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: 3.83 / 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 Introduction to Data Science Statistical Learning Database Management Systems Big Data Management*
Business Intelligence*
Machine Learning**
Deep Learning**
*to be completed Dec 2017
**to be completed May 2018

UNDERGRADUATE

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

SKILLS

PROGRAMMING

Python - TensorFlow for Deep Learning, Scikit-Learn for Machine Learning • R -Statistical Learning, Graph mining, ggplot2 • Shell • MATLAB - General Programming • LTFX • SQL - PostgreSQL

for research, **SQLPLUS** for DBMS Course

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 Machine Learning/Deep Learning on Time Series Data.

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., Rundensteiner, E. (January, 2018). Early Prediction of MRSA Infections using Electronic Health Records. HealthInf 2018.
- Sen, C., Hartvigsen, T., Claypool, K., Rundensteiner, E. (2017, September).
 CREST Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining. ECML/PKDD 2017.
- Hartvigsen, T., Sen, C., Teeple, E., Rundensteiner, E. (2017). 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 October 2017.