Tom Hartvigsen

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

INTERESTS: Data Mining, Recurrent Neural Networks, Time Series, Reinforcement Learning, Interpretability.

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 — 2016
BioMathematics minor
Advisors: Prof. Chris Leary, Prof. Kirk Anne

EXPERIENCE

PhD Research Fellow, *Worcester Polytechnic Institute* — 2016 present
Solving time series classification problems for time-sensitive domains. I supervised two masters theses, led three summers of NSF-funded undergrad research, and wrote one NSF grant based on my research.
Advisors: Prof. Elke Rundensteiner, Prof. Xiangnan Kong

Research Intern, *UMass Medical School* — 2018 - 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, *NSF REU, University of Arizona, Department of Environmental Science* — 2015
Built an instance-segmentation model for remotely-captured images of creosote bushes to model the effects of drought over time in the sonoran desert.
Supervisor: Prof. Shirley Papuga

PUBLICATIONS

FULL PAPERS UNDER REVIEW

25. *Continuous-Time Attention Networks for Irregularly-Sampled Time Series Classification.*
Thomas Hartvigsen, Jidapa Thadajarassiri, Xiangnan Kong, Elke Rundensteiner.
In submission to ACM SIGKDD 2021.

24. *Learning Saliency Maps to Explain Deep Time Series Classifiers.*
Prathyush Parvatharaju, Ramesh Doddiah, **Thomas Hartvigsen**, Elke Rundensteiner.
In submission to ACM SIGKDD 2021.

23. *Recurrent Bayesian Classifier Chains for Exact Multi-label Classification.*
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner.
In submission to ACM SIGKDD 2021.

22. *Explainable Text Classification with Partially-Labeled Human Attention.*
Dongyu Zhang, Cansu Sen, Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
In submission to ACM SIGKDD 2021.

21. *Positive Unlabeled Learning with a Sequential Selection Bias.*
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Kavin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, Elke Rundensteiner, Emmanuel Agu.
In submission to ACM SIGKDD 2021.

20. *Exact Multi-Label Classification with Incompletely Labeled Data.*
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner, Emmanuel Agu.
In submission to ACM SIGKDD 2021.

19. *Multi-State Brain Network Discovery.*
Hang Yin, Xinyue Liu, Xiangnan Kong, **Thomas Hartvigsen**, Yanhua Li.
In submission to ACM SIGKDD 2021.

18. *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.
In submission to ACM SIGKDD 2021.

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

PEER-REVIEWED

16. *Semi-Supervised Knowledge Amalgamation for Sequence Classification.*
Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
AAAI 2021, main track.

15. *Recurrent Halting Chain for Early Multi-label Classification.*
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
ACM SIGKDD 2020, research track.

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.
ACL 2020, long paper.

13. *Learning to Selectively Update State Neurons in Recurrent Networks.*
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
CIKM 2020, long paper.

12. *Learning Similarity-Preserving Word Meta-Embedding.*
Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
IEEE BigData 2020, long paper.

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.
HEALTHINF 2020, long paper, best poster award.

10. *Adaptive-Halting Policy Network for Early Classification.*
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
ACM SIGKDD 2019, research track.

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

8. *Learning Temporal Relevance in Longitudinal Medical Notes.*
Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
IEEE 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 BHI 2019.

6. *Early Diagnosis Prediction with Recurrent Neural Networks.*
   Daniel Johnston†, Liubou Klindziuk†, Lolita Nazarov†, **Thomas Hartvigsen**, Elke Rundensteiner.
   IEEE 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.
   BIOSTEC 2018.

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

3. *Early Prediction of MRSA Infections using Electronic Health Records.*
   Thomas Hartvigsen, Cansu Sen, Sarah Brownell†, Erin Teeple, Xiangnan Kong, Elke Rundensteiner.
   HEALTHINF 2018.

2. *MRSA Infection Prediction System.*
   Sarah Brownell†, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
   IEEE URTC 2017.

1. *CREST - Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining.*
   Cansu Sen, **Thomas Hartvigsen**, Kajal Claypool, Elke Rundensteiner.
   ECML 2017.

†undergraduate advisee.

Honors & Awards

CIKM Student Travel Award, ACM	2020
KDD Student Travel Award, NSF and ACM	2020
Best Poster, HEALTHINF	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

Selected Talks

Harvard University, Data to Actionable Knowledge Group, invited speaker	Cambridge, MA
Adaptive-Halting Policy Networks for Early Classification	2020
Florida State University, Data Science Group, invited speaker	Panama, FL
Adaptive-Halting Policy Networks for Early Classification	2020
MITRE, Data Science Group	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

Worcester Polytechnic Institute, NSF REU Tutorial Worcester, MA
Introduction to Deep Learning with PyTorch 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

MENTORING/TEACHING

Students Advised

- Prathyush Parvatharaju (MS Thesis), Worcester Polytechnic Institute 2019-Now
  - 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 2019-2020
  - Thesis: *SIFT - A Deep Network for Irregular Multivariate Time Series*
- Liubuo 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

Teaching Assistant, SUNY Geneseo, Modeling Biological Systems (2x) and BioStats (1x). 2015-2016

SERVICE

Program Committee/Reviewing: AAAI ('21), CVPR ('21), ICCV ('21), ACL ('21)
External Reviewer: KDD ('19, '20)
Organized/led the Deep Learning Reading Group at 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