Proposal:

Automated diagnosis of coronary artery disease using electrocardiograms

Jiachen Guo, Ashwin Vazhayil

Date submitted: 2021.07.15

Mechanistic Data Science Tutorials, Training and Research Experience

Automated diagnosis of coronary artery disease using electrocardiograms

Jiachen Guo Ashwin Vazhayil

Abstract: Coronary artery disease (CAD) is one of the major causes of death around the world. The diagnosis of CAD is of paramount importance for early prevention and treatment. In current clinical practice, electrocardiograms (ECG) have been widely used as a non-invasive method to detect CAD. However, the accuracy of diagnosis is highly intuitive and dependent on the physicians' training and experience. As a result, designing an automated diagnosis system of CAD can be very beneficial. In this project, the idea of mechanistic data science is used to design such a system. Heart rate variability (HRV) will be extracted from ECG signals. Feature engineering will be conducted on the time domain, frequency domain, and time-frequency domain. The support vector machine (SVM) algorithm is employed to classify normal subjects and CAD patients.

Keywords: coronary artery disease; electrocardiogram; heart rate variability; mechanistic data science; support vector machine

1. Introduction

Heart disease is a leading cause of mortality in the United States [1]. Coronary artery disease (CAD) is the most common type of heart disease, killing 365914 people in 2017 [2]. It creates distress on the arteries by creating levels of stenosis, cholesterol buildup, which leads to restricted blood flow and, in many cases, death. With current medical procedures, cardiac electrocardiograms (ECG) have been widely used in medical diagnosis. However, diagnosing CAD based on ECG depends on the subjective interpretation by the reading physicians, and the accuracy is highly related to the physician's training and knowledge [3][4]. Meanwhile, ECG signals in the time domain and frequency domain may be insufficient to make accurate judgments since the nonlinear characteristics are not included [5]. As a result, time-frequency analysis can be carried out to extract more important features. Therefore, an automated ECG diagnostic system that considers the features from both the time domain and time-frequency domain can be advantageous in clinical practice to facilitate the diagnosing process.

2. Objective

In this project, an automatic diagnostic system for CAD will be developed based on analyzing signals from patients' HRV curves. With this system, one can directly give diagnoses automatically without any experience in the field of cardiovascular disease.

3.Approach

3.1 Data collection

In this project, the open database provided by PhysioNet will be used [6]. This database contains many available ECG raw signals which have been labeled by experienced physicians. The ECG signals used include 40 normal subjects and 7 CAD patients.

3.2 Extraction of mechanistic features and feature engineering

The timing of the heartbeat is controlled by the heart's electrical system. As a result, if the abnormality is found in the cardiac electrical system, the heart may not function properly. Therefore, HRV can be treated as mechanistic features to uncover the heart's functionality. In the frequency domain, power density distribution will be used to characterize autonomic functionality. For example, the low frequency (0.04~0.15Hz) power and high frequency (0.15~0.40Hz) power are usually reduced for CAD patients. Meanwhile, such differences can also be observed in the time-frequency domain. As a result, with the aim of extracting useful features to distinguish CAD patients from healthy people, the features from the time domain, frequency domain, and time-frequency domain will be systematically extracted.

3.3 Classification

With the obtained input features, supervised machine learning algorithms like support vector machine (SVM) will be used to characterize the relationship between the input features and the diagnosing result.

3.4 System and design

With the aim of achieving an automated diagnosis of CAD using ECG signals, a mini-app will be provided as the outcome of this project. In this mini-app, one can easily upload raw ECG signals, and the program will give the corresponding diagnosis result. Therefore, this mini-app will highly facilitate the CAD diagnosing process in clinical practice.

4. Conclusion

In this project, an automated diagnostic program will be developed to diagnose patients with coronary artery disease both efficiently and accurately. Electrical signals obtained through electrocardiograms will be used as the raw data to uncover the functionality of the heart. The derived HRV signals will be analyzed in the time domain, frequency domain, and time-frequency domain. Support vector machine will be used to find the hidden relationship between input features and the diagnosis result.

Reference:

[1] Centers for Disease Control and Prevention. Underlying Cause of Death, 1999–2018. CDC WONDER Online Database. Atlanta, GA: Centers for Disease Control and Prevention; 2018.

[2] Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2019). Heart disease and stroke statistics—2019 update: a report from the American Heart Association. Circulation, 139(10), e56-e528.

[3] Leisy, P. J., Coeytaux, R. R., Wagner, G. S., Chung, E. H., McBroom, A. J., Green, C. L., ... & Sanders, G. D. (2013). ECG-based signal analysis technologies for evaluating patients with acute coronary syndrome: A systematic review. Journal of electrocardiology, 46(2), 92-97.

[4] Morrison, W. G., & Swann, I. J. (1990). Electrocardiograph interpretation by junior doctors. Emergency Medicine Journal, 7(2), 108-110.

[5] Sridhar, C., Acharya, U. R., Fujita, H., & Bairy, G. M. (2016, October). Automated diagnosis of Coronary Artery Disease using nonlinear features extracted from ECG signals. In 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 000545-000549). IEEE.

[6] Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. circulation, 101(23), e215-e220.