EXTRACTION OF PARAMETERS FROM DYSGRAPHIC HANDWRITING FOR CDSS SYSTEMS

Zuzana DANKOVIČOVÁ*, Matúš UCHNÁR**
Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice,
E-mail: zuzana.dankovicova@tuke.sk*, matus.uchnar@tuke.sk**

ABSTRACT

In this study we address the issue of the handwriting processing by extracting parameters from the written speech. The work applies machine learning method – the decision trees method which aims to recognize the impaired handwriting, particularly dysgraphia. 55 features (e.g. total time, pen movement, pressure, speed, acceleration) were extracted from each out of 80 handwriting samples while analyzing the performance of classifier for the dominant parameter – minimal speed, and without the dominant parameter as well. The experimental results of the classifier are compared to the results of the statistical test – Mann Whitney U-test as a complex and challenging endeavor to create an accurate classification.

Keywords: *classification, data mining, decision trees, dysgraphia, handwriting.*

1. INTRODUCTION

Clinical decision support systems (CDSS) are information systems or computer applications used in health care to assist clinicians with clinical decisions. The systems themselves are used to process, analyze and retrieve information from available sources. The aim of such systems is to cooperate with a physician and to support him / her in decision-making processes, to prevent mistakes in decision making and to provide early warning [1]. We focus on CDSS for handwriting, since handwriting is non-invasive, inexpensive, and reflects the symptoms of many other disorders such as Parkinson's disease [2] [3], Alzheimer's disease [4], Huntington's disease [5], dysgraphia [6] [7] and many others.

In response to increasing concerns about handwriting disorders, I focus specifically on dysgraphia, since it manifests itself in childhood. The authors of [7] claim that children spend 31% to 60% of their school day by writing and other tasks of fine motoring. Difficulties in writing and the consequences of bad handwriting can therefore significantly disrupt their academic results, and they can even have a serious impact on the emotional state and behavior of the individual. This discovery encourages timely identification of handwriting disorders as well as preventive and remedial assistance.

The goal of this research is to extract parameters from written text obtained with the Wacom graphic tablet in order to perform their processing, comparison and evaluation using the Python programming language. We compare the handwriting of the dysgraphic subjects to the handwriting of people without neurological disease or any other handwriting disorder. We have 80 samples of the handwriting, out of which 39 samples are provided by dysgraphic subjects (test subjects) aged 8 to 19 and 41 by healthy subjects (control subjects) aged 10 to 14. In both groups there is a greater incidence of males than the females, and the right-handed over the left-handed.

The paper is organized as follows. In the next section we briefly describe similar works, focused on dysgraphia and handwriting processing. Then we describe preprocessing of the dataset and handwriting analysis. Later, we provide briefly overview of classification models and present the experimental resuls. Finally, some conclusions are drawn.

2. RELATED WORK

There are several existing studies focusing on analysis and handwriting processing [8] [9] [10].

For example, Mekyska et. al in their work [11] describes a method that can be used for automated diagnostics of difficulty in writing text. In the study to suppress individuality in the handwriting and to emphasize the signs of dysgraphia, they introduced a simple method of normalization based on subtraction [12]. They found that children with dysgraphia are unable to hold the pen in a stable position and therefore the pen slope varies greatly over time. They were able to design a dysgraphic evaluation system with a HPSQ (Handwriting Proficiency Screening Questionnaire) [13] error estimate of about 10%.

Rosenblum and Dror in their work [6] use data mining methods to derive a statistical model that is able to distinguish the dysgraphic handwriting from an experienced handwriting with an accuracy of about 90% based on their performance characteristics.

Paz-Villagran et al. in their work [14] compared in the same written assignment 26 adults and children of the third year of elementary school – 39 proficient children and 16 dysgraphic children. The aim of the study was to focus exclusively on low-level graphomotor processes. The result of this study is that stopping during writing is more appropriate than an identifier, as opposed to pen strokes when differentiating the fluency of the handwriting to identify dysgraphy.

