

INFOSCIENCE TECHNOLOGY: THE IMPACT OF INTERNET ACCESSIBLE MELANOID DATA ON HEALTH ISSUES

JW Grzymala-Busse¹, ZS Hippe², M Knap² and W Paja²

¹ Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence KS 66045, USA
Email: Jerzy@ku.edu

² Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management, Suchbarskiego 2, 35-225 Rzeszow, Poland
Email: {ZHippe, MKnap, WPaja}@wsiz.rzeszow.pl

ABSTRACT

In this paper the development of a new internet information system for analyzing and classifying melanocytic data, is briefly described. This system also has some teaching functions, improves the analyses of datasets based on calculating the values of the **TDS** (**Total Dermatoscopy Score**) (Braun-Falco, Stolz, Bilek, Merkle, & Landthaler, 1990; Hippe, Bajcar, Blajdo, Grzymala-Busse, Grzymala-Busse, & Knap, et al., 2003) parameter. Calculations are based on two methods: the classical ABCD formula (Braun-Falco et al., 1990) and the optimized ABCD formula (Alvarez, Bajcar, Brown, Grzymala-Busse, & Hippe, 2003). A third method of classification is devoted to quasi-optimal decision trees (Quinlan, 1993). The developed internet-based tool enables users to make an early, non-invasive diagnosis of melanocytic lesions. This is possible using a built-in set of instructions that animates the diagnosis of the four basic lesions types: benign nevus, blue nevus, suspicious nevus and melanoma malignant. This system is available on the Internet website: <http://www.wsiz.rzeszow.pl/ksesi>.

Keywords: teledermatology, TDS, internet-based tool, melanocytic lesions, information system

1 INTRODUCTION. TELEDERMATOLOGY

Increasing subspecialization in medicine results in a concentration of experts at few medical centers and a lack of expert knowledge in other, small, predominantly rural areas. Physicians in these areas may have problems with the diagnosis and management of equivocal pigmented skin lesions. The patient therefore has to be sent to an expert in a medical center for further diagnosis and, depending on the results, for adequate treatment. This transportation means inconvenience for the patient and costs for the health care system because of re-examination by experts.

In the last few years continuous progress in information technology has led to the introduction of a revolutionary diagnostic tool, known as *telemedicine*, that is improving communication between physicians and medical specialists and which will help reduce costs for the citizen and the health care system. In several medical specialties, where digital images are crucial in diagnosis and management decisions, such as radiology or internal medicine (endoscopy; ultrasound) to name but a few, telemedicine already represents a well-integrated part of daily medical life. In addition, it has recently been shown (Piccolo, Smolle, Wolf, Peris, Hofmann-Wellenhof, & Dell'Eva, 1999) that there is a subdomain of telemedicine known as *teledermatology* (or *teledermoscopy*). It enables the sending of dermoscopic images of pigmented skin lesions over telematic networks (Provost, Kopf, Rabinovitz, Stolz, De David, Wasti et al., 1998) and a fast comparison of diagnosis of the same skin lesions by different experts.

Another encouraging approach due to the fast and easy exchange of information via the internet is the possibility of discussing dermoscopic images with experts all over the world. An example would be The New York University Group's "e-Room" project called *Dermnetwork* (<http://www.dermnetwork.org>) where information on interesting dermoscopy cases can be shared.

Our main goal was to use selected machine learning methods to create a hierarchy of importance of melanocytic symptoms. These symptoms are part of a well-known parameter called **TDS** (**Total Dermatoscopy Score**) that is a useful diagnostic tool for melanoma.

The TDS, according to Braun-Falco et al. (1990), is computed using the following formula (known as the *ABCD formula*):

$$\mathbf{TDS} = 1.3 * \mathbf{Asymmetry} + 0.1 * \mathbf{Border} + 0.5 * \sum \mathbf{Colors} + 0.5 * \sum \mathbf{Diversity} \quad (1)$$

where **A** is a description of lesion's asymmetry, **B** is a description of lesion's border, **C** is a description of colors appearing in considered lesion, and **D** is a specification of lesion's diversity.

