# Tom Hartvigsen

https://wpi.edu/people/doctoral-student/twhartvigsen
twhartvigsen@wpi.edu | 585 · 245 · 2767

## EDUCATION

### WORCESTER POLYTECHNIC INSTITUTE

PH.D. IN DATA SCIENCE
Expected May 2021 | Worcester, MA
Cum. GPA: 3.78 / 4.0
Advisor: Dr. Elke Rundensteiner

### SUNY GENESEO

BA IN APPLIED MATHEMATICS
Minor in BioMathematics
May 2016 | Geneseo, NY
Cum. GPA: 3.07 / 4.0

## 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
*to be completed May 2018

### UNDERGRADUATE

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

## SKILLS

### PROGRAMMING

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

## 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 on Deep Learning 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.
- Used MATLAB to design Machine Learning algorithms to segment sequences of images of Creosote Bushes to record changes through time in regards to drought.
- Presented findings at the Undergraduate Research Opportunities Consortium (UROC) in August 2015.

## RESEARCH

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

- Currently focusing on **early time series classification** with application to infection prediction.
- Developing deep recurrent models for using TensorFlow, moving towards Deep Reinforcement Learning.
- Developed CREST, a python-driven machine learning tool for infection detection in hospitals using SVMs, Random Forests, and Logistic Regression.
- Used the publicly available MIMIC III Database, Python, and PostgreSQL.

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

- Used R to build population networks driven by differential equation models. Used Python to collect lyrical data from song lyric websites. Used Python to scrape IMDB and build graphs relating movies.
- Built interdisciplinary projects, mentored younger students, and brought 6 projects to undergraduate research conferences.

## PUBLICATIONS

- **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.

- 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**. Preparing for submission the the New England Journal of Medicine in January 2018.