3. PROCESSING OF RECEIVED DATA

3.1. Dataset

The use of tablets is pervading almost all areas in our complex society. In order to collect the data, we use the

graphic tablet – the Wacom Intuos Pro Large and the template with the handwriting tasks created by us. We received 80 samples of the handwriting, out of which 39 samples belong to patients with dysgraphia (27 men, 12 women) and 41 samples belong to healthy ones (26 men, 15 women). The age of all subjects is between 8-14 years old. We created the template based on the previous research [6] [14] [15] [16].
The created template with the handwriting tasks consists of repeating the letter *l*, the syllable *le*, the words like *leto, lamoken, hračkárstvo* and the sentence *V lete bude teplo a sucho*.

3.2. Handwriting analysis

The collected samples are processed by the Python programming language. It is necessary to calculate the individual parameters for each sample separately and then to compute the calculations. Then the results are analyzed and the preliminary outcomes of the sample are determined. Table 1 shows parameters of the calculated sample. Samples no. 1-5 are obtained from the dysgraphic subjects and samples no. 8-12 are obtained from the healthy subjects. As expected, the results differ significantly, for example in the total time of the writing and in the average writing speed as well.
We also use classification models for analysing and processing the received data. In more detail they are described in the section 4.

4. CLASSIFICATION MODEL

Currently, there are multiple classification algorithms. Within this study we decided to use the classification model – decision trees.

4.1. Decision tree method

Decision tree method is most commonly used for data classification. It is one of the most preferred techniques in the data mining processes. The most important decision tree criterion is the choice of the most appropriate classifier. It is able to distinguish the objects from one another and divide them into different branches. To select the correct classifier, the statistical property – entropy (information gain) is used, which indicates the rate of division of the object into the target classification. Entropy is calculated according to the relation 1, where S is the set containing all possible states, and p_i is the ratio of the set S, falling into the class i.

$$Entropy(S) = \sum_{i=1}^{n} -p_i log_2 p_i \tag{1}$$

In general, the method of the decision trees applies multiple different algorithms. The most well-known decision tree algorithms include ID3, C4.5 and CART algorithm.

A. ID3 (Iterative Dichotomizer 3) algorithm
It is an iterative two-class classifier that creates the decision tree in a top-down way. When creating a tree, it follows the rule that if there are several explanations for any phenomenon, it is advisable to prefer the least demanding. We can simply describe the ID3 algorithm through the following three steps:

- Takes all unused features and calculates their entropy with respect to the test set of data.
- Selects the attribute with the smallest entropy (highest information gain).
- Creates nodes with the selected attribute.

B. C4.5 algorithm
It is an improved ID3 algorithm. It recognizes nominal (discrete) and real (continuous) features. C4.5 creates a tree from the top down using the "divide and dominate" principle. If some values are missing in the tree, they ignore them and continue "in learning". It uses a relative information gain that already takes into account the number of variables of the given attribute. At present, there is a newer version of this algorithm – the C5.0 (See5) algorithm.

C. CART (Classification and Regression Trees) algorithm
The decision trees created by the CART algorithm have a binary topology, meaning that each node has just two offspring. CART algorithm can process any input data - it can work with both, qualitative and quantitative data. In principle, this algorithm works on the most appropriate allocation of input data according to one of the features.

5. EXPERIMENTAL RESULTS

In order to provide an accurate classification, we selected the decision tree method as the most appropriate for our study. We used the scikit-learn [18] – a free software machine learning library for the Python programming language, which offers all the tools required for deep data analysis. After comparing the above algorithms, we chose the CART algorithm because the library was used particularly to classify the parameters, and uses the just-optimized version of the CART algorithm. We also conducted statistical testing using the Mann-Whitney U test. It should confirm the assumption that there are characteristic differences between the handwriting of the control subjects and the test subjects.

5.1. Mann-Whitney U test

It is the non-parametric statistical test [19] that is used to compare two samples and to test whether the two sample means are equal or not.
In our case, we compare the calculated parameters of the control subjects with the parameters of the tested subjects – people with dysgraphia. The calculated parameters represent a set of data. The variable represents whether it is a control subject or a person with dysgraphia.

