Naeemul Hassan
Assistant Professor
Department of Computer and Information Science
The University of Mississippi

+1 (817) 437 4518
nhassan@olemiss.edu
`www.cs.olemiss.edu/~nhassan`
216 Weir Hall, University, MS. 38677

RESEARCH INTERESTS

My research interests encompass database, data mining, and natural language processing. Computational journalism and social media sensing are my current research areas.

EDUCATION

Ph.D. in Computer Science. University of Texas at Arlington   August 2010 – August 2016
- CGPA: 4.0 out of 4.0
- Dissertation Title: Toward Automated Fact Monitoring and Checking. O1

PROFESSIONAL EXPERIENCE

Assistant Professor. University of Mississippi   August 2016 – Present
Directing the Data Exploration and Research laboratory (*dear.lab* `http://dear.cs.olemiss.edu`), a cross-domain facility for data-intensive research and application development. Developing and teaching the following undergraduate and graduate level courses.

Engr 691 Seminar on Natural Language Processing	Fall 2018
CSci 582 Computational Journalism	Spring 2017
CSci 543 Data Mining	Spring 2018, Fall 2016
CSci 517 Natural Language Processing	Fall 2018
CSci 444 Information Visualization	Fall 2017
CSci 387 Software Design & Development	Spring 2018, Spring 2017

Research Associate. Qatar Computing Research Institute (QCRI)   March 2013 – August 2013
Analyzed the problem of community and conflict detection in friend-foe network. It has many applications in social network analysis. Developed a probabilistic algorithm to solve the problem and validated its efficacy using Twitter data and Correlates of War data.

Network Planning & Engineering Intern. AT&T   May 2012 – August 2012
Voice call quality maintenance is a key operation for AT&T. Knowing the source (indoor/outdoor) of voice calls helps reduce troubleshooting cost. I analyzed AT&T's voice call data of three months, engineered features and developed a prediction model to determine indoor/outdoor status of a voice call. The method has been patented.

Course Instructor. University of Texas at Arlington
CSE 4335/5334 Data Mining   Fall 2014

Graduate Research Assistant. University of Texas at Arlington
Assisting with research activities such as, experimentation, data collection, documentation, model development, etc.   Fall 2012, 2015; Summer 2014, 2015; Spring 2015

Graduate Teaching Assistant. University of Texas at Arlington

CSE 5311 Design & Analysis of Algorithms	Fall 2013, 2011
CSE 3330 Database Systems & File Structures	Spring 2012, 2011, Summer 2011, Fall 2010
CSE 6324 Advanced Topics in Software Engineering	Spring 2014
CSE 1310 Introduction to Computers and Programming	Spring 2011

Lecturer. Daffodil International University, Bangladesh   Spring 2010 – Summer 2010

- **Learning Latent Representation**

  Sentence Representation C4: Vector representation of sentences is important for many text processing tasks that involve clustering, classifying, or ranking sentences. Recently, distributed representation of sentences learned by neural models from unlabeled data has been shown to outperform the traditional bag-of-words representation. However, Most of the existing representation learning methods consider only the content of a sentence and disregard the relations among sentences in a discourse by and large. The objective of this project is to explore novel models for learning latent representations of sentences that consider the content of a sentence as well as inter-sentence relations.

- **Computational Journalism**

  Clickbait Practice in the Media C2, C9, C8: The term clickbait refers to a form of web content that employs writing formulas and linguistic techniques in headlines to trick readers into clicking links, but does not deliver on its promises. The key purpose of this project is to systematically quantify the extents to which traditional print and broadcast media as well as the "unreliable" media use clickbait properties in contents published on the web.

  Automated Fact-checking C5, C6, C10, C11, C13, O2: Journalists spend a lot of time fact-checking claims made by politicians. I am developing an automated live fact-checking platform named **ClaimBuster** (`http://idir.uta.edu/claimbuster`). It has a text mining component that detects factual claims which are worth checking from political discourses such as presidential debates, a social media component that monitors check-worthy tweets and a community component that helps fact-checking enthusiasts discuss and collaborate with each other. This project have received media attention from multiple news outlets, including Austin American Statesman, Poynter and New Scientist. It has also won the Knight Foundation Prototype Fund.

