Justin Amevor, Matthew Kaminski, Wai Khumwang, Peter Subacz

Analyzing different machine learning models to detect Arrhythmia

Arrhythmia is an extremely common health problem that affects over 3 million people annually in the United States. Patients with arrhythmia will have too quick, too slow, or irregular pattern heartbeat. Arrhythmia can be dangerous because it shows little to no symptoms yet can have significant impacts on a person's health from mild seizures to heart attacks. This project will compare various "state-of-the-art" machine learning algorithms on detecting Arrhythmia.

Experienced doctors can detect arrhythmia using common medical instruments however, various consumer products like smartphone applications, blood pressure monitors, and smart-watches have been designed to detect symptoms of the condition. Despite these technologies, arrhythmia detection can still be difficult because of the complexity of heart rate patterns. Arrhythmia most commonly is detected using an electrocardiogram (or ECG), which measures the electrical impulse waves that travel through the heart. Heartbeat ECG waves are non uniform in nature and can be separated into J, P, Q, R, R', S, S', and T waves. The multiplex characteristic of detecting arrhythmia demonstrates the ever growing importance of machine learning in the medical field, which can more accurately help detect and diagnose this condition.

Our team plans to use data analysis techniques such as neural networks, support vector machines, and bayesian probability to determine the best method to detect arrhythmia and reduce false positives during diagnosis. The team plans to use common statistical methodologies including accuracy, performance, pattern recognition

etc. to analyze the various algorithms in order to determine the most suitable implementation for the detection agent.

The data source to be used with this project will be provided by the University of California, Irvine, which contains 452 patient records that described by 279 attributes. Most of the attributes are the ECG measurement from various electrodes, i.e. wavelength, wave interval. This data set was chosen because of its strong variation amongst the data points.

Data source : http://archive.ics.uci.edu/ml/datasets/Arrhythmia