Hypothesis definition:
H0: The handwriting of control subjects is not statistically different from the handwriting of persons with dysgraphia.

Table 1 Sample parameters

sample	1	2	3	4	5	8	9	10	11	12
gender	M	M	M	M	M	F	F	M	F	F
L/R-handed	R	R	R	R	R	R	R	R	R	R
total time [s]	369,2	461,2	462,3	427,6	446,0	98,0	183,5	135,9	119,1	93,0
writing time [s]	214,3	228,6	258,0	216,0	212,9	72,7	99,9	47,3	69,1	48,3
in air time [s]	155,0	232,7	204,3	211,7	233,1	25,3	83,6	88,6	50,1	44,7
average pressure	425,6	326,8	737,7	254,6	428,9	372,4	441,5	516,4	518,1	345,3
average speed [cm/s]	3,7	2,6	2,9	2,3	3,7	5,1	4,8	5,9	7,1	5,5
average acceleration [cm/s^2]	751,7	512,3	586,5	464,8	734,4	1033,2	969,2	1183,1	1426,3	1102,6

H1: The handwriting of control subjects is statistically different from the handwriting of persons with dysgraphia.

We have chosen the significance level (p-value) to 0.05. That is, we are working with a 5 percent risk that we mark the right hypothesis as the wrong one. The results of the Mann-Whitney U test of the significance level for the individual parameter are shown in table 2, table 3 and table 4. The green color in the tables highlights the significance levels that confirm the assumption – hypothesis H1.

Table 2 Results of significance levels for time parameters

$p < 0{,}05$	Mann-Whitney U test
time over the tablet	0,00008
writing time	0,00006
total time	0,00001

The table 2 shows that for all time parameters there is a significant difference between the tested subjects and the control subjects.

Table 3 Results of significance levels for pen movement length and pressure parameter

$p < 0{,}05$	Mann-Whitney U test			
	p(d)	p(d_x)	p(d_y)	p(pressure)
minimum	0,726	0,719	0,012	0,134
maximum	0,041	0,027	0,001	0,569
average	0,007	0,021	0,001	0,107
median	0,003	0,011	0,001	0,075

The table 3 shows that the pen pressure parameter does not show the difference between the tested subjects and the control subjects. Conversely, for the pen movement length parameter, the hypothesis H1 was confirmed in all the statistical values except the horizontal and total minimum of this parameter.

Table 4 shows the results for speed (v), acceleration (a) and for jerk (j) coefficient. Significant changes between the tested and control subjects can be observed especially for the minimum and maximum statistical values.

5.2. Classification using the decision tree method

Before the classification itself, we created a dataset consisting of all respondents (80) and their calculated parameters (55). The success of the decision tree algorithm is strongly dependent on how we divide the dataset into a test and training subset. This is done using the *train_test_split* feature, which is part of the scikit learn library. We defined that the size of the test and training subset was the same. The result is four subsets of data: *x_train, x_test, y_train* and *y_test*. X-subsets contain values (individual parameters) and Y-subsets contain labels (healthy or dysgraphic).
We have computed the test subsets for multiple possible sizes. The resulting success rate, to be as real as possible, is calculated as the average of all the successes that we have obtained through different dataset distributions. The success of the decision trees algorithm in the classification of a set of all parameters was 83.4%. It follows that with the accuracy of 83.4% we can determine whether it is a healthy (control) subject or a person with dysgraphia.
While creating the classification using decision trees, we found that the dominant parameter is the minimum speed. Based on this single parameter, the classification score would be 93.9%. Conversely, if we removed this parameter from the dataset, the success rate of the algorithm would drop to 67.9%. For better clarity, the results are shown in table 5.

