

# NABIL HOSSAIN

ML/NLP/DL expert looking for Senior position in product/research

[nabil.hossain@gmail.com](mailto:nabil.hossain@gmail.com)  
<https://cs.rochester.edu/u/nhossain>  
LinkedIn: <https://bit.ly/37gpElf>  
Phone: +1 (518) 368-7718  
Austin, TX

## EXPERIENCE

### SENIOR MACHINE LEARNING ENGINEER

Roku, Austin, TX

Jun 2022 – Present



- Working in Roku Voice Team: Natural Language Understanding (NLU), NLU evaluation, Training ML models to improve voice search experience & user engagement for Roku devices
- Model Deployment, Online Testing, Efficient Training Data Generation, Analysis and Visualization
- Frequently Used:** AWS, Gitlab CI/CD, Presto (SQL), Hive (Hadoop), Docker, Airflow, Streamlit, Dash

### MACHINE LEARNING ENGINEER

Intel, Folsom, CA

Aug 2020 – May 2022



- Worked at OmniBridge, a startup within Intel focused on providing a real-time translation solution to bridge the communication gap between the Deaf and the English-speaking communities
- Design, Train, Evaluate, Analyze & Deploy AI systems for American Sign Language Translation
- Machine Translation, Transformers, Dataset construction, Computer Vision, Python, PyTorch

### RESEARCH INTERN

Facebook, Seattle, WA

May 2019 – Dec 2019



- Research project focused on Natural Language Processing & Deep Learning for Text Generation
- Published Paper:** Simple and Effective Retrieve-Edit-Rerank Text Generation. [ACL 2020]
- Advisers:** Marjan Ghazvininejad (Mentor), Luke Zettlemoyer (Manager)

### RESEARCH INTERN

Amazon, Seattle, WA

May 2018 – Aug 2018



- Title:** Paraphrase Generation for Natural Language Understanding
- Augmented Alexa grammar with machine-synthesized utterances to achieve robust NLU
- Advisers:** Can Liu (Mentor), Markus Dreyer (Manager)

### RESEARCH INTERN

Microsoft, Redmond, WA

Feb 2018 – May 2018



- Invented a computational humor task: Machine Understanding of Humorous News Headlines
- Carefully created a dataset of news headlines edited for humor using efficient crowdsourcing, analyzed humor in the data, trained & evaluated LSTM classifiers for funny headline detection. [NAACL 2019]
- Used this dataset to launch an ML shared task, where 80+ teams participated & significantly improved the state-of-art for humor detection, leading to a **Best Task Award**. [SemEval 2020]
- Advisers:** John Krumm, Michael Gamon

### RESEARCH INTERN

Microsoft, Redmond, WA

May 2016 – Aug 2016



- Developed an algorithm to generate humorous Mad Libs. Suggested words for each blank are obtained using a Language Model and ranked for funniness using an ML classifier. [EMNLP 2017]
- Data Released:** Used Amazon Mechanical Turk to fill in Mad Libs and to judge their humor.
- Advisers:** John Krumm, Lucy Vanderwende, Eric Horvitz

### RESEARCH ASSISTANT

Toyota Technological Institute at Chicago, IL

Jul 2014 – June 2015



- Audio Search:** Developed algorithms to search for spoken queries in long audio tracks
- Advisers:** Karen Livescu, Greg Shakhnarovich

## SKILLS

---

**Python ML:** PyTorch, Huggingface, OpenNMT, Fairseq, Gensim, NLTK, Scikit-learn, Numpy, Pandas

- **Data Science/Mining/Analysis/Visualization:** Flask, Dash, Plotly, Matplotlib, Streamlit, Jupyter, Jinja, Django

**Deep Learning:** Generative AI, Transformer, GPT, Bert, LSTM, RNN, T5, Word2Vec, CNN, Autoencoders

**Misc:** Java, C/C++, SQL, Git, CI/CD, AWS, Docker, Linux, JavaScript, CSS, HTML, Latex

## EDUCATION

---

**University of Rochester** **Rochester, NY** **Jul 2015 – Aug 2020**

- PhD in Computer Science (Natural Language Processing, Machine Learning, Deep Learning)
- Thesis: *Creative Natural Language Generation: Humor and Beyond*
  - **Nominated for Best Dissertation Award**
- **Advisers:** Henry Kautz (University of Rochester), John Krumm (Microsoft Research AI)

**University of Houston** **Houston, TX** **Aug 2013 – Jun 2014**

- Graduate student & research assistant (Natural Language Processing, Machine Learning)

**Bard College** **Annandale-on-Hudson, NY** **Aug 2009 – May 2013**

- BA in Computer Science, BA in Mathematics (GPA – 3.90, Computer Science – 3.96, Math – 4.0)
- Thesis: *Implementation of the Solution to the Conjugacy Problem in Thompson's Groups*
  - **Nominated for Outstanding Thesis Award**

## SELECT PUBLICATIONS & PATENTS (550+ citations. Google Scholar: [qoo.g1/Awg1EY](https://scholar.google.com/citations?user=qoo.g1/Awg1EY))