The variable *<Asymmetry>* has three different values: *symmetric spot* (counted as 0), *one-axial asymmetry* (counted as 1), *two-axial asymmetry* (counted as 2). *<Border>* is a numerical attribute, with values from 0 to 8. A lesion is partitioned into eight segments. The border of each segment is evaluated: the sharp border contributes 1 to *<Border>*, the gradual border contributes 0. *<Color>* has six possible values: *black, blue, dark brown, light brown, red* and *white*. Similarly, *<Diversity>* has five values: *pigment dots, pigment globules, pigment network, structure-less areas* and *branched streaks*. In our data set *<Color>* and *<Diversity>* were replaced by binary single-value variables: *present* or *absent*, for example, the *pigment dots* structure is *absent*, the *black* color is *present*, etc. In this way, our data-set contains objects described by 13 descriptive attributes.

2 MELANOCYTIC DATASET

The source data describing the melanomas were collected in the Outpatient Center of Dermatology in Rzeszow, Poland. Every case was collected in the form of a hand-written description (history of disease) and was performed by physicians. From the anonymous record of 548 patients, a data set (in the form of text file) was created in the Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management, in Rzeszow (Hippe *et al.*, 2003). Each case is described by 13 attributes, divided into 5 groups of attributes (presented in previous paragraph), and another descriptive attribute known as TDS parameter, computed based on another 13 attributes. The formula in Eq. (1) is used to calculate TDS attribute. In this way, verification of our research was possible with use of the concept of constructive induction (Michalski, Bratko, & Kubat, 1998). Each investigated case was assigned to one of four decision categories: *benign nevus, blue nevus, suspicious nevus* or *melanoma malignant*. The number of different diagnoses in the database are presented in Table 1. The diagnoses have been confirmed histologically.

| Diagnose | Number of cases | % |
|--------------------|-----------------|----|
| Benign nevus | 248 | 45 |
| Blue nevus | 78 | 14 |
| Suspicious nevus | 108 | 20 |
| Melanoma malignant | 114 | 21 |

Table 1. The number of different diagnoses in the database

3 STRUCTURE AND OPERATION OF THE SYSTEM

The described system has an internet-based user interface (see Figure 3) that is used to gain access to two main working modules Figure 1. The first module is a multimedia module which fulfills the role of a virtual consultation on the features of melanocytic lesions. It is a set of instructions about early diagnosis and the determination of descriptive attributes.

Using the first module, a user can learn about correct methods of diagnosing selected attributes, based on authentic patient cases. The second module is a kind of melanocytic calculator for the non-invasive diagnosis of melanocytic lesions. The input values for this module are values for a set of 13 describing attributes. These values, introduced by users, are processed by two different formulas to achieve the 14-th attribute, TDS. The first formula is a standard ABCD formula (Eq. (1)), however the second formula is an optimized formula (Alvarez *et al.*, 2003; Bajcar, Grzymała-Busse, Grzymała-Busse, & Hippe, 2002) that makes possible to calculate New_TDS, which is described in detail in (Alvarez *et al.*, 2003), which states:

$$\begin{aligned}
 \text{New_TDS} = & (0.8 * \text{Asymmetry}) + (0.11 * \text{Border}) + (0.5 * \text{C_White}) + (0.8 * \text{C_Blue}) \\
 & + (0.5 * \text{C_DarkBrown}) + (0.6 * \text{C_LightBrown}) + (0.5 * \text{C_Black}) + (0.5 * \text{C_Red}) \\
 & + (0.5 * \text{Pigment_Networks}) + (0.5 * \text{Pigment_Dots}) + (0.6 * \text{Pigment_Globules}) \\
 & + (0.6 * \text{Branched_Streaks}) + (0.6 * \text{Structureless_Areas})
 \end{aligned}
 \tag{2}$$

The values of TDS attribute gained are the basis of making a diagnosis using classic ABCD formula. In this way, according to second Stolz's algorithm (Stolz, 1997) (with flexible thresholds of classification) and to other sources (<http://www.dermatology.com/dermoscopy/dermoscope.htm>), the result is defined as:

TDS < 4.76 indicates *benign nevus*,
4.76 ≤ TDS ≤ 5.45 indicates *suspicious nevus*,
 and **TDS > 5.45** indicates *malignant nevus*

