A DIAGNOSTICATION METHOD OF OPENNESS USING THE NON-LINEAR INTEGRALS

Mădălina GIURGESCU (MANEA)[1], Ion CHIȚESCU[2], Titi PARASCHIV[3], Cătălin ȘTEFAN[4]

This article describes a mathematical model through which the level of EEG type waves are processed in order to characterize the level of openness. Our idea is to use the Choquet integral with respect to a monotone measure. We consider the data resulting from the EEG wave measurements for a group of subjects. We describe a procedure by using different monotone measures to calculate the openness level of a subject using the Choquet integral. For each patient we have the level of openness given by psychologists. We compare the results obtained by this method with the results of psychologists. Of all the measures used, we choose the measure that provides the closest results to the real ones.

Keywords: Openness, EEG, Choquet integral, monotone measure, Big Five, C++.

1. Introduction

In this paper we describe a mathematical model through which the level of EEG type waves is processed in order to characterize the level of openness, which represents one of the 5 characteristics of the BigFive model. The goal was to determine a mathematical tool through which to diagnose the level of openness.

In writing this article we have worked in collaboration with the Institute of Studies, Research, Development and Innovation of Titu Maiorescu Faculty in Bucharest, as well as with specialists of the Military Technical Academy in Bucharest.

We considered the data resulting from the EEG wave measurements. The measurements of the values of EEG waves were measured in 14 subjects. In order to carry out the measurements a NeuroSky device, with two sensors, of which one active was used. The specialists in psychology state that the openness is characterized by Delta and MidGamma waves. The input data used were specific values of EEG waves, as well as classification data of the openness level, provided by the Psychology Research Institute. The classification is given by integer numbers from 0 to 100.

As a mathematical procedure, the nonlinear integrals are used as a fusion instrument. We used 8 monotone measures to calculate the openness level of each subject, using the Choquet integral. For each measure, we compared the results obtained by our method with the results of psychologists. Of all the measures used, we chose the measure that provided the closest results to the real ones.

To calculate the values, we have created a C++ programme. The source code is written in C++ in the CodeBlocks development medium, 17.12 version on Windows 10 operating system, combined with GNU GCC Compiler in MinGW distribution, 6.3 version. For the matrix operations the Eigen library, version 3.3 was used.

---

[1] PhD Candidate, University of Pitești, Romania, e-mail: madalina_giurgescu@yahoo.com
[2] Professor PhD, University Politehnica of Bucharest, Romania, e-mail: ionchitescu@yahoo.com
[3] Professor PhD, Technical Academy of Bucharest, Romania, e-mail: titi.paraschiv@siemens.com
[4] PhD, University Titu Maiorescu, Secvento Consulting, Romania, e-mail:catastef_ro@yahoo.com

Finally, we have determined a monotonous measure to provide the closest results in relation to psychological results in terms of openness.

The determined instrument will be used to draw conclusions regarding the level of openness of other subjects who have been measured with NeuroSky.

In the sequel we explain how the nonlinear integral was used for the aggregation of data for the above-mentioned problem, and, of course, we explain the obtained results.

2. Preliminary Facts

Throughout the paper, the positive integer numbers will be $\boldsymbol{N} = \{0, 1, 2, \dots\}$, the real numbers will be $\mathbb{R}$ and the positive real numbers will be
$\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$. As usual, $\overline{\mathbb{R}_+}$=$\mathbb{R}_+ \cup \{\infty\}$. For any set T, the Boolean of T is $\mathcal{P}(T)$=$\{A|A\subset T\}$.
A measurable space is a couple (T, τ), where T is a non-empty set and $\tau \subset \mathcal{P}\,(T)$ is a σ-algebra.
If (T, τ) is a measurable space, a monotone measure is a function $\mu: \tau \to \mathbb{R}_+$ having the properties:
i) $\mu(\emptyset) = 0$
ii) $\mu(A) \leq \mu(B)$ for any A, B in τ such that $A \subset B$.

Now, let us consider a measurable space (T, τ), a monotone measure $\mu: \tau \to \mathbb{R}_+$ and a positive τ-measurable function $f: T \to \mathbb{R}_+$.
For any $a \in \mathbb{R}_+$, we consider the inferior level set F_a=$\{t \in T\ |f(t) \geq a\}$ =
$f^{-1}([a, \infty)) \in \tau$. Because $F_a \supset F_b$, whenever $0 \leq a < b < \infty$, it is seen that the function $\varphi: \mathbb{R}_+ \to \mathbb{R}_+$, given via $\varphi(a) \overset{\text{def}}{=} \mu(F_a)$, is decreasing.
Considering the Lebesgue measurable sets of $\mathbb{R}_+$ which we denote by $\mathcal{L}$ and the Lebesgue measure on $\mathcal{L}$ which we denote by L: $\mathcal{L} \to \overline{\mathbb{R}_+}$, we can compute $\int \varphi \mathrm{dL} \in \overline{\mathbb{R}_+}$, because φ is $\mathcal{L}$- measurable. For the sake of concreteness and respecting the traditional notations, we shall write $\int_0^\infty \mu(F_a)\, da \in \overline{\mathbb{R}_+}$ instead of $\int \varphi \mathrm{dL}$.