---

1. **N. Hossain**, J. Krumm, M. Gamon and H. Kautz. "SemEval-2020 Task 7: Assessing Humor in Edited News Headlines." In *Proceedings of the 14th Workshop on Semantic Evaluation (SemEval)*. 2020. **Best Task Award**
2. **N. Hossain**, M. Ghazvininejad and L. Zettlemoyer. "Simple and Effective Retrieve-Edit-Rerank Text Generation". In *ACL*, 2020.
3. **N. Hossain**, J. Krumm, T. Sajed and H. Kautz. "Stimulating Creativity with FunLines: A Case Study of Humor Generation in Headlines". In *ACL (System Demonstrations)*, 2020.
  - Demo: <https://youtu.be/5OXJMxDBaLY> • TV News Interview: <https://bit.ly/2SCmqcD>
4. **N. Hossain**, Minh Tran and H.Kautz. "A Framework for Political Portmanteau Decomposition." In *ICWSM*, 2020.
5. **N. Hossain**, J. Krumm and M. Gamon. "President Vows to Cut Taxes Hair: Dataset and Analysis of Creative Text Editing for Humorous Headlines." In *NAACL*, 2019.
6. **N. Hossain**, T. Tran and H. Kautz. Discovering Political Slang in Readers' Comments. In *ICWSM*, 2018.
7. R. Magu, **N. Hossain** and H. Kautz. "Analyzing Uncivil Speech Provocation and Implicit Topics in Online Political News." *The 3rd International Workshop on News and Public Opinion (NECO) at ICWSM*, 2018.
8. **N. Hossain**, J. Krumm, L. Vanderwende, E. Horvitz and H. Kautz. "Filling the Blanks (hint: plural noun) for Mad Libs Humor." In *EMNLP*, 2017.
9. **N. Hossain**, T. Hu, R. Feizi, A.M. White, J. Luo and H. Kautz. "Precise Localization of Homes and Activities: Detecting Drinking-While-Tweeting Patterns in Communities." In *ICWSM*, 2016.
  - Press Reports: [MIT Tech Review](#), [CBS News](#), [UofR NewsCenter](#), [Gizmodo](#), [Business Insider](#), [Engadget](#), [Mashable](#), [Ars Technica](#), [Tech Spot](#), [Ubergizmo](#), [PC Magazine](#), [Digital Trends](#), [IBTimes](#), [Tech Times](#), etc.
10. R. Verma, N. Shashidhar, **N. Hossain** and N. Rai. "Automatic phishing email detection based on natural language processing techniques." *U.S. Patent 10,404,745*.

11. R. Verma and N. Hossain. "Semantic Feature Selection for Text with Application to Phishing Email Detection." *Information Security and Cryptology--ICISC 2013*. Springer International Publishing, 2014. 455-468.
12. R. Verma, N. Shashidhar and N. Hossain. "Detecting Phishing Emails the Natural Language Way." *Proceedings of the 17th European Symposium on Research in Computer Security (ESORICS)*, 2012.

## AWARDS & HONORS

---

- **Best Task Award**, SemEval **2020**
- **Conference Travel Grants**: ICWSM 2016, ICWSM 2018, NAACL 2019, ACL 2020
- **University of Rochester GSA Conference Travel Award**: May 2016, June 2019
- **Distinguished Scientist Scholarship**, Bard College – Full tuition scholarship **2009 – 2013**
- **John Bard Scholarship**, Bard College – Awarded annually to no more than two students **2011**  
     "For outstanding academic achievement in the field of major interest"
- **3<sup>rd</sup> Best Result in Bangladesh**, University of Cambridge A Level Examinations **2009**
- **3<sup>rd</sup> Best Result in Bangladesh**, University of Cambridge AS Level Examinations **2008**
- **Divisional Math Olympiad Champion**, Bangladesh **2007**
- **World Highest Score in Mathematics**, University of Cambridge O Level Examinations **2007**

## SERVICE

---

- **Reviewer**: NAACL '19, EMNLP '19, ACL '20, AAAI '20, SemEval '20, AAAI '21, SemEval '22, ACL '22
- **Steering Committee**: SemEval 2021
- **Organizer**: SemEval-2020 Task 7: Assessing Humor in Edited News Headlines
  - This shared task attracted over 80 participating teams

## RELEVANT GRADUATE COURSEWORK

---

Natural Language Processing	Machine Learning	Statistical Speech & Language Processing
Data Mining & Statistical Learning	Computer Vision	Advanced Algorithms

## MENTORING

---

- **Minh Tran** (Senior, University of Rochester, Fall 2019) – **Political Portmanteau Decomposition** (in ICWSM 2020)
- **Edan Meyer** (Junior, University of Rochester, Fall 2018) – **News Headline Generation & Style Transfer**
- **Marysia Tran** (Junior, Bard College, Summer 2017) – **Detection and Analysis of Political Slang** (in ICWSM 2018)

## TEACHING

---

**Teaching Assistant** **University of Rochester** **2016 – 2018**

- CSC 440 – Data Mining (Enrollment: 118, Languages: Python, Java)
  - Advise course projects of 3 teams, grade assignments, hold office hours
- DSC 450 – Data Science Practicum (Enrollment: 24, Languages: Python, SQL)
  - Advise capstone projects of 6 teams, lead SQL tutorial sessions, grade assignments & exams, hold office hours

**Teaching Assistant** **University of Houston** **Fall 2013**

- COSC 1410 Introduction to Computer Science (Enrollment: 34, Language: C++)
- Supervise lab classes twice a week, design & grade assignments and exams, office hours & exam review sessions

**Peer Tutor** **Bard College** **Fall 2011 – Spring 2013**

- Dedicated tutor for Math courses: Calculus 1 & 2, Proofs & Fundamentals (grade homework, hold office hours)
- Drop in tutor for all Computer Science courses (assist in understanding class lectures, assignments, programming)