Grace Tang
Springboard DSC - Capstone 2
September 2020

Capstone 2 Proposal: Detecting Heavy Drinking with Smartphone Accelerometer Data

Alcohol consumption has been associated with injury, short- and long-term health consequences, chronic and acute illnesses, and unintentional death. A 2010 CDC report found that an estimated 88,000 people die from alcohol-related causes annually, making alcohol the third leading preventable cause of death in the United States, preceded only by tobacco (first leading preventable cause) and poor diet and physical inactivity (second leading preventable cause).

The NIAAA reports that college-age students (ages 18-24) are a particularly high-risk group for developing problems related to alcohol, compared with other age groups. College-age students are the most likely to drink heavily, regardless of gender or ethnicity. The drinking is even more frequent amongst those enrolled as full-time college students. According to the 2018 NSDUH, 54.9% of full-time college students ages 18-22 drank alcohol in the past month compared with 44.6% percent of other persons of the same age. In the same report, it was found that 36.9% of college students ages 18-22 reported binge drinking in the past month compared with 27.9% of other persons of the same age. Drinking by college-age students has been associated with unintentional death (estimated 1,825 students/year), injury (~599,000 students/year), physical assault (~696,000 students/year), sexual assault (>97,000 students/year), drunk driving (~2.7 million students/year), and alcohol abuse disorders (~20 % of college students).

It is clear that college-age drinking is a kind of health crisis, and necessitates intervention. One particular method used to curb heavy drinking behavior is just-in-time adaptive interventions (JITAIs). This method sends prompts or cues in a strategically timed manner to the user (sent "just-in-time"), in combination with user feedback, to alter user behavior and change habits. It would be particularly beneficial to develop a method of detecting when a student may be engaged in heavy drinking, and prompt the JITAI to prompt the student. This can be implemented in mobile devices.

Can we use those same mobile devices to not only initiate JITAIs, but additionally detect heavy drinking? Killian et al. (2019) developed a dataset to answer that question. In the UCI Machine Learning Repository, Killian et al. collected smartphone accelerometer data from 13 anonymized Ohio State University college students and corresponding Transdermal Alcohol Concentration (TAC) data using SCRAM ankle bracelets, during a bar crawl. The SCRAM bracelets provided 24/7 transdermal alcohol testing, and functions like a breathalyzer for the ankle. SCRAM bracelets show better TAC accuracy at higher alcohol concentrations, making the dataset unsuitable for determining all instances of drinking, but suitable for detecting heavy drinking.

Our criteria for success in this project is to develop a machine learning algorithm that can classify, using only smartphone accelerometer data, when someone is considered legally impaired from drinking alcohol. This is defined as when a person's blood alcohol content (BAC) is 0.08% or higher. We would split the data into test/train sets, with the accelerometer data as

the independent variables, and the TAC data as the dependent variable. The accelerometer data could be cleaned and processed to produce other intermediary variables useful for predicting drunkenness, such as a person's gait, cadence, number of steps, etc. After being trained on the data, the algorithm would need to reliably classify the test data (>50% accuracy as a bear minimum) when someone is above or below the 0.08% BAC threshold.

The authors of the dataset noted that the smartphone accelerometer data and TAC data are both subject to lots of noise. This may require various smoothing and preprocessing methods before we can feed the data to our classifiers.

The deliverables of this Capstone 2 Project are: code for a reliable classifier, a slide deck of the project, and a project report. These will be collected in a GitHub repo and submitted at the end of the project. The results of this project would have a variety of applications, from better implementation of JITAIs for curbing heavy drinking, to law enforcement applications, to medical interventions in severe drinking circumstances.

References:
- https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/alcohol-facts-and-statistics
- https://pubs.niaaa.nih.gov/publications/aa74/aa74.htm
- https://pubs.niaaa.nih.gov/publications/aa87/aa87.htm

Dataset Citation:
- Killian, J.A., Passino, K.M., Nandi, A., Madden, D.R. and Clapp, J., Learning to Detect Heavy Drinking Episodes Using Smartphone Accelerometer Data. In Proceedings of the 4th International Workshop on Knowledge Discovery in Healthcare Data co-located with the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019) (pp. 35-42).
- Downloaded from the UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/datasets/Bar+Crawl%3A+Detecting+Heavy+Drinking

Note: When this proposal is finalized, it will be converted to a PDF and uploaded to a GitHub repo.