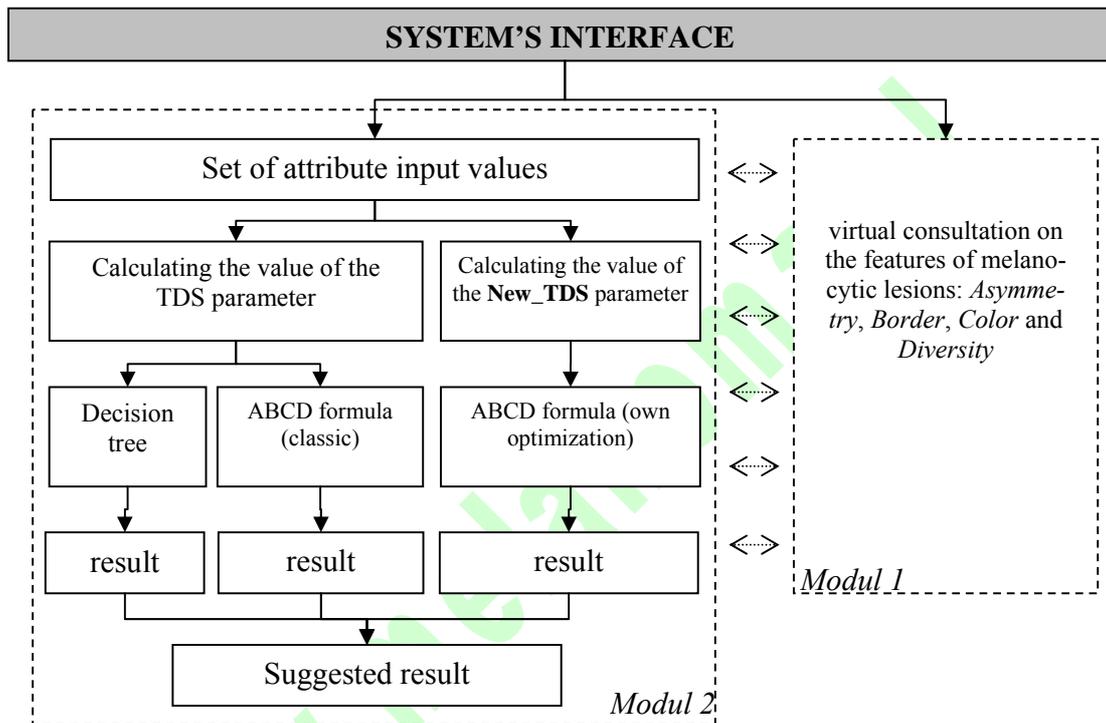


Figure 1. Structure of the system

Similar thresholds are used in case of the ABCD formula with its own optimization of the TDS parameter (**New_TDS**). The learning model developed using the standard TDS parameter, classifies unseen objects with an error rate equal to 9-11%. However, the learning model developed using the optimized TDS parameter, classifies the same set of unseen objects with an error rate of about 5%.

The third way to diagnose lesions is by using a decision tree (see Figure 2). This tree was developed using the data set presented in paragraph 2. In the process of developing the decision tree the ID3/C4.5 algorithm was used (Quinlan, 1993). The developed tree is shown below:

