Tom Hartvigsen

twhartvigsen@wpi.edu thartvigsen.github.io Cambridge, Massachusetts

INTERESTS: Data Mining, Time Series, Deep Learning, Reinforcement Learning, Explainability, AI for Medicine.

EDUCATION

Worcester Polytechnic Institute, *Worcester, MA*
Ph.D., Data Science — Expected 2021
Dissertation: *Observation and Prediction Timing in Time Series Classification*.
Committee: Elke Rundensteiner (Advisor), Xiangnan Kong (Advisor), Randy Paffenroth, Jenna Wiens.

SUNY Geneseo, *Geneseo, NY*
B.A., Applied Mathematics, minor in Biomathematics — 2016

EXPERIENCE

Research Intern, *UMass Medical School, Worcester MA* — 2019
Developed an automatic text summarization tool for clinical trial eligibility criteria to recommend inclusion and exclusion criteria for new clinical trials.
Supervisor: Dr. Jomol Matthew

Research Intern, *University of Arizona, Department of Environmental Science, Tucson AZ* — 2015
Built an instance-segmentation model for camera trap images of creosote bushes to model the effects of drought over time in the Sonoran desert.
Supervisor: Prof. Shirley Papuga

PUBLICATIONS

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

UNDER REVIEW

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

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

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

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

21. *Learning Saliency Maps to Explain Deep Time Series Classifiers.*
Prathyush Parvatharaju, Ramesh Doddiah, **Thomas Hartvigsen**, Elke Rundensteiner.

20. *Explainable Text Classification with Partially-Labeled Human Attention.*
Dongyu Zhang, Cansu Sen, Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.

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

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

PUBLISHED

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

16. *Semi-Supervised Knowledge Amalgamation for Sequence Classification.*
Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
AAAI Conference on Artificial Intelligence (**AAAI**), 2021.

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

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

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

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

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

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

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

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

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

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

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

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

3. *Early Prediction of MRSA Infections using Electronic Health Records.*
Thomas Hartvigsen, Cansu Sen, Sarah Brownell†, Erin Teeple, Xiangnan Kong, Elke Rundensteiner.

International Conference on Health Informatics (**HEALTHINF**), 2018. 🏆 Best Student Paper finalist.

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

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

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

Honors & Awards

CIKM Student Travel Award, ACM 2020
KDD Student Travel Award, NSF and ACM 2020
Graduate Student Travel Award ($1000), WPI 2020
IMA Travel Award ($500), University of Minnesota 2019
KDD Student Travel Award ($500), NSF and ACM 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
Graduate Student Travel Award ($1000), WPI 2018
Graduate Student Travel Award ($1000), WPI 2017
GAANN Fellowship (Tuition Award + Annual Stipend), U.S. Department of Education 2016

Grants

NSF-IIS: 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 Talks

Harvard University, invited speaker Cambridge, MA
Adaptive-Halting Policy Networks for Early Classification 2020

Florida State University, invited speaker Panama, FL
Adaptive-Halting Policy Networks for Early Classification 2020

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), Worcester Polytechnic Institute 2019-Now
  - **Masters Thesis**: *Learned Saliency Maps to Explain Deep Time Series Classifiers*
- Ramesh Doddaiah (PhD student), Worcester Polytechnic Institute 2020-Now
- Aleksa Perucic (MS Thesis), Worcester Polytechnic Institute 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), CVPR ('21), ICCV ('21), ACL ('21), EMNLP ('21)
External Reviewer: KDD ('18, '19, '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