DEFINITION: The Choquet integral of the function f with respect to the measure μ is the element $\int_0^\infty \mu(F_a)\mathrm{d}a \in \overline{\mathbb{R}_+}$. We shall write (C)$\int f \mathrm{d}\mu$=$\int_0^\infty \mu(F_a)\, da$.

We shall say that f is Choquet integrable with respect to μ in case (C)$\int f \mathrm{d}\mu < \infty$.
The Choquet integral is a generalization of the abstract Lebesgue integral. Namely in case μ is a classic measure (i.e. μ is σ- additive), we can see that the Choquet integral (C)$\int f \mathrm{d}\mu$ coincides with the abstract Lebesgue integral $\int f \mathrm{d}\mu$.
We have special formulae for the computation of the Choquet integral in case the function f is simple, i.e. f has the form

$$f = \sum_{i=1}^{n} a_i\, \varphi_{A_i}$$

where $a_i \in \mathbb{R}_+$, $A_i \in \tau$ are mutually disjoint and $\bigcup_{i=1}^{n} A_i$=T. Here $\varphi_A$is the characteristic (indicator) function of the set $A \subset T$.
Namely, for such f, one can put in order the numbers a_i such that $a_1 \leq a_2 \leq \dots \leq a_n$. Considering that f is in this situation, one has the formula

$$\text{(C)}\int f \mathrm{d}\mu = \sum_{i=1}^{n} (a_i - a_{i-1})\mu(A_i \cup A_{i+1} \cup \dots \cup A_n) \qquad (1)$$

with the convention $a_0 \overset{\text{def}}{=} 0$.
Caution: The reordering of the values a_i is unique in case a_i are distinct and all A_i are nonempty. Otherwise,
different reorderings can occur, but the value of (C)$\int f \mathrm{d}\mu$ given by the formula from above does not depend upon these different reorderings.

We shall be concerned with the case when T is finite, T=$\{x_1, x_2, \dots, x_n\}$, n≥1. Hence, for the function $f: T \to \mathbb{R}_+$, there exists a permutation $\sigma: T \to T$
(we write $\sigma(x_i) = x_i^*$ for any i=1, 2, …, n) such that $f(x_1^*) \leq f(x_2^*) \leq \dots \leq f(x_n^*)$.
In case the values $f(x_1), f(x_2), \dots, f(x_n)$ are distinct, such σ is unique.

Working for the measurable space $(T,\ \tau)$=$(T,\ \mathcal{P}(T))$ and for a monotone measure $\mu: \mathcal{P}(T) \rightarrow \mathbb{R}_+$, the formula$(FF)$ from Preliminary Facts gives

$$(\mathrm{C})\int f\mathrm{d}\mu = \sum_{i=1}^{n}(f(x_i^*) - f(x_{i-1}^*))\mu(\{x_i^*, x_{i+1}^*, \ldots, x_n^*\}) \quad (2)$$

with the convention $f(x_0^*) = 0$.

3. Determination of the Openness Degree

We made l=14 measurements. These are the l =14 functions $f_1, f_2, \ldots f_{14}$. The n=2 measured attributes are $x_1 = Delta, x_2 = MidGamma$.

Namely, for each of the 14 measurements (rows), we obtained the input values $f_p(x_1), f_p(x_2)$ and the output values $y_p, p = 1,2, \ldots ,14$. So, the fifth column contains the output values $y_p, p = 1,2, \ldots ,14$.

The input values are the averages of the measurements carried on the 14 subjects. The output values are obtained using the classification given by the psychologists to the subjects (measured individuals). These output values are represented by grades, from 0 to 100.

In the C++ program, we used a function to process the data from the CSV files, and to create a matrix.

Thus, we obtained the table *Table 1*, with 14 rows and 2+1=3 columns + one column.

In order to save typographical space, we exhibit below only one row of the table:

Table 1

Number of Subject	Delta	MidGamma	Grade
S1	33738.85	26911.79	10

Psychological results:

S1=30; S2=0; S3=10; S4=30; S5=20; S6=2; S7=20; S8=30; S9=20; S10=60; S11=30; S12=40; S13=30; S14=50.