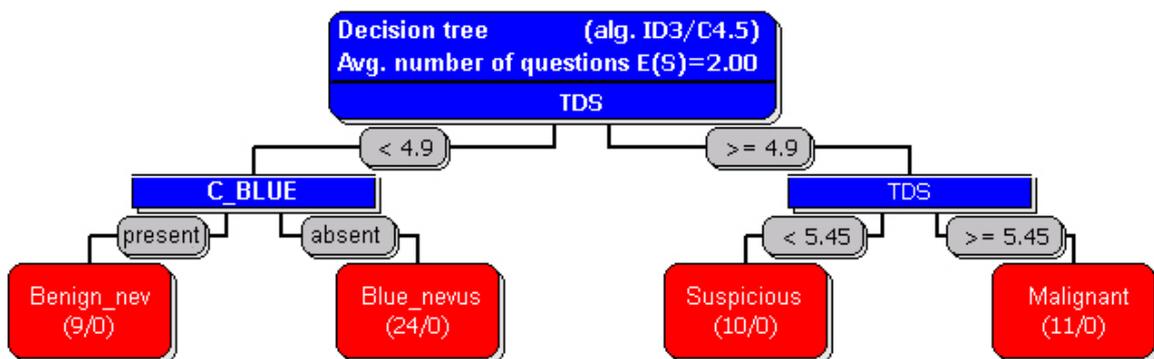


Figure 2. Learning model in the form of decision tree

In our recent research, we stated that the developed decision tree classified new, unseen melanoma cases with an error rate equal to exactly 1.4%. Finally, the system makes three independent diagnoses based on the *ABCD formula (classic)*, the *ABCD formula (own optimization)* and the *decision tree*. On the basis of these three decisions, the system suggests the final diagnosis by voting. In cases where all three diagnoses are different, the system leaves the decision to the user.

Figure 3. User interface

4 CONCLUSIONS

The correct classification of pigment skin lesions is possible using histopathological lesion research. The newest trend of diagnosing using non-invasive methods, has come about because of the dissemination of information technology tools that support this process.

In this paper, the practical development of a new internet-based information system for classifying melanocytic lesions, is briefly described. This system also has some teaching functions, and improves the analysis of datasets based on calculating values of the **TDS (Total Dermatoscopy Score)** parameter. The system uses three different methods to find a correct diagnosis of a lesion. The developed internet-based tool enables users to make an early, non-invasive diagnosis of melanocytic lesions. This system is available on the Internet website: <http://www.wsiz.rzeszow.pl/ksesi>.

5 REFERENCES

Alvarez, A., Bajcar, S., Brown, F.M., Grzymała-Busse, J.W., & Hippe, Z.S. (2003) Optimization of the ABCD Formula Used for Melanoma Diagnosis. In Kłopotek, M.A., Wierzchoń, S.T., & Trojanowski, K., (Eds.), *Advances in Soft Computing, (Intelligent Information Processing and Web Mining)*, Heidelberg: Springer-Verlag.

Bajcar, S., Grzymała-Busse, J.W., Grzymała-Busse, W.J., & Hippe, Z.S., (2002) Diagnosis of Melanoma Based on Data Mining and ABCD Formulas, *Proc. 3rd International Conference on Hybrid Intelligent Systems*, Melbourne, Australia.

Braun-Falco, O., Stolz, W., Bilek, P., Merkle, T., & Landthaler, M. (1990) *Dermatoscopy. Das Dermatoskop. Eine Vereinfachung der Auflichtmikroskopie von pigmentierten Haut-veränderungen*, *Hautarzt* (40)131-135.

Hippe, Z.S., Bajcar, S., Błajdo, P., Grzymała-Busse, J.P., Grzymała-Busse, J.W., Knap, M., Paja, W., & Wrzesień, M. (2003) Diagnosing Skin Melanoma: Current versus Future Directions. *TASK Quarterly*, 7(2), 289-293.

Michalski, R.S., Bratko, I., & Kubat, M. (Eds.) (1998) *Machine Learning and Data Mining: Methods and Applications*, Chichester: J. Wiley & Sons Ltd

Piccolo, D., Smolle, J., Wolf, I.H., Peris, K., Hofmann-Wellenhof, R., Dell'Eva, G., Burrioni, M., Chimenti, S., Kerl, H., & Soyer, H.P. (1999) "Face-to-face" versus remote diagnosis of pigmented skin tumors: a teledermoscopic study. *Arch Dermatol* 135, pp 1467-1471.

Provost, N., Kopf, A.W., Rabinovitz, H.S., Stolz, W., De David, M., Wasti, Q., & Bart, R.S. (1998) Comparison of conventional photographs and telephonically transmitted compressed digitized images of melanomas and dysplastic nevi. *Dermatology* 196, pp 299-304.

Quinlan, J.R. (1993) *C4.5 Programs for Machine Learning*, San Mateo, CA, Morgan Kaufmann.

Stolz, W., (1997) Auflichtmikroskopische Diagnose des malignen Melanoma. In Garbe, C., Dummer, R., Kaufmann, R., & Tilgen, W., (Eds.), *Dermatologische Onkologie*, Heidelberg: Springer-Verlag.

www.melanomapi