L-Labs (2023), **confidential**, 1–1

ARTICLE TYPE

Migraine Prediction

Abstract

In the project, we provided a solution for predicting migraines' occurrence and intensity using sparse sensor data from large populations and individual patient data. The approach involves a multi-scale time-series prediction model, collaboratively refined with clinical teams for data preparation, including missing data handling and feature engineering. The solution encompasses various binary prediction tasks for population subsets, with refined model configurations. New users receive a hybrid model, blending a general baseline with personalized versions. An ensemble model combining baseline and personalized models is used for users with historical data. Implementation and testing reveal significant improvements in treatment success rates for controlled and deployed user groups. The latter phase shows increased device usage, indicating user preference for the proposed predictive model.

The Problem

Predicting the occurrence and intensity of migraines using sparse sensor data from large-size populations ($\sim 100,000$ with ~ 70 samples each) and personalized data from the patient itself. The data is gathered using a medical device that also provides a treatment, which causes inherent in the user's clinical dynamics.

The Solution

The proposed model takes advantage of multi-scale time-series prediction model. First, multiple iterations with the client's clinical team allowed us to prepare the client's data to be used as part of an algorithmic solution, including missing data filling, feature selection, feature engineering, etc. Second, for the entire population, we define several binary prediction tasks that differ by the duration of prediction and the amount of data they have to make the prediction. Based on these results, we allocate the better model configurations and study them. Afterward, we used these models, after further refinement, as the baseline model for new users. However, for users with some historical use, we over-fitted our model on the user's data alone. With a large number of such over-fitted personal models, we trained an ensemble model that gets the baseline (general) model and the personalized one and computed the best mixture of the two. At this point, our solution does not take into consideration its own concept drift in the users' behavior and therefore clinical state. Hence, finally, the client deployed the proposed model for small-size user groups and collected more data. This data is used to refine the proposed model with a robustness metric in mind.

The Outcome

The proposed solution was implemented and tested in two phases. First, a control and case group of 150 users are used to compare the treatment success rate, as reported by the users, over three months. The results of this experiment reveal statistically significant improvements ($p < 0.01$) with an average increase of around 12%. Second, deployment of the proposed solution for 20,000 users is provided and evaluated for three-month. Following this duration, a test to measure the re-usage of the device with and without the prediction model revealed an increase of 6% of the device usage which indicates the users like it better.