  Significant Fact Monitoring J2, C15, C16: Developed an award winning system **FactWatcher**, which helps journalists identify data-backed, attention-seizing facts from ever-growing database. These facts serve as leads to news stories. Demonstration is available at `http://idir.uta.edu/factwatcher`.

- **Multi-dimensional Pareto-optimal Analysis**

  Skyline Group C14, J3, C17: We formulated the novel *Skyline Group* problem. One of its many applications is finding expert groups to complete tasks. Demonstrations on three domains (sports, question-answering, paper review) are available at `http://idir.uta.edu/crewscout`.

  Crowdsourcing Pareto-Optimal Objects Finding C12: This is the first study on crowdsourcing Pareto-optimal object finding, which has applications in public opinion collection, group decision making, and information exploration. Departing from prior studies on crowdsourcing skyline and ranking queries, it considers the case where objects do not have explicit attributes and preference relations on objects are strict partial orders. The partial orders are derived by aggregating crowdsourcers' responses to pairwise comparison questions. The goal is to find all Pareto-optimal objects by the fewest possible questions.

PUBLICATIONS

- **Peer-Reviewed Conference Publications**:

  C1 Mohammad Yousuf, Naeemul Hassan, Md Main Uddin Rony. Social Media as Source: A Large-scale Study of Mainstream and Unreliable News Sites. *In Proceedings of the Association for Education in Journalism and Mass Communication (AEJMC) Midwinter conference*, Oklahoma, USA, March, 2018.

C2 Md Main Uddin Rony, Naeemul Hassan, Mohammad Yousuf. BaitBuster: A Clickbait Identification Framework. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18)*, New orleans, Louisiana, USA, February, 2018.

C3 Mohammad Yousuf, Naeemul Hassan, Md Main Uddin Rony. Boon or Bane for Political Engagement: A Large-Scale Study of Normalization of Facebook. In *Proceedings of the 68th International Communication Association (ICA) annual conference*, Prague, Czech Republic, May, 2018.

C4 Tanay Kumar Saha, Shafiq Joty, Naeemul Hassan, Mohammad Al Hasan. Learning Sentence Representation with Context. In *Proceedings of the 26th ACM International Conference on Information and Knowledge Management (CIKM)*, Singapore, November, 2017.

C5 Naeemul Hassan, Gensheng Zhang, Fatma Arslan, Josue Caraballo, Damian Jimenez, Siddhant Gawsane, Shohedul Hasan, Minumol Joseph, Aaditya Kulkarni, Anil Kumar Nayak, Vikas Sable, Chengkai Li, Mark Tremayne. ClaimBuster: The First-ever End-to-end Fact-checking System. In *Proceedings of the 43rd International Conference on Very Large Databases (VLDB)*, 4 pages, Munich, Germany, August 2017.

C6 Naeemul Hassan, Fatma Arslan, Chengkai Li, and Mark Tremayne. Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster. In *Proceedings of the 23rd SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 10 pages, Halifax, Nova Scotia, Canada, August 2017.

C7 Naeemul Hassan, Mohammad Yousuf, Md Mahfuzul Haque, Javier A. Suarez Rivas, Md Khadimul Islam. Towards A Sustainable Model for Fact-checking Platforms: Examining the Roles of Automation, Crowds and Professionals. In *Proceedings of the 2017 Computation+Journalism Symposium (C+J2017)*, Illinois, October, 2017.

C8 Md Main Uddin Rony, Naeemul Hassan, Mohammad Yousuf. BaitBuster: Destined to Save You Some Clicks. In *Proceedings of the 2017 Computation+Journalism Symposium (C+J2017)*, Illinois, October, 2017.

C9 Md. Main Uddin Rony, Naeemul Hassan, Mohammad Yousuf. Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects? In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 8 pages, Sydney, Australia, 2017.

C10 Naeemul Hassan, Mark Tremayne, Fatma Arslan and Chengkai Li. Comparing Automated Factual Claim Detection Against Judgments of Journalism Organizations. In *Proceedings of the 2016 Computation+Journalism Symposium*, 5 pages, San Francisco, CA, USA, September 2016.

