Todd Bodnar

Last updated 7/31/2015
CV on GitHub

Contact Information

E-mail: ToddBodnar@gmail.com
Web: ToddBodnar.com
Phone: (650) 731-2586

Work Experience

Research Assistant at the Salathé Group (2012 - 2015)
Datamined large datasets including Twitter (≈ 3 Billion Tweets) and Open Street Map
Text classification and Sentiment analysis involving Weka and hand-built tools

Consultant at Human Longevity (Spring 2015)
Social media analytics for trend analysis.

Contractor at Wright Patterson Air Force Base (May - July 2014)
Developed large-scale, unsupervised event detection system using social network data.

Education

Pennsylvania State University

PhD, Biology, 2015 (Expected)
Advisor: Marcel Salathé

Santa Fe Institute

Complex Systems Summer School, 2013
Surveys of the fields of: complexity, nonlinear dynamics, network structure and dynamics, computation theory and adaptive computation techniques, and machine learning.

Pennsylvania State University

B.S., Computer Science, May, 2012
Junior/Senior GPA 3.90
Minor in Mathematics

Select Software Developed

Weka Hadoop
https://github.com/ToddBodnar/weka-hadoop
Map-Reduce port of Weka's experiment platform

Moocdemic
https://www.moocdemic.com
Mobile game simulating outbreaks (≈10,000 users)
Developed press release tools using Amazon Turk and A/B testing
Created visualizations of game play elements (samples available here)

Crowdbreaks
www.crowdbreaks.com
Online disease surveillance system using social media.
Worked on back end, developed system for fast processing of large geospatial data sets.

Computer Skills

- Cluster Computing / Hadoop
- Web based psychological game development through Amazon Mechanical Turk
- Machine Learning, Data Mining, Database Management
- Java, Python, BASH, R, Hive, LaTeX
- Basic familiarity with C, JavaScript and PHP

Select Publications

- **Todd Bodnar**, Conrad Tucker, Kenneth Hopkinson, and Sven G. Bilén. *Increasing the Veracity of Event Detection on Social Media Networks Through User Trust Modeling* IEEE BigData 2014 Link

  Social media, such as Twitter, is being used by many parties to measure something in the real world. However, many users are not experts and can often be wrong, which lowers the value of social media analytics. We looked at several rumors, some of which were true, and analyzed users that either accurately or not accurately tweeted about the rumor. This allows an end user to filter out inaccurate users from their analysis. My contribution was developing the method to classify users as likely or unlikely to be providing useful information.

- **Todd Bodnar**, Victoria Barclay, Nilam Ram, Conrad Tucker and Marcel Salathé. *On the Ground Validation of Online Diagnosis with Twitter and Medical Records* WWW 2014 Link

  We contacted individuals that had recently been to a local health provider to see if they had a Twitter account. If they did, and consented to a data release, we collected their Tweets along with their medical data. The collected Twitter data was then used to build machine learning models to classify a user's medical state. My contribution was collecting Twitter data and performing analysis.

- Cheryl Abundo, **Todd Bodnar**, John Driscoll, Ian Hatton, and Jody Wright *City population dynamics and fractal transport networks* Proceedings of Santa Fe Institute Complex Systems Summer School Link

  We used OpenStreetMap to measure fractal patterns of the world's roads. This measure was then compared against various social, geographic, or political factors. My contribution was scaling up the calculation of fractals to cover the entire dataset by implementing a custom map-reduce program that could be run on a supplied cluster.