Personalized prediction of migraine attacks using deep learning on weather data

Context

Migraine is a prevalent, multifactorial brain disease, characterized by recurrent attacks of invalidating headaches with or without aura, each attack lasting 1 to 3 days. Migraine affects 1 in 7 adults, predominantly women (80%), and has a severe impact on the lives of those. Migraine is most incapacitating to patients when they are in their productive years, so not surprisingly WHO classified migraine among the top 5 most disabling disorders in people below the age of 50. For the Netherlands alone yearly costs of migraine to society are estimated at €1.7 billion due to costs for treatment and lost work time. Especially the unpredictability of migraine attacks has a major impact on the lives of patients and families.

Migraine patients strongly express the wish for measures to objectify the triggers that influence their susceptibility for attacks and for obtaining early markers for a lowered threshold for upcoming attacks so they gain better control over their lives. Personalized interventions focusing on combined triggers are expected to reduce the susceptibility to attacks.

Therefore, much research is done to better understand triggers for migraine. A commonly reported trigger for migraine is (changes in) the weather and it is frequently reported as possible trigger both in the literature and in clinical practice. Multiple studies have investigated the relationship between migraine attacks and the weather, but the results remain inconclusive. It is particularly difficult to study this association using traditional statistics because of the enormous number of available meteorological variables, the non-specificity of triggers, the inter-patient differences, and the need for large quantities of 'headache diaries', i.e., daily headache 'scores' as reported by migraine patients.

In the context of the LUMINA (Leiden University Migraine Neuro-Analysis) research program, headache e-diary data were collected since August 2018 from over 2,000 clinically diagnosed migraine patients in the Netherlands. Patients are recruited at the LUMC headache clinic and the LUMINA research website. In the LUMINA database we have access to the headache diaries containing daily reports on the presence of headache, medication use and, if applicable, the headache characteristics. In addition, a wide variety of daily (or hourly) weather data can be freely accessed at the KNMI (Koninklijk Nederlands Meteorologisch Instituut) website.

Objectives

This project has two complementary objectives:

1. To build a predictive model based on both LUMINA headache diaries and weather using temporal models (e.g. recurrent neural networks) taking into account the inter-patient variability so that it ultimately becomes possible to predict migraine attacks for individual migraine patients.
2. Using explanatory AI methods to gain insight into the NN predictions and identify weather factors showing associations with an elevated migraine incidence, ultimately leading to a better understanding of the relationship between migraine and the weather.

There are substantial challenges in this project. One of the main challenges will be the personalized approach that will need to be taken due to the inter-patient differences. Besides this, the temporal aspect of the problem will require innovative approaches. Finally, the diversity of the input data adds an additional challenge.

Approach

The LUMINA database and KNMI weather data together provide a very rich source of temporal data that provides both ample opportunities and challenges. This project continues on research of an earlier bachelor thesis which implemented non-temporal classical machine learning methods. As such, the data has already been preprocessed and formatted to a large extend. Once the student has an adequate understanding of the problem and the data, the main task will be to build a (hopefully accurate) predictive model using RNNs. The performance will be assessed by comparison to a baseline predictor.

Student profile

Excellent knowledge and understanding of machine learning and data mining problems and algorithms is essential; programming skills in Python and experience with data science packages (pandas, sklearn, keras) are required; interest in applying data science in the medical domain. Hands-on experience with setting up a novel data science project is a plus. Good communicative skills are key.

Collaboration & supervision

This project is a collaboration with the group of Prof. dr. Gisela Terwindt at the Department of Neurology at the Leiden University Medical Center (LUMC). The main supervisor of this project will be Dr. Matthijs van Leeuwen (LIACS), daily supervision will be by Drs. Hermes Spaink (LUMC/LIACS). You will be recuired to take an active part in both research groups and get the opportunity to collaborate with and present your work regularly to both a data science and clinical audience. Moreover, you will get the opportunity to work at the LUMC in an interdisciplinary group where you work on applying state-of-the-art CS/AI to solve real-world clinical problems.

Contact

For any questions about this project please contact Hermes Spaink at: h.a.j.spaink@lumc.nl.