Table 4 Results of significance levels for kinematics parameters

$p < 0{,}05$	Mann-Whitney U test								
	p(v)	p(v_x)	p(v_y)	p(a)	p(a_x)	p(a_y)	p(j)	p(j_x)	p(j_y)
minimum	0,276	0,05	0,689	0,003	0,004	0,091	0,001	0,001	0,004
maximum	0,007	0,091	0,02	0,0006	0,073	0,021	0,006	0,063	0,021
average	0,704	0,865	0,697	0,749	0,697	0,704	0,741	0,682	0,704
median	0,849	0,704	0,992	0,865	0,741	0,873	0,889	0,795	0,873

Table 5 Results of classification accuracy using decision trees method

parameter input	accuracy
only a minimum speed parameter	93,9%
all parameters	83,4%
all parameters except minimum speed	67,9%

We also calculated statistical measures of the performance of a binary classification test – sensitivity and specificity.
Sensitivity is intuitively the ability of the classifier to find all the positive samples. It refers to the test's ability to correctly detect ill patients who do have the condition. Mathematically, this can also be written as:

$$sensitivity = \frac{tp}{tp + fn} \tag{2}$$

where tp is the number of true positives and fn the number of false negatives. The results are shown in table 6.

Table 6 Results of classification sensitivity using decision trees method

parameter input	sensitivity
only a minimum speed parameter	89%
all parameters	84%
all parameters except minimum speed	76%

Specificity is intuitively the ability of the classifier not to label as positive a sample that is negative. It relates to the test's ability to correctly reject healthy patients without a condition.. Mathematically, this can be expressed as:

$$specificity = \frac{tn}{tn + fp} \tag{3}$$

where tn is number of true negatives and fp the number of false positives. The results are shown in table 7.

Table 7 Results of classification specificity using decision trees method

parameter input	specificity
only a minimum speed parameter	90%
all parameters	85%
all parameters except minimum speed	76%

We also plotted a decision tree built from the handwriting parameters. Figure 1 depicts the decision tree generated from the entire dataset (55 parameters). The tree visualization shows that the minimum speed parameter is dominant. However, the decision tree does not contain any other parameter except the minimum speed, which is not very correct. That is why we have chosen to remove this parameter from the dataset and to build another tree.