We considered $t = 8$ monotone measures. Using each measure μ_k $(k = 1, t)$, for each subject p, we calculated the level of openness $z_{k,p}$ ($k = 1, \ldots, t$ and $p = 1, \ldots ,14$).As we have said, we decided to choose as fusion instrument the Choquet integral of the functions $f_p, p = 1,2, \ldots ,14$, with respect to a monotone measure μ_k. So, for any $k = 1, t$ and for any $p = 1, 14$ one has: $z_{k,p} = (C)\int f_p d\mu_k$

For each measure μ_k, we compared the results obtained by our method with the results of psychologists, using The Least Squares Method. Actually, for each measure μ_k, we calculated $E_k = \sum_{p=1}^{14}(z_{k,p} - y_p)^2$.

We considered the measures:

$$\mu_1(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.0000143333, if\ E = \{x_1\} \\ 0.0000406178, if\ E = \{x_2\} \\ 0.0000471675, if\ E = \{x_1, x_2\} \end{cases}, \quad \mu_2(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.0000245333, if\ E = \{x_1\} \\ 0.0000507178, if\ E = \{x_2\} \\ 0.00671685, if\ E = \{x_1, x_2\} \end{cases},$$

$$\mu_3(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.000014, if\ E = \{x_1\} \\ 0.00004, if\ E = \{x_2\} \\ 0.000047, if\ E = \{x_1, x_2\} \end{cases}, \mu_4(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.001, if\ E = \{x_1\} \\ 0.00004, if\ E = \{x_2\} \\ 0.002, if\ E = \{x_1, x_2\} \end{cases}, \mu_5(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.03, if\ E = \{x_1\} \\ 0.1, if\ E = \{x_2\} \\ 0.2, if\ E = \{x_1, x_2\} \end{cases},$$

$$\mu_6(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.21, if\ E = \{x_1\} \\ 0.22, if\ E = \{x_2\} \\ 0.3, if\ E = \{x_1, x_2\} \end{cases}, \mu_7(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.00000249806, if\ E = \{x_1\} \\ 0, if\ E = \{x_2\} \\ 0.00000558338, if\ E = \{x_1, x_2\} \end{cases} \mu_8(E) = \begin{cases} 0, if\ E = \emptyset \\ 0.0123, if\ E = \{x_1\} \\ 0.6178, if\ E = \{x_2\} \\ 0.7675, if\ E = \{x_1, x_2\} \end{cases}.$$

For $k=1$, we obtained the conclusions: $z_{1,1}$=16.8855, $z_{1,2}$=0.0000406178, $z_{1,3}$=3.93688, $z_{1,4}$=21.251, $z_{1,5}$=11.8435, $z_{1,6}$=1.36402, $z_{1,7}$=13.8231, $z_{1,8}$=18.3126, $z_{1,9}$=16.3625, $z_{1,10}$=39.0766, $z_{1,11}$=16.8855, $z_{1,12}$=24.3088, $z_{1,13}$=20.7074, $z_{1,14}$=35.1075. And $E_1 = \sum_{p=1}^{14}(z_{1,p} - y_p)^2$ = 1704.34.

For $k=2$, we obtained the conclusions: : $z_{2,1}$=26.2941, $z_{2,2}$=0.0000507178, $z_{2,3}$=6.30102, $z_{2,4}$=32.191, $z_{2,5}$=19.9034, $z_{2,6}$=2.23469, $z_{2,7}$=23.0759, $z_{2,8}$=27.897, $z_{2,9}$=25.0873, $z_{2,10}$=56.8422, $z_{2,11}$=26.2941, $z_{2,12}$=37.457, $z_{2,13}$=32.007, $z_{2,14}$=50.4147. And $E_2 = \sum_{p=1}^{14}(z_{2,p} - y_p)^2$ =106.418.

For $k=3$, we obtained the conclusions: $z_{3,1}$=16.6715, $z_{3,2}$=0.00004, $z_{3,3}$=3.8753, $z_{3,4}$=21.0435, $z_{3,5}$=11.5934, $z_{3,6}$=1.33915, $z_{3,7}$=13.5417, $z_{3,8}$=18.123, $z_{3,9}$=16.182, $z_{3,10}$=38.8559, $z_{3,11}$=16.6715, $z_{3,12}$=24.0279, $z_{3,13}$=20.4613, $z_{3,14}$=34.954.
And $E_3 = \sum_{p=1}^{14}(z_{3,p} - y_p)^2$ =1761.04.

