The Iby and Aladar Fleischman
Faculty of Engineering
Tel Aviv University

Pulse waves Scorer

By: Noa Gelber, Amit Gavish, Amit Hadar
Advisors: Lior Toledano & Sam Zibman
Project Number: 22-0-251

Avertto™
IF YOU CAN'T SEE IT- YOU CAN'T TREAT IT

Introduction

Avertto's goal is to develop and provide systems to handle strokes within at-risk populations. Their vision is to improve life quality by reducing the impact of a silent stroke. Avertto will develop, create, and introduce the first easy-to-use stroke warning device for high-risk populations. The company targets two main markets;

1. The hospital market - use of the sensor during operations with a risk of a blood clot drifting. these may create a stroke while the patient is anesthetized.
2. The private market - people who have experienced a stroke, have a 30% chance of having another stroke in the following 2 months, and a 10% chance in the following year. The uniqueness of Avertto's product is the ability to detect the clinical hemodynamic changes in anesthetized people, and thus warn of the formation of a stroke and give treatment in time.

Motivation & Goals

In order to detect the clinical hemodynamic changes, Avertto will use a model that can recognize and differentiate a standard pulse. To do this, it must be ensured that the model will be trained on quality data, with signals that meet defined criteria and are free of noise - and this is where we are stepping in.

Our Project's main goal, is to create a deep learning model that can determine the quality of the data. This model will selectively funnel high-quality data to the main model, allowing it to refine its training solely with clean and relevant data. Furthermore, our model will handle with the problem of receiving bad signals because of bad placement of the sensor.

We strive to refine the ability to identify the signal and its quality, in order to give a real-time feedback on the quality of the data the sensor receives, so it can alert the user if something is not right.

Methodology

Collecting The Data: Avertto collected data for us of patients from Val d'Hebron Barcelona, Beilinson Hospital and Galilee Medical Center in Nahariya.

Tagging The Data: We learned the characteristics of a pulse wave, and labeled our data using CVAT - according to the quality of the wave and the characteristics we learned.

Building The Baseline: Creating a work environment suitable for our data type and the models we plan to run. Using Python and Pytorch library, we built 2 collab notebooks (binary classification & 3 classes classification) which contain all the necessary steps of the process, from importing the data to building and evaluating the models.

Deep Learning Models: We selected several DL models and trained them on our data. Each time, we tested different parameters and examined the performances. With the help of the Sklearn library, we evaluated the results and made a comparison between the different experiments.

WorkFlow

1. Tagging an initial amount of data, and creating a labeled data base.
2. Searching online for the best methods and models to work with this kind of data.
3. Building two types of Baselines - one for 3 classes classifying, and one for binary classifying.
4. Running the different DL models, evaluation and drawing insights.
5. Tagging more data and increasing the database.
6. Testing our best-performing models again, this time with the newly added data, to see how it affects their performance.
7. Evaluating the best model and drawing final conclusions.

Binary Classification
Model: CNN_LSTM
Loss Function: BCELoss
Number of epochs: 100
Score Function: f1
Best validation score: 0.8644

3 Labels Classification
Model: ClassifierCNN
Loss Function: CrossEntropy
Number of epochs: 500
Score Function: f1
Best validation score: 0.7335

Conclusion

Looking at our expiriments' results tracking sheet, we can determine that all three models yield comparable scores. We are aware that the results are not very high, but since the data was very difficult to work with, we think we did our best and achieved good results considering the problematic data. Medical data is something that is very difficult to work with from several reasons.

First of all, the recorded data was significantly affected by high levels of noise, which posed a considerable challenge for the models during analysis. The data was affected not only by noise, but also because of variations in pulse wave patterns between the peaks.

Another problem we had to deal with, was the fact that the tagging process was conducted subjectively, involving four different individuals who tagged different waves. As a result, the reliability of the tagged data may be compromised due to varying interpretations.

Overall, we got to work in the med field - which none of us had ever dealt with before. We explored the topic, read articles and learned a lot. Afterwards, using some of the new techniques we learned, combined with the material we were thought in the ML course, we created the models.

Future Work

- As we said in our conclusion, the data was affected by high level of noise, we think that the data should be recorded from individuals in non-clinical settings, such as those working in a relaxed environment.
- In order to reduce bias arising from multiple taggers, we would suggest to appoint a single individual for tagging the data. Alternatively, can also choose few individuals who will tag the same waves to create some kind of a mean score. Another way to solve the tagging problem is to create rules for grading as strict as possible.
- To take care of the variations between the peaks, we suggest to consider a preprocessing step that involves isolating and retaining only the desired peaks of the pulse waves. This approach will help remove noise and ensure that the data primarily consists of relevant pulse wave information
- If there is no time limit, we can recommend to examine other models and methods, such as Spectrogram and FFT, and other complex DL models who might handle the noisy data by creating a neural network with many layers.