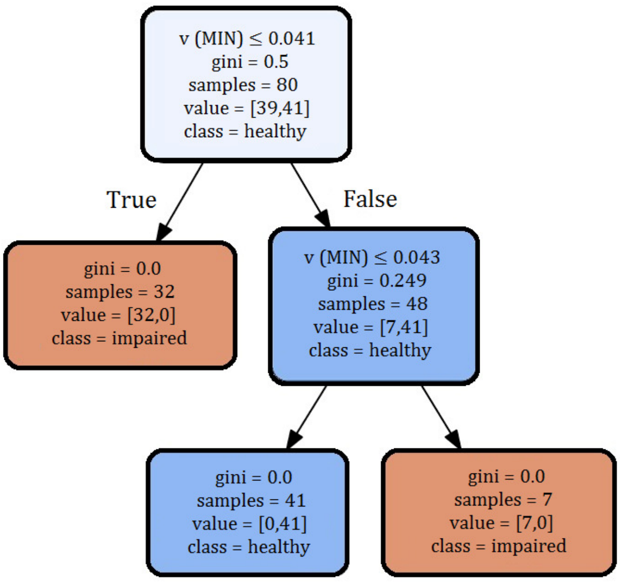

Fig. 1 Visualization of the decision tree plotted from all parameters

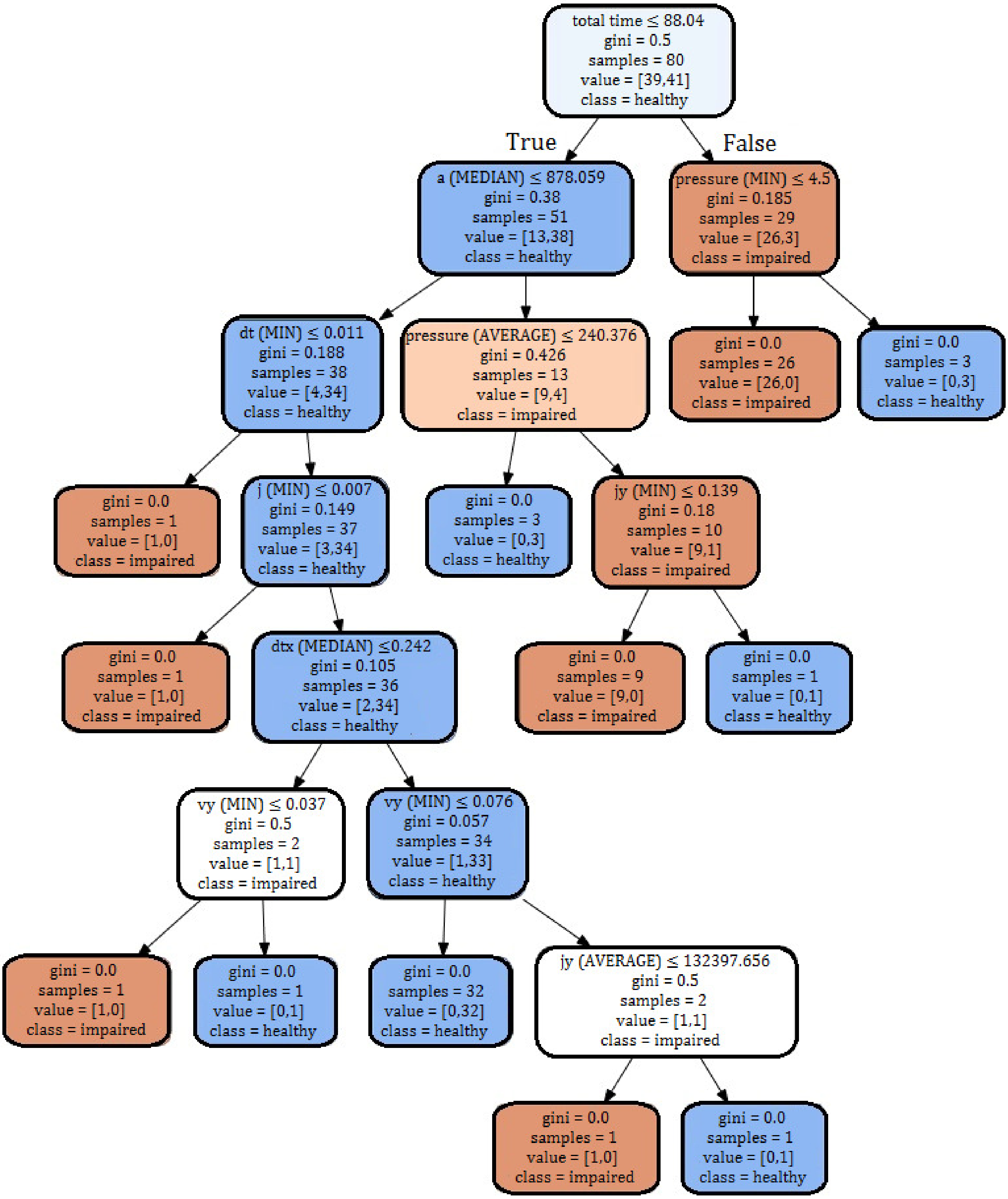

Fig. 2 Visualization of the decision tree plotted without a minimum speed parameter

In the same way, we plotted a decision tree containing all parameters except the minimum speed (a dataset consisting of 54 parameters). This tree is shown in figure 2. In this case, the classifier points to the total writing time as the most dominant parameter.

