Artificial Intelligence in BrainRT

14th of April 2020

By Sabine Wuytens

Artificial Intelligence to Improve Automatic Hypnogram

As stated in previous newsletters, OSG is working on a vast improvement of the automatic analysis. In Dec 2020, our results for the automatic hypnogram were up to a level of accuracy that allowed us to proceed towards bringing this solution to our customers directly. How? By providing a new AI toolbox that will bring the new AI hypnogram analysis to the RT Software Suite application! This new addition to our product portfolio is expected by 2021.

Using novel techniques to automate our software production, we will be able to provide continuous improvement of our AI toolbox through automatic updates. OSG aims to focus its next improvement on the respiratory analysis.

BrainRT AI results for Automatic Hypnogram

High levels of accuracy: The automatic Hypnogram of our AI Solutions has been evaluated on more than 3000 measurements. We find following results as visualized in the yellow distribution graph below:
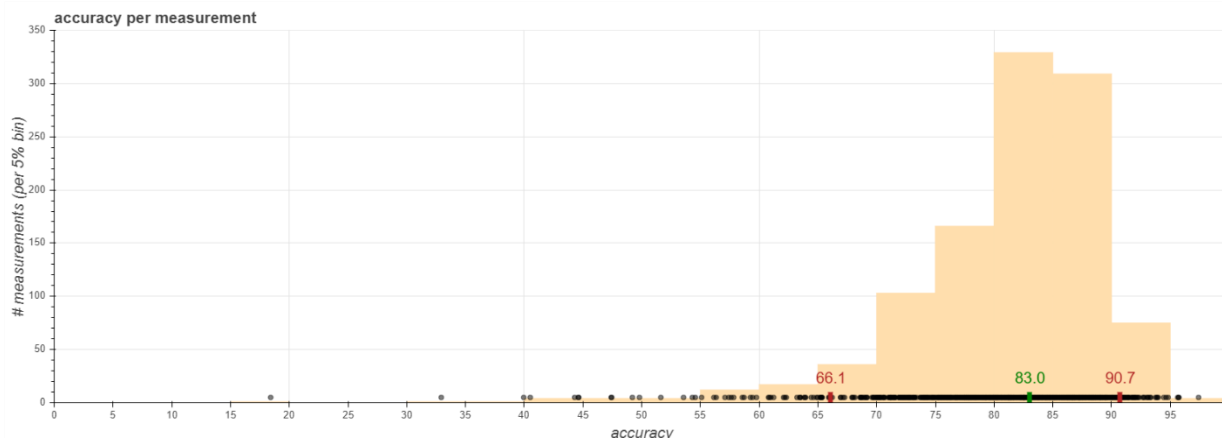
From the graph we can read the distribution of accuracy for individual measurements, where the three values on the graph are the 5th, 50th and 95th percentile respectively. For 95% of measurements, the accuracy is above 66,1%, and most measurements have an accuracy above 75%. The reason for poor results on a limited number of measurements is still under investigation but has various reasons such as very low signal quality and neurological disorders influencing the EEG curves to a level that feature detection becomes impossible. This is part of our Anomaly project, performed in PhD research at OSG.

2

Error distribution: We know that not all errors in sleep scoring are equally problematic. Therefore, in our construction of the AI decision process, we've taken care to reduce Sleep-Wake errors to a minimum, followed by a reduction of REM-NREM errors. This is visualized in the matrix below, where we show manual scorings results in the different rows, and the corrresponding automatic scoring results in the corresponding columns.

Automatic scoring

	Wake	N1	N2	N3	REM	
Wake	23.9	1.0	1.0	0.0	0.5	26.3
N1	1.6	2.7	2.9	0.0	0.8	8.0
N2	0.8	1.2	30.8	2.6	1.1	36.4
N3	0.1	0.0	3.3	12.1	0.0	15.5
REM	0.3	0.3	1.3	0.0	11.8	13.8
	26.6	5.1	39.3	14.8	14.2	**81.3**

(Manual scoring — left axis label)

How to read this matrix correctly? Let's focus on the most important part first, the parts where the automatic scoring is equal to the manual scoring: this is represented by the diagonal line where Wake – Wake, N1 – N1 etc. are located. Note that the sum of all values correspond to 81.3% of accuracy.

Now let's continue to the errors for Sleep-Wake. In the first row, we can see that in total, 26.3% of all stages was manually scored as Wake. 23.9% were correctly scored as Wake by the automatic hypnogram, and the others were scored incorrectly. We can see the errors made by the Automatic scoring in the rest of the row: 1.0 was scored as N1, 1.0% as N2, 0.0% as N3 and 0.5% as REM.

Note that we can also read which stages were incorrectly scored as Wake by the Automatic hypnogram: 1.6% as N1, 0.8% as N2, 0.1% as N3 and 0.3% as REM. In total, this gives us a Sleep/Wake accuracy of over 90% for our Automatic scoring.

We can use the same methods to see that REM/NREM errors are low. Most errors are related to N2/N3 errors, which can be tricky for human scorers as well.

Personalize your AI hypnogram

Throughout the AI hypnogram design process, we've been cautious not to overfit the AI hypnogram on our data, by using a separate test & validation set. This allows us to guarantee that the AI hypnogram will perform equally well on most data sets. However, we also offer the service to provide a personalized model for the AI hypnogram at individual sites, as we know it will outperform the global data set by a few percent of accuracy. Curious to know how this works? Get into contact with us at sales@osg.be to learn more (no charges involved!).