C11 Naeemul Hassan, Chengkai Li, and Mark Tremayne. Detecting Check-worthy Factual Claims in Presidential Debates. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 1835-1838, Melbourne, Australia, October 2015. DOI 10.1145/2806416.2806652.

C12 Abolfazl Asudeh, Gensheng Zhang, Naeemul Hassan, Chengkai Li, and Gergely Zaruba. Crowd-sourcing Pareto-Optimal Object Finding by Pairwise Comparisons. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 753-762, Melbourne, Australia, October 2015. DOI 10.1145/2806416.2806451.

C13 Naeemul Hassan, Bill Adair, James Hamilton, Chengkai Li, Mark Tremayne, Jun Yang and Cong Yu. The Quest to Automate Fact-Checking. In *Proceedings of the 2015 Computation+Journalism Symposium*, 5 pages, New York City, USA, October 2015.

C14 Naeemul Hassan, Huadong Feng, Ramesh Venkataraman, Gautam Das, Chengkai Li, and Nan Zhang. Anything You Can Do, I Can Do Better: Finding Expert Teams by CrewScout. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM)*, pages 2030-2032, Shanghai, China, November 2014. DOI 10.1145/2661829.2661839.

C15 Brett Walenz, You (Will) Wu, Seokhyun (Alex) Song, Emre Sonmez, Eric Wu, Kevin Wu, Pankaj K. Agarwal, Jun Yang, Naeemul Hassan, Afroza Sultana, Gensheng Zhang, Chengkai Li, and Cong Yu. Finding, Monitoring, and Checking Claims Computationally Based on Structured Data. In *Proceedings of the 2014 Computation+Journalism Symposium*, 5 pages, New York City, USA, October 2014.

C16 Afroza Sultana, Naeemul Hassan, Chengkai Li, Jun Yang, and Cong Yu. Incremental Discovery of Prominent Situational Facts. In *Proceedings of the 30th International Conference on Data Engineering (ICDE)*, pages 112-123, Chicago, Illinois, April 2014. DOI 10.1109/ICDE.2014.6816644.

C17 Chengkai Li, Nan Zhang, Naeemul Hassan, Sundaresan Rajasekaran, and Gautam Das. On Skyline Groups. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM)*, pages 2119-2123, Maui, Hawaii, November 2012. DOI 10.1145/2396761.2398585.

C18 Zamshed Iqbal Chowdhury, Masudul Haider Imtiaz, Muhammad Moinul Azam, Mst Sumi, Rumana Aktar, Md Rahman, Farzana Alam, Ishtiak Hussain and Naeemul Hassan. Design and Deployment of a Robust Remote River Level Sensor Network. *IEEE Sensors Applications Symposium (SAS)*, pages 244-249, San Antonio, Texas, February 2011. DOI 10.1109/SAS.2011.5739776.

C19 SM Rifat Ahsan, Mohammad Saiful Islam, Naeemul Hassan and Ashikur Rahman. Exploiting Packet Distribution for Tuning RTS Threshold in IEEE 802.11. *IEEE 25th Biennial Symposium on Communications (QBSC)*, pages 369-372, Kingston, Ontario, Canada, May 2010. DOI 10.1109/BSC.2010.5472955.

C20 SM Rifat Ahsan, Mohammad Saiful Islam, Naeemul Hassan and Ashikur Rahman. Packet Distribution Based Tuning of RTS Threshold in IEEE 802.11. *IEEE Symposium on Computers and Communications (ISCC)*, pages 1-6, Riccione, Italy, June 2010. DOI 10.1109/ISCC.2010.5546703.

- **Peer-Reviewed Journal Publications**:

J1 Mohammad R. Hasan, Naeemul Hassan , Rayan Khan, Young-Tae Kim, Samir M. Iqbal. Classification of Cancer Cells using Computational Analysis of Dynamic Morphology. In *Proceedings of the Computer Methods and Programs in Biomedicine*, 2018.