6. CONCLUSION

In the current study, we extracted basic handwriting features from samples of handwriting obtained by a digital tablet. The handwriting features provided by control subjects were compared with those of dysgraphic subjects in order to confirm the difference in handwriting. Individual parameters of the evaluation of handwriting have been calculated and also visualized using the Python programming language.
Above the extracted handwriting parameters, we determined classification using the decision tree classifier. This helped us find the dominant parameter, in our case it was the minimum writing speed. By computing the classifier's success, we found that the data is classified into classes (healthy/damaged) with an accuracy of 83.4% If we determined the classification according to the dominant parameter only, we would achieve a success rate of 93.9%. Conversely, if we did not have the minimum speed parameter available, the classification success would fall to 67.9%.
We also calculated classification function for sensitivity and specificity, which can correctly detect ill, respectively reject healthy patients. The success of the sensitivity calculation in the classification for a set of all parameters was 84%. But we obtained the best sensitivity results when calculating only the dominant parameter - 89%. When calculating without a minimum speed parameter, the success rate was 76%. We did the same calculations for the specificity. The success of the specifiticy for all parameters was 85%, calculation with only a minimum speed parameter was 90% and calculation of all parameters except minimum speed was 76%.
The obtained parameters were also subjected to statistical testing by the Mann-Whitney U test. The test confirmed the hypothesis that the handwriting of control subjects is statistically different from the handwriting of people with dysgraphia. From the calculated significance levels, we found that the most significant difference between the control subjects and the test subjects was in time parameters, as well as in the case of materiality levels.
After examining the individual parameters, we can confirm the difference in the handwriting of people with dysgraphia and people without handwriting impaires. The results show that people with dysgraphia write more slowly and need much more time to write the same task than the healthy individuals. Likewise, people with dysgraphia perform excessively long strokes during writing.
Further research is intended to focus on the calculation of a larger number of samples and to create a classification with the various other classifiers (e.g. using the support vector machine [20], random forest method or neural networks [21]), to optimize the results. It is also is intended to compare the results from classifiers with the results of statistical testing, e.g. using Bonferroni correction or Mann-Whitney U test.

ACKNOWLEDGEMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-16-0211 and by the Technical University of Košice under Grant No. FEI-2017-43 "Handwriting analysis focused on dysgraphy".

References

[1] JI, S. Y. – NAJARIAN, K. – SMITH, R. – HUYNH, T.: Computer Aided Traumatic Pelvic Injury Decision Making. Proceedings of the Southern Association for Information Systems Conference. Richmond, VA, USA, 2008

[2] RAZA, M. A. – CHAUDRY, Q. – ZAIDI S. M. T. – KHAN, M. B.: Clinical decision support system for Parkinsons disease and related movement disorders. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1108-1112, New Orleans, LA, 2017

[3] GANGADHAR, G. – JOSEPH, D. – SRINIVASAN, A. – SUBRAMANIAN, D. – SHIVAKESHAVAN, R. – SHOBANA, N. – CHAKRAVARTHY, V.: A Computational Model of Parkinsonian Handwriting that Highlights the role of the indirect pathway in the basal ganglia. Human Movement Science. vol. 28, no. 5, pp. 602-618, ISSN 0167-9457, 2009

[4] CARVALHO, C. M. – CHRISTINA, D. – SAADE, M. – CONCI, A. – SEIXAS, F. L. – LAKS, J.: A clinical decision support system for aiding diagnosis of Alzheimers disease and related disorders in mobile devices. 2017 IEEE International Conference on Communications (ICC). pp. 1-6, Paris, 2017

[5] PLEROU, A. – BOBORI, C. – VLAMOS, P.: Molecular basis of Huntingtons disease and brain imaging evidence. 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). pp. 387-391, Abu Dhabi, 2015

[6] ROSENBLUM, S. – DROR, G.: Identifying Developmental Dysgraphia Characteristics Utilizing Handwriting Classification Methods. 2017 IEEE Transactions on Human-Machine Systems. vol. 47, no. 2, pp. 293-298, April 2017

[7] KUSHKI, A. – SCHWELLNUS, H. – ILYAS, F. – CHAU, T.: Changes in kinetics and kinematics of handwriting during a prolonged writing task in children with and without dysgraphia. Research in Developmental Disabilities. vol. 32, no. 3, pp. 1058-1064, 2011

