Thomas Hartvigsen
Cambridge, Massachusetts
https://thartvigsen.github.io
twhartvigsen@wpi.edu

Research Interests

Data Mining, Time Series, Deep Learning, Reinforcement Learning, Explainability

Education

Worcester Polytechnic Institute, *Worcester, MA* 2016-present
PhD in Data Science, research focused on deep learning for time series
MS in Data Science 2018
Advised by Elke Rundensteiner and Xiangnan Kong

SUNY Geneseo, *Geneseo, NY*
BA, Applied Mathematics, minor in Biomathematics 2016

Employment Experience

Worcester Polytechnic Institute 2016-present
Research assistant with Elke Rundensteiner and Xiangnan Kong

Microsoft, Azure Machine Learning 2021
PhD Intern with Dipankar Ray
Detecting hate speech generated by large language models

UMass Medical School 2019
Research Intern with Jomol Matthew
Machine Learning to help doctors write clinical trials faster

University of Arizona 2015
NSF REU Intern with Shirley Papuga
Modeling the effects of drought on creosote bushes in the Sonoran desert via camera-trap images.

Grants

NSF-III: Timely Classification for Actionable Predictions (Under Review)
PI: Elke Rundensteiner, Co-PI: Xiangnan Kong.
This grant proposal is written based on my research (KDD'19 and KDD'20) and I am responsible for 90% of the writing.

Selected Honors & Awards

KDD Student Travel Award, NSF and ACM 2020
🏆 **Best Poster**, International Conference on Health Informatics 2020
IMA Travel Award ($500), University of Minnesota 2019
Graduate Student Travel Award ($1000), WPI 2019
Outstanding Graduate Research Award, WPI Data Science 2019
🏆 **Best Poster ($500)**, Graduate Research Innovation and Exchange, WPI 2019
People's Choice Poster Award, Graduate Research Innovation and Exchange, WPI 2017
GAANN Fellowship (Tuition Award + Annual Stipend), U.S. Dept. of Education 2016

Publications

I have published in KDD, AAAI, ACL, CIKM, ECML, IEEE BigData, HEALTHINF, and IEEE BHI.

Under Review

27. **Thomas Hartvigsen**, Walter Gerych, Jidapa Thadajarassiri, Xiangnan Kong, Elke Rundensteiner. *Early Classification of Irregular Time Series*.

26. **Thomas Hartvigsen**, Jidapa Thadajarassiri, Xiangnan Kong, Elke Rundensteiner. *Continuous-Time Attention Network for Irregularly-Sampled Time Series Classification*.

25. Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner. *Recovering The Propensity Score from Biased Positive Unlabeled Data*.

24. Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner. *Recurrent Bayesian Classifier Chains for Exact Multi-label Classification*.

23. Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Kavin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, Elke Rundensteiner, Emmanuel Agu. *Positive Unlabeled Learning with a Sequential Selection Bias*.

22. Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner, Emmanuel Agu. *Exact Multi-Label Classification with Incompletely Labeled Data*.

21. Dongyu Zhang, Cansu Sen, Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Human-like Explanation for Text Classification with Limited Attention Supervision*.

20. Hang Yin, Xinyue Liu, Xiangnan Kong, **Thomas Hartvigsen**, Yanhua Li. *Multi-State Brain Network Discovery*.

19. Luke Buquicchio, Walter Gerych, Kavin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, **Thomas Hartvigsen**, Elke Rundensteiner. *Variational Open-Set Recognition*.

REFEREED

18. Hang Yin, John Boaz Lee, Xiangnan Kong, **Thomas Hartvigsen**, Sihong Xie. *Energy-Efficient Models for High-Dimensional Spike Train Classification using Sparse Spiking Neural Networks*. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**), 2021 (238/1541 = 15.4% acceptance rate).

17. Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Semi-Supervised Knowledge Amalgamation for Sequence Classification*. AAAI Conference on Artificial Intelligence (**AAAI**), 2021 (1692/7911 = 20% acceptance rate).

16. Prathyush Parvatharaju, Ramesh Doddaiah, **Thomas Hartvigsen**, Elke Rundensteiner. *Learning Saliency Maps to Explain Deep Time Series Classifiers*. ACM International Conference on Information and Knowledge Management (**CIKM**), 2021 (271/1251 = 21% acceptance rate).

15. **Thomas Hartvigsen**, Cansu Sen, Xiangnan Kong, Elke Rundensteiner. *Recurrent Halting Chain for Early Multi-label Classification*. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**), 2020 (216/1279 = 16.9% acceptance rate).

14. Cansu Sen, **Thomas Hartvigsen**, Biao Yin, Xiangnan Kong, Elke Rundensteiner. *Human Attention Maps for Text Classification: Do Humans and Neural Networks Focus on the Same Words?* Annual Meeting of the Association for Computational Linguistics (**ACL**), 2020 (571/2244 = 17.6% acceptance rate).

13. **Thomas Hartvigsen**, Cansu Sen, Xiangnan Kong, Elke Rundensteiner. *Learning to Selectively Update State Neurons in Recurrent Networks*. ACM International Conference on Information and Knowledge Management (**CIKM**), 2020 (18% acceptance rate).

12. Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Learning Similarity-Preserving Word Meta-Embedding*. IEEE International Conference on Big Data (**BigData**), 2020 (15.5% acceptance rate).

11. Erin Teeple, **Thomas Hartvigsen**, Cansu Sen, Kajal Claypool, Elke Rundensteiner. *Clinical Performance Evaluation of a Machine Learning System for Predicting Hospital-Acquired Clostridium Difficile Infection*. International Conference on Health Informatics (**HEALTHINF**), 2020. 🏆 **Best Poster**.

