# COLLEGE OF ENGINEERING AND COMPUTER SCIENCE
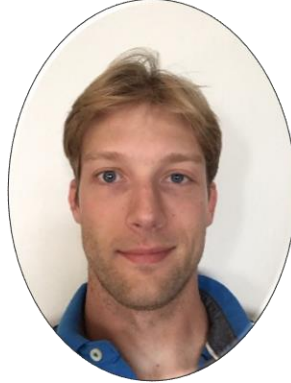## FLORIDA ATLANTIC UNIVERSITY

Announces the Ph.D. Dissertation Defense of

# Matthew A. Herland

for the degree of Doctor of Philosophy (Ph.D.)

## "Big Data Analytics and Engineering for Medicare Fraud Detection"

**March 20, 2019, 1:30 p.m.**
**EE 405**
**777 Glades Road**
**Boca Raton, FL**

DEPARTMENT:
Computer and Electrical Engineering and Computer Science

ADVISOR:
Taghi M. Khoshgoftaar, Ph.D.

PH.D. SUPERVISORY COMMITTEE:
Taghi M. Khoshgoftaar, Ph.D., Chair
Martin Solomon, Ph.D.
Hanqi Zhuang, Ph.D.
Xingquan Zhu, Ph.D.

ABSTRACT OF DISSERTATION
Big Data Analytics and Engineering for Medicare Fraud Detection

The United States (U.S.) healthcare system produces an enormous volume of data with a vast number of financial transactions generated by physicians administering healthcare services. This makes healthcare fraud difficult to detect, especially when there are considerably less fraudulent transactions than non-fraudulent. Fraud is an extremely important issue for healthcare, as fraudulent activities within the U.S. healthcare system contribute to significant financial losses. In the U.S., the elderly population continues to rise, increasing the need for programs, such as Medicare, to help with associated medical expenses. Unfortunately, due to healthcare fraud, these programs are being adversely affected, draining resources and reducing the quality and accessibility of necessary healthcare services. In response, advanced data analytics have recently been explored to detect possible fraudulent activities. The Centers for Medicare and Medicaid Services (CMS) released several `Big Data' Medicare claims datasets for different parts of their Medicare program to help facilitate this effort. In this dissertation, we employ three CMS Medicare Big Data datasets to evaluate the fraud detection performance available using advanced data analytics techniques, specifically machine learning. We use two distinct

approaches, designated as anomaly detection and traditional fraud detection, where each have very distinct data processing and feature engineering. Anomaly detection experiments classify by provider specialty, determining whether outlier physicians within the same specialty signal fraudulent behavior. Traditional fraud detection refers to the experiments directly classifying physicians as fraudulent or non-fraudulent, leveraging machine learning algorithms to discriminate between classes. We present our novel data engineering approaches for both anomaly detection and traditional fraud detection including data processing, fraud mapping, and the creation of a combined dataset consisting of all three Medicare parts. We incorporate the List of Excluded Individuals and Entities database to identify real-world fraudulent physicians for model evaluation. Regarding features, the final datasets for anomaly detection contain only claim counts for every procedure a physician submits while traditional fraud detection incorporates aggregated counts and payment information, specialty, and gender. Additionally, we compare cross-validation to the real-world application of building a model on a training dataset and evaluating on a separate test dataset for severe class imbalance and rarity.

BIOGRAPHICAL SKETCH
Born in Miami, Florida
B.S. 2011, Florida Atlantic University, Boca Raton, Florida
M.S. 2012, Florida Atlantic University, Boca Raton, Florida
Ph.D. 2019, Florida Atlantic University, Boca Raton, Florida

CONCERNING PERIOD OF PREPARATION
& QUALIFYING EXAMINATION

**Time in Preparation:** 2015 - 2019

**Qualifying Examination Passed:** Spring 2013

**Selected Published Papers:**

Matthew A. Herland, Richard A. Bauder, and Taghi M. Khoshgoftaar. The Effects of Class Rarity on the Evaluation of Supervised Healthcare Fraud Detection Models. Journal of Big Data, 2019.

Matthew A. Herland, Taghi M. Khoshgoftaar, and Richard A. Bauder. "Big Data Fraud Detection using Multiple Medicare Data Sources." Journal of Big Data 5.1 (2018): 29.

Matthew Herland, Richard A. Bauder, and Taghi M. Khoshgoftaar. "Approaches for Identifying US Medicare Fraud in Provider Claims Data." Health care management science (2018): pages 1-18.

Matthew A. Herland, Richard A. Bauder, and Taghi M. Khoshgoftaar. "Medical Provider Specialty Predictions for the Detection of Anomalous Medicare Insurance Claims." 2017 IEEE International Conference on Information Reuse and Integration (IRI). IEEE, 2017, pages 579-588.

Richard A. Bauder, Taghi M. Khoshgoftaar, Aaron Richter, and Matthew Herland. "Predicting Medical Provider Specialties to Detect Anomalous Insurance Claims." Tools with Artificial Intelligence (ICTAI), 2016 IEEE 28th International Conference on. IEEE, 2016, pages 784-790.