

Tom Hartvigsen

twhartvigsen@wpi.edu

[thartvigsen.github.io](https://github.com/thartvigsen)

INTERESTS: Machine Learning, Recurrent Neural Nets, Conditional Computation, Computational Sustainability.

EDUCATION

Worcester Polytechnic Institute, *Worcester, MA*

Ph.D., Data Science

2021

Advisors: Elke Rundensteiner, Xiangnan Kong

Earned MS in 2018

SUNY Geneseo, *Geneseo, NY*

B.A., Applied Mathematics

2016

BioMathematics minor

Advisors: Chris Leary, Kirk Anne

EXPERIENCE

GAANN PhD Fellow, *Worcester Polytechnic Institute*

2016-2021

Studying and developing recurrent models for a variety of challenging sequence classification tasks.

Supervisor: Elke Rundensteiner

Machine Learning Research Intern, *University of Massachusetts Medical School*

2018 - 2019

Worked on auto-summarization of clinical trial eligibility criteria for recommendation in new trials.

Supervisor: Jomol Matthew

NSF REU Intern, *University of Arizona, Department of Environmental Science*

Summer 2015

Built an image-segmentation model to process remotely-captured images to understand effects of drought on creosote bushes over time.

Supervisor: Shirley Papuga

Research Assistant, *SUNY Geneseo*

2014 - 2016

Modeled infection spread on graphs, built a graph dataset from IMDB, mined song lyrics for text features useful for discriminating genres and artists.

Supervisor: Chris Leary

PUBLICATIONS

PEER-REVIEWED PUBLICATIONS

1. *Predicting Hospital-Acquired Clostridium Difficile Infection Using Electronic Health Record Information.*
Erin Teeple, **Thomas Hartvigsen**, Cansu Sen, Kajal Claypool, Elke Rundensteiner.
International Conference on Health Informatics (**HEALTHINF**), 2020.
2. *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.
3. *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.
4. *Learning Temporal Relevance in Longitudinal Medical Notes.*
Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
IEEE International Conference on Big Data (**BigData**), Special Session on Intelligent Data Mining, 2019.

5. *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.**
7. *Detecting MRSA Infections by Fusing Structured and Unstructured Electronic Health Record Data.*
Thomas Hartvigsen, Cansu Sen, Elke Rundensteiner.
Communications in Computer and Information Science (**CCIS**) 1024, 2018.
8. *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. Short-listed for Best Student Paper.
9. *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.
10. *CREST - Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining.*
Cansu Sen, **Thomas Hartvigsen**, Kajal Claypool, Elke Rundensteiner.
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML**), 2017.

MANUSCRIPTS

1. *Reducing Computation in Recurrent Networks by Selectively Updating State Dimensions.*
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
2. *Recurrent Halting Chain for Early Multi-label Classification.*
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
3. *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.
4. *Deep Biased Positive Unlabeled Learning of Sequential Data.*
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Elke Rundensteiner.
5. *Explainable Document Classification with Human-guided Attention*
Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
6. *Similarity-Preserving Meta-Embedding.*
Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.

*Student under my supervision.

SELECTED AWARDS

Graduate Student Travel Grant, WPI	2020
Best Paper Runner Up, IEEE/MIT URTC.	2019
IMA Travel Grant, Institute for Mathematics and its Applications, U. of Minn.	2019
KDD 2019 Student Travel Grant, NSF and ACM	2019
Graduate Student Travel Grant, WPI	2019
Best Poster, Graduate Research Innovation and Exchange, WPI	2019
People's Choice Poster Award, Graduate Research Innovation and Exchange, WPI	2018
Graduate Student Travel Grant, WPI	2018
People's Choice Poster Award, Graduate Research Innovation and Exchange, WPI	2017
Graduate Student Travel Grant, WPI	2017
GAANN Ph.D. Fellowship, U.S. Department of Education	2016-2021

TEACHING

NSF REU Project Advisor, WPI.

Summers of 2017-19

Students: Y. Klindziuk, D. Johnston, L. Nazarov, J. Friend, A. Hauck, S. Kurada, S. Brownell, S. Tocci.

Outcomes: One paper per summer.

Teaching Assistant, SUNY Geneseo, Modeling Bio. Systems (2x) and BioStats (1x).

2015-2016

Modeling Biological Systems, SUNY Geneseo

2016

Guest lecturer: taught Percolation Models, created and led in-class exercise in R.

TECHNICAL SKILLS

Programming: Python, R, L^AT_EX, SQL.

Frameworks: PyTorch, TensorFlow, Scikit-learn, NumPy.