10. **Thomas Hartvigsen**, Cansu Sen, Xiangnan Kong, Elke Rundensteiner. *Adaptive-Halting Policy Network for Early Classification*. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**), 2019 (170/1200 = 14.2% acceptance rate).

9. Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Patient-Level Classification of Clinical Note Sequences Guided by Attributed Hierarchical Attention*. IEEE International Conference on Big Data (**BigData**), 2019 (19.3% acceptance rate).

8. Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Learning Temporal Relevance in Longitudinal Medical Notes*. IEEE International Conference on Big Data (**BigData**), 2019 (19.3% acceptance rate).

7. Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *Comparing General and Locally-Learned Word Embeddings for Clinical Text Mining*. IEEE International Conference on Biomedical and Health Informatics (**BHI**), 2019.

6. Daniel Johnston†, Liubou Klindziuk†, Lolita Nazarov†, **Thomas Hartvigsen**, Elke Rundensteiner. *Early Diagnosis Prediction with Recurrent Neural Networks*.* IEEE MIT Undergraduate Research Technology Conference (**URTC**), 2019. 🏆 **Best Paper runner up**.

5. **Thomas Hartvigsen**, Cansu Sen, Elke Rundensteiner. *Detecting MRSA Infections by Fusing Structured and Unstructured Electronic Health Record Data*. International Joint Conference on Biomedical Engineering Systems and Technologies (**BIOSTEC**), 2018.

4. Julia Friend†, Alec Hauck†, Sruthi Kurada†, Cansu Sen, **Thomas Hartvigsen**, Elke Rundensteiner. *Handling Missing Values in Multivariate Time Series Classification*.* IEEE MIT Undergraduate Research Technology Conference (**URTC**), 2018.

3. **Thomas Hartvigsen**, Cansu Sen, Sarah Brownell†, Erin Teeple, Xiangnan Kong, Elke Rundensteiner. *Early Prediction of MRSA Infections using Electronic Health Records*. International Conference on Health Informatics (**HEALTHINF**), 2018. 🏆 **Best Student Paper finalist**.

2. Sarah Brownell†, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner. *MRSA Infection Prediction System*.* IEEE MIT Undergraduate Research Technology Conference (**URTC**), 2018.

1. Cansu Sen, **Thomas Hartvigsen**, Kajal Claypool, Elke Rundensteiner. *CREST - Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining*. European Conference on Machine Learning (**ECML**), 2017.

†undergraduate advisee.
*undergraduate paper under my supervision (3 total).

Selected Talks

Harvard University, invited speaker — Cambridge, MA
Adaptive-Halting Policy Networks for Early Classification — 2020
Host: Prof. Finale Doshi-Velez

Florida State University, invited speaker — Panama, FL
Adaptive-Halting Policy Networks for Early Classification — 2020
Host: Prof. Karen Works

The MITRE Corporation, invited speaker — Bedford, MA
Adaptive-Halting Policy Networks for Early Classification — 2020

Computational Sustainability Doctoral Consortium — Virtual Event
Adaptive-Halting Policy Networks for Early Classification — 2020

Worcester Polytechnic Institute, 3MT Competition — Worcester, MA
Early Classification of Clinical Time Series — 2020

University of Minnesota, Institute for Mathematics and its Applications — Minneapolis, MN
Adaptive-Halting Policy Networks for Early Classification — 2019

Northeastern University, New England Machine Learning Day — Boston, MA
Adaptive-Halting Policy Networks for Early Classification, poster — 2019

Worcester Polytechnic Institute, Arts and Sciences Week, invited speaker — Worcester, MA
Recurrent Models for Clinical Time Series — 2019

TEACHING

I have been the primary advisor for two Masters Theses and three NSF-funded REU groups.

Students Advised

- Prathyush Parvatharaju (MS Thesis), WPI 2019-Now
  - **Masters Thesis**: *Learned Saliency Maps to Explain Deep Time Series Classifiers*
- Ramesh Doddaiah (PhD student), WPI 2020-Now
- Aleksa Perucic (MS Thesis), WPI 2020
  - **Masters Thesis**: *SIFT - A Deep Network for Irregular Multivariate Time Series*
- Liubuo (Yuuna) Klindziuk (Undergraduate), Amherst College, NSF REU 2019
- Daniel Johnston (Undergraduate), Columbia University, NSF REU 2019
- Lolita Nazarov (Undergraduate), StonyBrook University, NSF REU 2019
- Julia Friend (Undergraduate), Oberlin College, NSF REU 2018
- Alex Hauck (Undergraduate), Valporaiso University, NSF REU 2018
- Sruthi Kurada, Advanced Math & Science Academy Charter School, NSF REU 2018
- Sarah Brownell (Undergraduate), Simmons University, NSF REU 2017
- Sean Tocci (Undergraduate), UMass Dartmouth, NSF REU 2017

Developed and led workshop on Deep Learning with PyTorch for Undergrads, WPI. 2019

SERVICE

Program Committee: AAAI ('21, '22), CVPR ('21), ICCV ('21), ACL ('21), EMNLP ('21)
External Reviewer: KDD ('18, '19, '20)
Conference Volunteer: KDD ('19, '20, '21), NeurIPS ('20)
Organized Deep Learning Reading Group, WPI 2019-2020
Graduate Student Advisory Council to the Dean of Arts & Sciences, WPI 2018-2020
Graduate Student Government Senate, WPI 2018
Data Science Graduate Student Council, WPI 2016-2019