My Summer Experience: "Data Science Across Disciplines" REU Program

Sarah McDougall '21

Program Information

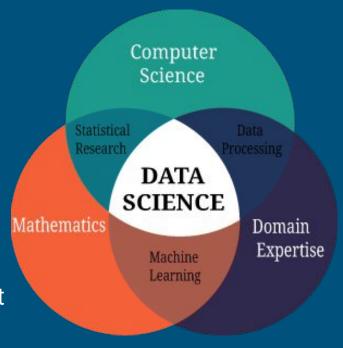
- Name: Data Science Across Disciplines REU Program
- Hosted by: Marquette University (Milwaukee, WI)
- Sponsored by: National Science Foundation (NSF)
- Duration: June 1 August 7 (10 weeks)
- Stipend: \$600/week, \$750 for food, ~\$1000 for housing
- ** held remotely due to COVID-19 **





Program Details

- Several research areas:
 - health care
 - data science ethics
 - exercise science
 - social science
 - o computer science education
 - bioinformatics
 - spatial analytics, security
- Work with mentor on a broadly defined project
 - Conduct background research
 - Raise research questions
 - Conduct research
 - Write research paper, deliver presentation, create poster



My Project

- "Predicting Opioid Overdose Readmission and Opioid Use Disorder (OUD)
 with Machine Learning"
- Used deidentified hospital data from Medical College of Wisconsin
- Constructed and analyzed machine learning models:
 - Likelihood of hospital readmission following opioid overdose
 - Likelihood of OUD diagnosis following prescription of opioids

My Project





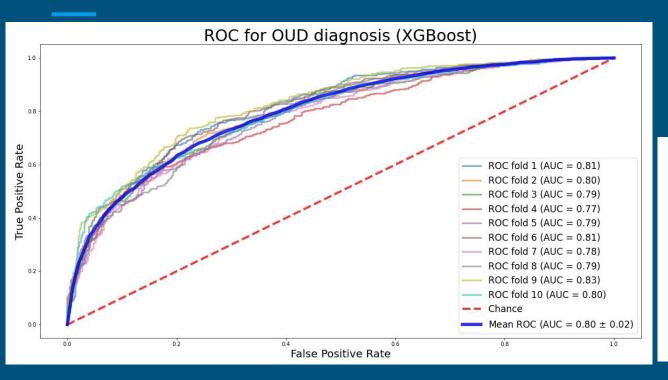




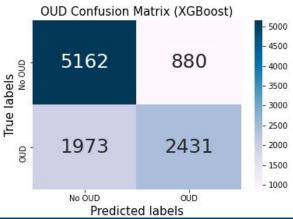


- Used Python packages pandas, scikit-learn
- Supervised learning
 - Did patient have a hospital readmission after 30/90 days?
 - Did patient receive OUD diagnosis within 6 months of being prescribed opioids?
- Dataset makeup:
 - Demographics
 - Social history
 - Procedures
 - Lab results
 - Diagnoses

My Project - OUD Study Results



- Random Forest
- Support Vector Machine (SVM)
- AdaBoost
- XGBoost



Breakdown of a Typical Day/Week

- 40 hours/week
- Weekly breakdown
 - Monday professional development discussion
 - Tuesday meeting with project team
 - Wednesday weekly check-in
 - Occasional presentations
- Daily tasks may include
 - Conducting background research
 - Data preprocessing, running tests, following tutorials
 - Communicating with project team
 - Updating documentation and work log
 - Working on final paper, presentation, and poster
- Overall, very loose structure time management and organization are key

COVID-19 Adjustments

- Zoom / Microsoft Teams
- GroupMe
- Jackbox Game Nights
- Pizza Socials
- Virtual Presentations
- Virtual Poster Session



Going Beyond the Summer

Predicting Opioid Overdose Readmission and Opioid Use Disorder with Machine Learning

Sarah McDougall Priyanka Annapureddy **Bucknell University** Marquette University Email: snm009@bucknell.edu Email: privanka.annapureddy@marquette.edu Email: praveen.madiraiu@marquette.edu

Dr. Prayeen Madiraiu Marquette University

Nicole Fumo, MPH Medical College of Wisconsin Email: nfumo@mcw.edu

Dr. Stephen Hargarten, MD, MPH Medical College of Wisconsin Email: hargart@mcw.edu

Abstract-Opioid use disorder (OUD) is a medical condition associated with problematic patterns of opioid use that cause interpersonal and social impairment. This research demonstrates how supervised machine learning can be used to predict patients at risk of hospital readmission following opioid overdose, and to predict patients at risk of developing OUD. Two labeled datasets were built from deidentified hospital data provided by a Level I Trauma Center Hospital. Several machine learning models were constructed (logistic regression, random forest, support vector machine, AdaBoost, XGBoost) and validated with 10 iterations of 10-fold cross validation. The XGBoost classifier can sufficiently predict patients at risk for OUD (AUC = 0.80, precision = 0.73, recall = 0.55). This work can assist providers in determining appropriate preventive care and resources for at-risk patients.

Index Terms-opioid use disorder, opioid dependence, patient readmission, machine learning, electronic health records

I. INTRODUCTION

The Centers for Disease Control and Prevention (CDC) define opioid use disorder (OUD) as a "problematic pattern of opioid use that causes significant impairment or distress" [1]. OUD diagnosis relies on several criteria, including excessive and the state of t

deaths decreased 4% from 2017 to 2018. However, nearly 70% of these drug deaths involved an opioid. While the prescription opioid and heroin death rates decreased 13.5% and 4%, respectively, the synthetic opioid-involved death rate (excluding methadone) increased 10% from 2017 to 2018 [4]. The use of illegally manufactured fentanyl, which is 50 times more potent than heroin and 100 times more potent than morphine, greatly contributes to the increased death rate [5].

We can potentially reduce opioid overdose and death rates by providing OUD patients with preventive interventions and specialty treatments that align with the patients' needs [6]. Nationally, only 23% of OUD patients receive specialty treatment service [7]. These services include counseling and behavioral therapies, recovery support services, and medicationassisted treatment (MAT) with methadone, buprenorphine, or naltrexone [5]. Further efforts must be made to provide at-risk patients with the appropriate treatment options.

The Wisconsin Onioid Enidemic

- Periodic meetings with project team
- Acceptance to 2020 IEEE Conference on Big Data (IEEE BigData 2020)
 - **December 10 13**
 - Special Session on Health Care Data
 - More information:

https://bigdataieee.org/BigData2020/

Application Process

- 2-page essay
 - o Interest in data science, goals for the future, goals for the REU program
- College transcript
- 2 recommendation letters
 - At least 1 from faculty member at your home institution

*** APPLY FOR NEXT YEAR: Applications open December 2020, close February 2021 ***

Contact Information

- Presenter: Sarah McDougall
 - o Email: snm009@bucknell.edu
 - LinkedIn: https://www.linkedin.com/in/sarahnmcdougall/
- Program Director: Dr. Praveen Madiraju
 - Email: <u>praveen.madiraju@marquette.edu</u>
- Program Information:
 - o REU site: https://reu.cs.mu.edu/index.php/Main_Page
 - o Email: reu@marquette.edu