J2 Naeemul Hassan, Afroza Sultana, You Wu, Gensheng Zhang, Chengkai Li, Jun Yang, and Cong Yu. Data In, Fact Out: Automated Monitoring of Facts by FactWatcher. In *Proceedings of the VLDB Endowment (PVLDB)*, pages 1557-1560, Hangzhou, China, September 2014. DOI 10.14778/2733004.2733029. **Excellent demonstration award**.

J3 Nan Zhang, Chengkai Li, Naeemul Hassan, Sundaresan Rajasekaran, and Gautam Das. On Skyline Groups. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 26(4):942-956, April 2014. DOI 10.1109/TKDE.2013.119.

- **Other Publications (Dissertation, Technical Reports, Datasets)**:

O1 Naeemul Hassan. Toward Automated Fact Monitoring and Checking. Ph.D. Dissertation. University of Texas at Arlington. August 2016.

O2 Naeemul Hassan, Chengkai Li, Mark Tremayne. ClaimBuster: An Automated Live Fact-Checking Platform. *Graduate Poster*, ACES 2016. **President's Award**.

O3 Abolfazl Asudeh, Gensheng Zhang, Naeemul Hassan, Chengkai Li, and Gergely V. Zaruba. Crowdsourcing Pareto-Optimal Object Finding by Pairwise Comparisons. *Technical Report*, arXiv:1409.4161, September 2014.

O4 Facebook Posts of 150+ U.S. Media Organizations. `https://data.world/naeemulhassan/facebook-posts-of-150-u-s-media-organizations`

O5 1991 – 2004 NBA Boxscore Dataset. `http://naffi.blogspot.com/2012/03/nba-boxscore-dataset.html`

O6 Dataset of factual claims in all U.S. presidential debates. `http://idir-server2.uta.edu/classifyfact_survey`

AWARDS, FELLOWSHIPS & SCHOLARSHIPS

1. Sally McDonnell Barksdale Honors College (SMBHC) Faculty Summer Stipend. Spring 2018
2. President's Award for Graduate Poster Presentation. ACES 2016 (out of 40 submissions)
3. Ph.D. Dissertation Fellowship. University of Texas at Arlington. Spring 2016
4. Graduate Fellowship. University of Texas at Arlington. Fall 2015, Spring 2016
5. Outstanding Teaching Assistant Award. University of Texas at Arlington. Spring 2015
6. CIKM 2015 Travel Grant
7. Computation+Journalism Symposium 2015 Travel Grant
8. Excellent Demonstration Award. VLDB 2014 (out of 115 submissions)
9. STEM Fellowship. University of Texas at Arlington. Fall 2010 – present
10. Dean's List Award. Bangladesh University of Engineering & Technology (BUET). 2007 – 2009
11. Computer Science & Engineering Academic Merit Scholarship. BUET. 2007 – 2009
12. University Engineering Scholarship. BUET. 2006 – 2009

PATENT

Classification of Indoor and Outdoor Telecommunications Events of Mobile Telecommunications Networks. Mario Kosseifi, Roswell, GA (US); Mark Austin, Roswell, GA (US); Jeremy Fix, Acworth, GA (US); Naeemul Hassan, Arlington, TX (US); and Sheldon Kent Meredith, Marietta, GA (US). US20140155085 A1. December 4, 2012.

GRANT PREPARATION

1. NSF Innovation Corps (I-Corps^TM): ClaimBuster– Risk Assessment for Crowdfunding Platforms, Backers and Creators. Naeemul Hassan (Entrepreneurial Lead), Gensheng Zhang (Entrepreneurial Lead), Chengkai Li (Principal Investigator), Harold Strong (Mentor). National Science Foundation, $50,000, 11/15/15-04/30/16.

2. Knight Prototype Fund: ClaimBuster. Chengkai Li (PI), Naeemul Hassan, Mark Tremayne, Bill Adair, Jun Yang, Cong Yu. Knight Foundation, $35,000, 11/1/15-04/30/16.