[8] ROSENBLUM, S. – WEISS, P. – PARUSH, S.: Temporal measures of poor and proficient handwriters. Proceedings of the 10th Biennal Conference of the Interna-

tional Graphonomics Society. pp. 119-125, Nijmegen, The Netherlands, 2001

[9] ROSENBLUM, S. – WEISS, P. – PARUSH, S.: Product and process evaluation of handwriting difficulties. Educational Psychology Review. vol. 15, pp. 41-81, 2003a

[10] ROSENBLUM, S. – PARUSH, S. – WEISS, P.: Computerized temporal handwriting characteristics of proficient and poor handwriters. American Journal of Occupational Therapy. vol. 57, pp. 129-138, 2003b

[11] MEKYSKA, J. – FAUNDEZ-ZANUY, M. – MZOUREK, Z. – GALAZ, Z. – SMEKAL, Z. – ROSENBLUM, S.: Identification and Rating of Developmental Dysgraphia by Handwriting Analysis. IEEE Transactions on Human-Machine Systems. vol. 47, no. 2, pp. 235-248, April 2017

[12] TUCHA, O. – MECKLINGER, L. – THOME, J. – REITER, A. – ALDERS, G. L – SARTOR, H. – NAUMANN, M. – LANGE, K. W: Kinematic Analysis of Dopaminergic Effects on Skilled Handwriting Movements in Parkinsons Disease. Journal of Neural Transmission. vol. 113, pp. 609-623, ISSN 0300-9564, 2006

[13] ROSENBLUM, S.: Development, reliability, and validity of the handwriting proficiency screening questionnaire (HPSQ). American Journal of Occupational Therapy. vol. 62, no. 3, pp. 298-307, 2008

[14] PAZ-VILLAGRAN, V. – DANNA, J. – VELAY, J. L.: Lifts and stops in proficient and dysgraphic handwriting. Human Movement Science. vol. 33, pp. 381-394, ISSN 0167-9457, 2014

[15] DROTAR, P. – MEKYSKA, J. – REKTOROVA, I. – MASAROVA, L. – SMEKAL, Z. – FAUNDEZ-ZANUY, M.: Decision support framework for parkinsons disease based on novel handwriting markers. IEEE Transactions on Neural Systems and Rehabilitation Engineering. vol. 23, no. 3, pp. 508-516, 2015

[16] HAMSTRA-BLETZ, L. – BLOTE, A.: A longitudinal study on dysgraphic handwriting in primary school. Journal of Learning Disabilities. vol. 26, no. 10, pp. 689-699, 2003

[17] BREIMAN, L.: Random forests. Machine Learning. vol. 45, pp. 5-32, 2001

[18] PEDREGOSA, F. et al.: Scikit-learn: Machine Learning in Python. JMLR 12. pp. 2825-2830, 2011 `http://scikit-learn.org/stable/`

[19] UPTON, G. – COOK, I.: A Dictionary of Statistics. Oxford University Press. rev. 2nd ed., pp. 453, ISBN 9780199541454, 2008

[20] VAVREK, J. – FECIĽAK, P. – JUHÁR, J. – ČIŽMÁR, A.: Classification of broadcast news audio data employing binary decision architecture. Computing and Informatics. vol. 36, no. 4, pp. 857-886, 2017

[21] ÁDÁM, N. – PIETRIKOVÁ, E. – CHOVANCOVÁ, E. – FECIĽAK, P.: The Impact of Data Representationson Hardware Based MLP Network Implementation. Acta Polytechnica Hungarica. vol. 15, no.2, 2018

Received October 26, 2018, accepted January 11, 2019

BIOGRAPHIES

Zuzana Dankovičová graduated (Ing.) in 2016 at the Department of Computers and Informatics of the Faculty of Electrical Engineering and Informatics at Technical University in Košice, Slovakia. She is currently student of postgraduate study in an informatics major at the same department of the Technical University in Košice. Her scientific research is focusing on biomedical informatics, data processing, decision support systems and handwriting analysis.

Matúš Uchnár graduated (Ing.) from a Technical University in Košice with a major in business informatics in 2015 and is currently in the fourth year of postgraduate study in an informatics major. His thesis is aimed towards malware analysis and malware analysis tools.