Project description Atrial Fibrillation (raw data)

Data Science Course

Authors: Kicky van Leeuwen & Mannes Poel

Project Introduction

Atrial fibrillation (AF) occurs as a complication postoperatively form cardiac surgery. AF results in stasis of the blood. In the postoperative period AF can induce delirium and neurocognitive decline, thereby prolonging the hospital stay. [1] On the long term serious complications like thromboembolic diseases, stroke and heart failure can be induced by AF. These complications result in increased morbidity and mortality and prolonged hospital stays. [2-7] Precise ECG monitoring is important to detect AF as soon as possible. Then complications caused by AF can be obviated due to a fast intervention.

The challenge of this project it is to develop an algorithm/method that can detect automatically episodes of AF (minimum of 30 seconds) from (preprocessed) ECG data. Framing it differently, the research questions is: "To what extend can one automatically detect episodes of AF?"

Background

Manual detection of AF in ECG record is time-consuming, especially in the case of large datasets consisting of 24-hour ECGs. When automating the detection, the physician can be deprived of work and research can be accelerated. Also, such an algorithm may result in the direct detection of AF during ECG monitoring, thereby creating the possibility for a fast treatment of AF. This underlines the need for an algorithm to automatically detect AF for analysis purposes. Automatic AF detection provides a faster analysis of long term ECGs. Hereby opportunities arise for better diagnostics and for gaining more insight into postoperative AF on a larger scale. Automatic quantification of AF may help to get insight in the yet unsolved underlying problem of AF.

AF is defined as a period of at least 30 seconds in which an irregular ventricular rate and P peaks are absent. [8] These two ECG characteristics indicate the rapid abnormal atrial activity seen in AF. An AF detection algorithm based on R-R interval irregularity is preferred, due to the prominence of QRS complexes, making it more robust to noise. [9] Therefore this algorithm is also based on the R-R interval irregularity.

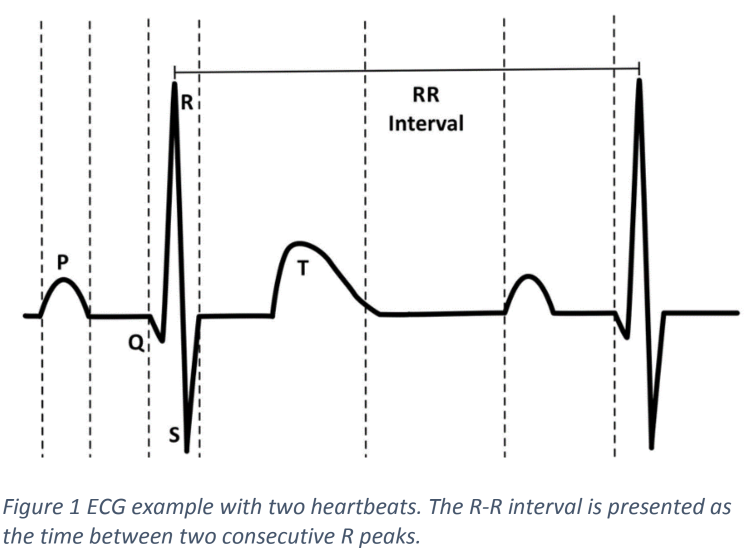

Figure 1 ECG example with two heartbeats. The R-R interval is presented as the time between two consecutive R peaks.

The Data

In this project you will work with preprocessed ECG data from the Erasmus Medical Centre in Rotterdam of the department electrophysiology. Data was obtained within 10 days post-operatively of CABG surgery. Atrial fibrillation (an arrhythmia) occurs frequently after cardiac surgery. Atrial fibrillation is an arrhythmia that does not have to sustain constantly, it comes and goes, and therefore often passes by unnoticed in the patient’s hospital stay.

ECGs

From the patient a 12 lead ECG is recorded. A semi-automatic program (Synescope) analyzed the ECGs for annotation of the R peaks (see green dots in Figure 2). R peak detection was manually audited by a physician and atrial fibrillation or other arrhythmias were labeled.

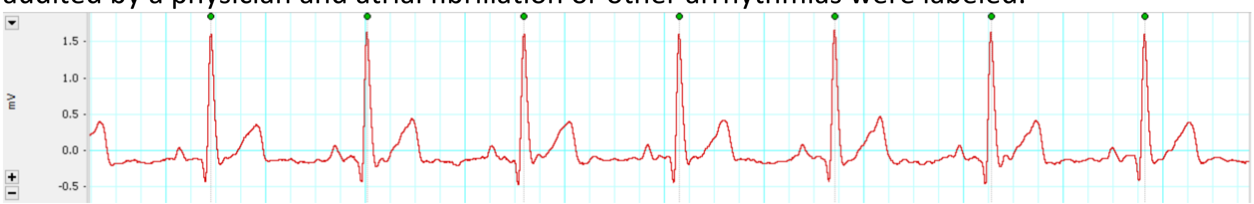

Figure 2 Example ECG. The green dots annotate the R peaks

From these R peak data text files are formed named Data1 to 804 which can be found in the ECG_data folder (see Figure 3 for an example). The time is presented in the first column. The second column contains the R-R intervals in milliseconds. The third column holds letters that refer to the type of heartbeat. This has been automatically annotated by the Synescope program and is not considered reliable information. The fourth column consists of annotations made by the physician. PAUSE is used to annotate missing data (you can see that here R-R intervals are unphysiologically large). PAUSE has been manually placed. With the automatic detection you want to make this manual step redundant, so try to avoid using this information and detect these artifacts yourself with the R-R intervals.