MEDIA COVERAGE

1. *Clickbait headlines grow in mainstream media.* Deb Wenger & Khadimul Islam. March 2018. `http://newslab.org/clickbait-headlines-grow-mainstream-media/`

2. *Data Day: Journalism school hosts annual workshop.* The Mississippian. November 18, 2016. `http://thedmonline.com/section-ole-miss-data-2day-journalism-school-hosts-annual-workshop/`

3. *Naeemul Hassan Joins CIS Faculty.* Ole Miss News. November 1, 2016. `http://news.olemiss.edu/naeemul-hassan-joins-cis-faculty/`

4. *Engineering Students Win Awards in ACES Competition.* UT Arlington News. March 31, 2016. `http://www.uta.edu/engineering/news-events/news-archives/2016/03-ACES.php`

5. *Graduate students work to help make voting process easier.* Brittany Harborth. January 29, 2016.`http://www.theshorthorn.com/news/graduate-students-work-to-help-make-voting-process-easier/article_b7acf48c-c2e8-11e5-9091-d7aedf40fb3e.html`

6. *New Software Developed at UTA Tracks Candidates Statements.* Julie Fine. January 7, 2016. `http://www.nbcdfw.com/news/politics/New-Software-Developed-at-UTA-Tracks-Candidates-Statements-364565851.html`

7. *COLA Notes for December 2015.* College of Liberal Arts Newsroom. December 8, 2015. `https://utalibartsnews.wordpress.com/2015/12/08/cola-notes-for-december-2015/`

8. *In search of fact checkings Holy Grail: News outlets might not get there alone.* Craig Silverman. October 30, 2015. `https://medium.com/1st-draft/automated-fact-checking-and-verification-will-only-happen-if-organizations-other-than-newsrooms-can-84cf40689924`

9. *The Holy Grail of computational fact checking and what we can do in the meantime.* Alexios Mantzarlis. October 21, 2015. `http://www.poynter.org/news/international-fact-checking/379687/the-holy-grail-of-computational-fact-checking-and-what-we-can-do-in-the-meantime/`

10. *Reporters Lab projects featured at Computation + Journalism conference.* Julia Donheiser. October 6, 2015. `https://reporterslab.org/tag/university-of-texas-arlington/`

11. *Students work to develop new fact checking database.* Benjamin Owens. October 24, 2014. `http://www.theshorthorn.com/news/students-work-to-develop-new-fact-checking-database/article_23ce2b92-5bc3-11e4-9882-001a4bcf6878.html`

INVITED TALKS, PRESENTATIONS & GUEST LECTURES

1. Computational Solutions against Fake News: DB vs. AI Approaches, AAAI Tutorial, Feb, 2018.
2. Computational Journalism. 2017 Southeast Journalism Conference, Mississippi, Feb, 2017.
3. Detecting Check-worthy Factual Claims in Presidential Debates. Knight Foundation, Apr 2016.
4. ClaimBuster Demonstration. Tech & Check, Duke University, Mar 2016
5. Towards Automated Fact-checking. University of Mississippi, Mar 2016.
6. Computational Journalism. CSE 5334- Data Mining, University of Texas at Arlington, Feb 2016.
7. Data Science for Computational Journalism. The Dallas Morning News, Texas, Apr 2015.
8. Skyline Queries. CSE 5334- Data Mining, University of Texas at Arlington, Dec 2012.
9. Introduction to PHP. CSE 3330- Database, University of Texas at Arlington, Apr 2012.
10. On Skyline Groups. CSE 5334- Data Mining, University of Texas at Arlington, Apr 2012.

SERVICES

1. Guest Editor: ACM Journal of Data and Information Quality (ACM JDIQ)
2. Program Committee Member: Data Science and Systems (DSS-2018), The Web Conference 2018, CIKM 2017, SmartData 2017, SumPre 2017
3. Tutorial/Workshop Organizer: AAAI 2018, SEJC 2017
4. Sub-reviewer: VLDB 2018, WWW 2018, ACMSE 2018, DAPD 2018, ICDE 2017, VLDB 2017, ICDM 2016, CIKM 2016, SIGMOD 2015, VLDB 2014, EDBT 2014, ICDE 2014, ICDM 2013, ICDE 2013, CIKM 2013, WWW 2013, WAIM 2012, CIKM 2012, ICDM 2012, ICDM 2011, WAIM 2011.
5. Undergraduate Poster Judge: Mississippi Region VII Science Fair 2017, ACES 2016.