A Predictive Model to Detect Alcohol Use Disorder Ph.D. Project

Ph.D. Student: Ali Ebrahimi[1]

Principle Supervisor: Uffe Kock Wiil[2]

Co-Supervisor: Marjan Mansourvar[3]

Faculty of Engineering, Department of Health Informatics, The Maersk Mc-Kinney Moller Institute

[1]aleb@mmmi.sdu.dk, [2]ukwiil@mmmi.sdu.dk, [3]marm@mmmi.sdu.dk

Introduction

Alcohol use disorder (AUD) is associated with various diseases and it carries a high prevalence rate among Western populations particularly in Europe [1]. Excessive alcohol use leads to numerous diseases such as liver cirrhosis, chronic pancreatitis, upper gastrointestinal cancers, cardiomyopathy, polyneuropathy and dementia. Alcohol is also a contributing cause for many road accidents [2]. Recent studies have shown that among the Danish population, 20 percent are heavy drinkers and 14 percent are harmful alcohol users [2]. This indicates that AUD may cause shorter life expectancies among the Danes when compared to the general population. From their study, Westman, Wahlbeck [3] note that the mortality rate in Denmark was higher than other European countries with similar populations such as Sweden and Finland. Consequently, it was deduced that approximately 5 percent of all deaths in Denmark can be attributed to alcohol use, a common scenario in most Western countries [2, 3].

AUD-associated diseases have a poor prognosis in Nordic countries like Denmark as well as other developed and under-developed countries. To overcome this, a predictive machine learning model can be used to detect information based on the patient's medical records. This information can offer insights into the patient's lifestyles, which affect his/her medical conditions. Using this information, medical staff like physicians can become actively involved with their patients to understand their needs and to be able to discuss how treatments best can be arranged for the patients. Some earlier works such as [4, 5] have considered using predictive models to detect AUD, but no prior study had proposed or developed a machine learning model that can predict AUD by scanning through patients' data and reports which are stored in free-texts and structured fields in Danish.

To address this gap, this study aims to develop a predictive model by using machine learning algorithms to analyse patients' structured data and their free-text clinical reports. This model can help medical staff members to gain a better overview of patient data by maintaining caution when the patient's recorded data indicates a high probability of a complex alcohol relationship. The nucleus of this proposed project will be based on two groups of patients, those who were screened and tested positive for AUD and those who were tested negative. The project is developed from an ongoing study called Relay [6] which systematically collects data on diet, smoking, alcohol, and exercise from more than 5,000 patients in selected somatic wards at Odense University Hospital.

Aim

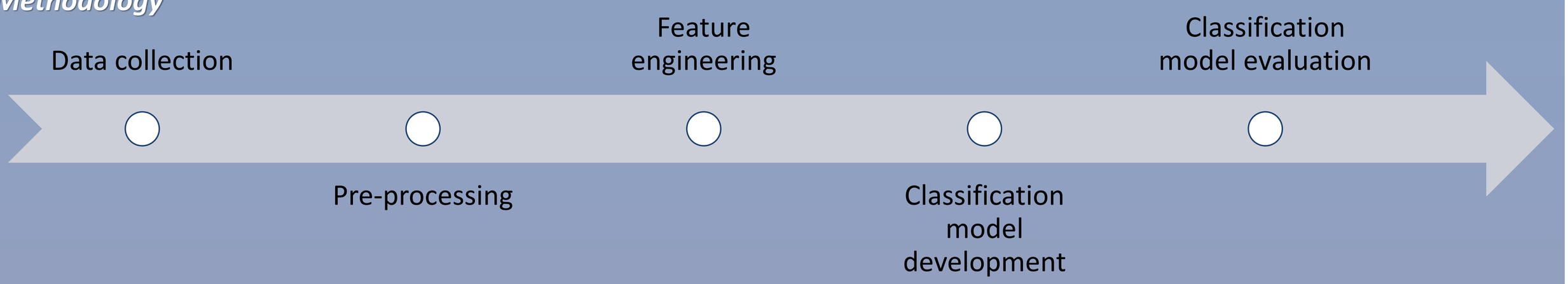
The primary aim of this project is to improve patient treatment by developing a predictive model that can read medical records in hospital information systems thereby, allowing health or medical staff members to suggest a better treatment for their patients. This kind of support can help the staff to utilise the information, which are extracted from the huge amount of compiled data kept in the electronic system of the patient's record. The model is expected to be able to read both the numerical data (test results, etc.) as well as the free-text data.

Objective

The core objectives of this research are:

- To improve the usability of data collected in the electronic patient record systems which are essential for patient treatment (and prognosis).
- To develop a predictive model which can indicate AUD.
- To evaluate the developed model in order to assess efficacy as well as societal cost impacts and other factors.

Methodology



The research methodology proposed for this study encompasses five phases:

i. Data collection: Based on an ongoing study [6] that systematically collects data on diet, smoking, alcohol, and exercise from more than 5,000 patients in selected somatic wards at Odense University Hospital.
ii. Pre-processing: The phase of removing meaningless data from the collected dataset.
iii. Feature engineering: The heart of this study in which the discriminative features will be designed so as to predict AUD. The feature engineering phase will also comprise three sub-steps including feature extraction, feature representation, and feature selection.
iv. Classification model development: Supervised machine learning based algorithms including generative and discriminative supervised machine learning based algorithms will be used to build the AUD prediction model.
v. Classification model evaluation: Performance metrics including precision, recall, F-measure, and accuracy will be applied to evaluate the classification model.

Publication and dissemination

The core of this project has a high potential for several publications. The results of the project will be presented as a publication in an international peer-reviewed journal, acknowledged in the field. The publication will be made openly accessible to facilitate the knowledge transfer.

The required institute work for this project accumulates to 300 hours of knowledge dissemination, passing 30 ECTS courses and participation in active research environments. The plan is to execute the majority of these in 2018 and 2019 as an instructor/co-teacher in various BSc/MSc courses in Health Informatics, attending international conferences and scientific courses.

References

1. Naik, A., D. Rozman, and A. Belič, SteatoNet: The First Integrated Human Metabolic Model with Multi-layered Regulation to Investigate Liver-Associated Pathologies. PLOS Computational Biology, 2014. 10(12): p. e1003993.
2. Gottlieb Hansen, A.B., et al., The number of persons with alcohol problems in the Danish population. Scandinavian Journal of Social Medicine, 2011. 39(2): p. 128-136.
3. Westman, J., et al., Mortality and life expectancy of people with alcohol use disorder in Denmark, Finland and Sweden. Acta Psychiatrica Scandinavica, 2015. 131(4): p. 297-306.
4. Rodríguez, J.A.G.D.C., et al., Predictive models of alcohol use based on attitudes and individual values. Journal of drug education, 2013. 43(1): p. 19-31.
5. Rosenström, T., et al., Prediction of alcohol use disorder using personality disorder traits: a twin study. Addiction, 2018.
6. Schwarz, A.-S., et al., Relay model for recruiting alcohol dependent patients in general hospitals-a single-blind pragmatic randomized trial. BMC health services research, 2016. 16(1): p. 132.

University of Southern Denmark