23:56:21	845	N
23:56:22	845	N
23:56:23	855	N
23:56:24	845	N
23:56:25	840	N
23:56:26	840	N
23:56:26	845	N
23:56:27	830	N
23:56:28	825	N
23:56:29	825	N
23:56:30	835	N
23:56:31	815	N
23:56:31	815	N
23:56:32	820	N
23:56:33	830	N
23:56:34	830	N
23:56:35	835	N
23:56:35	835	N
23:56:36	840	N
23:56:37	850	N
23:56:38	850	N
23:56:39	835	N
23:56:40	830	N

Figure 3 Example of text file that is the result of the processed ECG. In the first column record time, second column R-R interval in milliseconds and in the last columns annotations regarding the rhythm have been made.

Class

You have two options for knowing the class of the ECG data. Either per R peak or per 30 seconds episode. You can choose either option.

1. **Per R peak**: In the ECG data described above (Figure 3), annotations can be found of the beginning and end of an episode of AF: 'START AF' and 'END AF'. Note, these are manually typed annotations and spelling mistakes occur. If no 'START AF' is present, but an 'END AF' is, then from the beginning of the file to the first 'END AF' is AF. The other way around is also true, so if the 'START AF' is not closed off with an 'END AF' in the same file, it continues to the end of the file.
2. **Per 30 second episode**: There are also text files named Control from 1 to 804 which can be found in the Class folder. These are processed versions of the START AF and END AF label. Here episodes of AF and no-AF episodes have been transformed to 1 and 0 respectively. The timeline was divided into periods of 30 seconds. If more than 75% of a period was labelled as AF, the 30 second period was labelled as AF (thus 1). When -1 is stated the data for that period could not be transformed due to for instance a measurement error and no class could be given.

Report

Instructions and requirements for the final report can be found in the course guide for Big Data for eHealth.

References

[1] Alqahtani AA. Atrial fibrillation post cardiac surgery trends toward management. Heart Views. 2010 Apr 1;11(2):57.

[2] W.H. Maisel, J.D. Rawn, and W.G.Stevenson. Atrial Fibrillation after Cardiac Surgery; Review. *Annals of Internal Medicine*, 135(12):1061–1073, 2001.

[3] N. Echahidi, P. Pibarot, G. O'Hara, and P. Mathieu. Mechanisms, prevention, and treatment of atrial fibrillation after cardiac surgery. *Journal of the American College of Cardiology*, 51(8):793–801, February 2008.

[4] N.S. Peters, R.J. Schilling, P. Kanagaratnam, and V. Markides. Atrial fibrillation: strategies to control, combat, and cure. *Lancet*, 359(9306):593–603, February 2002.

[5] S.M. Narayan, M.E. Cain, and J.M. Smith. Atrial fibrillation. *Lancet*, 350(9082):943–50, September 1997.

[6] S.S. Chugh, R. Havmoeller, K. Narayanan, D. Singh, M. Rienstra, E.J. Benjamin, R.F. Gillum, Y.H. Kim, J.H. McAnulty, Z.J. Zheng, M.H. Forouzanfar, M. Naghavi, G. Mensah, M. Ezzati, and C.J.L. Murray. Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study. *Circulation*, 129(8):837–847, 2014.

[7] M.P. Turakhia, M.D. Solomon, M. Jhaveri, P. Davis, M.R. Eber, R. Conrad, N. Summers, and D. Lakdawalla. Burden, timing, and relationship of cardiovascular hospitalization to mortality among Medicare beneficiaries with newly diagnosed atrial fibrillation. *American Heart Journal*, 166(3):573–580, 2013.

[8] V. Fuster, L.E. Ryden, R.W. Asinger, D.S. Cannom, H.J. Crijns, R.L. Frye, J.L.

Halperin, and et al. *ACC/AHA/ESC Guidelines for the Management of Patients With Atrial Fibrillation: Executive Summary A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee*, volume 104. 2001, ISBN: 8006116083.

[9] N. Larburu, T. Lopetegi, and I. Romero. Comparative study of algorithms for atrial fibrillation detection. In *2011 Computing in Cardiology*, pages 265–268. IEEE, 2011.