For $k=4$, we obtained the conclusions: $z_{4,1}$=929.931, $z_{4,2}$=0.00004, $z_{4,3}$=233.04, $z_{4,4}$=1084.61, $z_{4,5}$=791.247, $z_{4,6}$=85.6484, $z_{4,7}$=908.82, $z_{4,8}$=949.589, $z_{4,9}$=863.795, $z_{4,10}$=1770.69, $z_{4,11}$=929.931, $z_{4,12}$=1301.01, $z_{4,13}$=1117.72, $z_{4,14}$=1528.61.
And $E_4 = \sum_{p=1}^{14}(z_{4,p} - y_p)^2$ =$1.36172 \cdot 10^7$.

For $k=5$, we obtained the conclusions: $z_{5,1}$=54810.5, $z_{5,2}$=0.01, $z_{5,3}$=11506.2, $z_{5,4}$=75709.6, $z_{5,5}$=27538.9, $z_{5,6}$=3601.57, $z_{5,7}$=33293.6, $z_{5,8}$=64067.6, $z_{5,9}$=56042.4, $z_{5,10}$=156767, $z_{5,11}$=54810.5, $z_{5,12}$=81868.5, $z_{5,13}$=68997.9, $z_{5,14}$=145726. And $E_5 = \sum_{p=1}^{14}(z_{5,p} - y_p)^2$ =$7.82091 \cdot 10^{10}$.

For $k=6$, we obtained the conclusions: $z_{6,1}$=172218, $z_{6,2}$=0.22, $z_{6,3}$=45068.4, $z_{6,4}$=190763, $z_{6,5}$=162903, $z_{6,6}$=17101.5, $z_{6,7}$=185685, $z_{6,8}$=168917, $z_{6,9}$=155572, $z_{6,10}$=283005, $z_{6,11}$=172218, $z_{6,12}$=236493, $z_{6,13}$=204320, $z_{6,14}$=235407.
And $E_6 = \sum_{p=1}^{14}(z_{6,p} - y_p)^2$ =$4.44816 \cdot 10^{11}$.

For $k=7$, we obtained the conclusions: $z_{7,1}$=2.43591, $z_{7,2}$=0, $z_{7,3}$=0.601086, $z_{7,4}$=2.8905, $z_{7,5}$=1.99253, $z_{7,6}$=0.218284, $z_{7,7}$=2.29558, $z_{7,8}$=2.52138, $z_{7,9}$=2.28419, $z_{7,10}$=4.85805, $z_{7,11}$=2.43591, $z_{7,12}$=3.42969, $z_{7,13}$=2.94089, $z_{7,14}$=4.23747. And $E_7 = \sum_{p=1}^{14}(z_{7,p} - y_p)^2$ = 11257.1.

For $k=8$, we obtained the conclusions: $z_{8,1}$=154248, $z_{8,2}$=0.6178, $z_{8,3}$=26824.8, $z_{8,4}$=242427, $z_{8,5}$=29904.7, $z_{8,6}$=6530.32, $z_{8,7}$=43170.6, $z_{8,8}$=200482, $z_{8,9}$=170500, $z_{8,10}$=571771, $z_{8,11}$=154248, $z_{8,12}$=243320, $z_{8,13}$=201948, $z_{8,14}$=548742.
And $E_8 = \sum_{p=1}^{14}(z_{8,p} - y_p)^2$ =$9.06956 \cdot 10^{11}$.

We chose the measure that provided the closest results to the real ones. Actually, we chose the minimum value of E_k. That is E_2.

So, μ_2 is a monotonous measure which provides the closest results in relation to psychological results in terms of openness. This is a good reason to chose μ_2 as a “good” measure for future measurements.

4. Conclusions

The studied level of openness which represents one of the 5 characteristics of the BigFive model., and its values were determined using EEG waves.

The determined instrument will be used to draw conclusions regarding the level of openness of other subjects who have been measured with NeuroSky.

To be more precise, this will be done using the already chosen measure μ_2.

R E F E R E N C E S

[1] *D.E.Hinkle, W.Wiersma, S.G.Jurs*., Applied Statistics for the Behavioral Sciences (5th edition). Houghton Mifflin Company, Boston. New York, 2003.

[2] *E. Niedermeyer*., “The normal EEG of the waking adult”, (E. Niedermeyer, F. Lopes da Silva, eds), 4th edition. Lippincott, Williams and Wilkins, Philadelphia, Pennsylvania, 1999

[3] *B. Stroustrup*., The C++ Programming Language, 4th edition. Addison-Wesley Professional, 2013.

[4] *Z.Wang, G. J. Klir*. Generalized Measure Theory. Springer, 2009.

[5] *Z.Wang, R. Yang, K-S. Leung*., Nonlinear Integrals and Their Applications in Data Mining. World Scientific, 2010.