016</i>
Bangladesh University of Engineering and Technology <i>M.Sc. in Computer Science and Engineering; GPA: 3.83</i>	Dhaka, Bangladesh <i>Mar. 2011 – Aug. 2013</i>
Bangladesh University of Engineering and Technology <i>B.Sc. in Computer Science and Engineering; GPA: 3.95</i>	Dhaka, Bangladesh <i>Jan. 2006 – Feb. 2011</i>

PROFESSIONAL EXPERIENCE

PhD Research Intern
Samsung Research America, CA *Present*

Open Domain Question Answering Developing a non-contextual question answering system using reading comprehension deep learning model. **Skills:** Python, Keras, Theano.

Applied Machine Learning Intern
Los Alamos National Laboratory, Los Alamos, NM *Summer 2017*

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

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

Efficient Test Collection Construction via Active Learning. (Rahman et al., arXiv:1801.05605, January 2018). **Skills:** Python, Indri, Active Learning.

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.

Graduate Research Assistant
University of Virginia, Charlottesville, VA *Summer 2015*

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

SELECTED PUBLICATIONS [[COMPLETE LIST](#)]

- 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.
- Md Mustafizur Rahman**, Mucahid Kutlu, Tamer Elsayed, and Matthew Lease, “Efficient Test Collection Construction via Active Learning,” Technical report, January 2018. arXiv:1801.05605.
- 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>
- Md Mustafizur Rahman** and Hongning Wang, “Hidden Topic Sentiment Model,” *25th International World Wide Web Conference (WWW 2016)*, Montreal, Canada, 2016. (*Acceptance Rate: 16%.*)

SKILLS

Programming language: Java, Python

Deep Learning: CNN, LSTM. Framework - Keras

Machine Learning: Naïve Bayes, Logistic Regression, Linear Regression, SVM, K-Nearest Neighbour, K-means, Expectation Maximization (EM), Hidden Markov Model (HMM), Active Learning. Co-training, Self-training, Semi-supervised Learning and Graph Regularization

Text mining: Probabilistic Latent Semantic Analysis (pLSA)