# Tom Hartvigsen

https://thartvigsen.github.io/
twhartvigsen@wpi.edu | 585 · 245 · 2767

## EDUCATION

**WORCESTER POLYTECHNIC INSTITUTE**
PH.D. IN DATA SCIENCE
Expected May 2021 | Worcester, MA
Advisor: Dr. Elke Rundensteiner

**SUNY GENESEO**
BA IN APPLIED MATHEMATICS
Minor in BioMathematics
May 2016 | Geneseo, NY

## LINKS

LinkedIn:// **in/thartvigsen**
Github:// **thartvigsen**

## COURSEWORK

### GRADUATE

Knowledge Discovery and Data mining
Statistical Learning
Deep Learning
Big Data Management
Database Management Systems
Introduction to Data Science
Business Intelligence

### UNDERGRADUATE

Modeling Biological Systems
Combinatorics
Differential Equations
Probability and Statistics
Linear Algebra
Calculus I-III

## SKILLS

### PROGRAMMING

**Python** - Deep Learning: **TensorFlow, PyTorch** , Machine Learning: **Scikit-Learn** , **Numpy** • **R** - Statistical Learning, Graph mining, ggplot2 • **Shell** • **LaTeX** • **SQL** - **PostgreSQL**

## AWARDS

2016- GAANN Research Fellowship
U.S. Department of Education

## EXPERIENCE

**WPI** | GAANN RESEARCH FELLOW
August 2016 – present | Worcester, MA

- Working with Dr. Elke Rundensteiner and Dr. Xiangnan Kong on sequential decision making for time series classification.

**UNIVERSITY OF ARIZONA** | RESEARCH EXPERIENCE FOR UNDERGRADUATES INTERN
June 2015 – Aug 2015 | Tucson, AZ

- School of Natural Sciences and the Environment advised by Dr. Shirley Papuga.
- Trained decision trees to segment images of Creosote Bushes with the aim of assessing phenological changes with respect to drought seasons (MATLAB).
- Presented findings at the Undergraduate Research Opportunities Consortium (UROC) in August 2015.

## RESEARCH

**DATA SCIENCE RESEARCH GROUP** Aug 2016 – Pres. | Worcester, MA

- Developing recurrent models for sequential classification tasks (PyTorch).
- Current work: **early time series classification** with application to data-driven improvements for healthcare.
- Developed CREST, a python-driven machine learning tool for infection detection in hospitals using SVMs, Random Forests, and Logistic Regression (Numpy/Scikit-Learn).

**BIOMATHEMATICS INNOVATION GROUP** Jan 2013 – May 2016 | Geneseo, NY

- Studied networks driven by differential equation models to model infection spread (R). Scraped song lyrics for sentiment analysis (Python). Scraped IMDB to build graphs relating films to one another (Python).
- Interdisciplinary focus, mentored younger students, brought 6 projects to undergraduate research conferences.

## PUBLICATIONS

- Teeple, E., **Hartvigsen, T.** , Sen, C., Rundensteiner, E. **Risk Stratification and Diagnostic Performance of a Machine Learning Algorithm for Clostridium Difficile Detection Using Electronic Health Records Data** . Currently in review at the Journal of Infection Control and Hospital Epidemiology.

- **Hartvigsen, T.** , Sen, C., Brownell, S., Teeple, E., Kong, X. and Rundensteiner, E. **Early Prediction of MRSA Infections using Electronic Health Records** . In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018) - Volume 5: HEALTHINF, pages 156-167, ISBN: 978-989-758-281-3. **Nominated for Best Student Paper** .

- Sen, C., **Hartvigsen, T.** , Claypool, K., Rundensteiner, E. (2017, September). **CREST - Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining** . ECML/PKDD 2017.