[7] Friedman-Hill E 1999 *JESS – the Java Expert System Shell* http://herzberg.ca.sandia.gov/jess/
[8] Negrete Martinez J and Gonzalez Perez P 1998 *Expert Systems with Applications* **14** 102
[9] Dobrowolski G and Nawarecki E 2001 *Management and Control of Production and Logistics 2000*, PERGAMON (Binder Z, Ed.), pp. 445–450
[10] Jeon H, Petrie C and Cutkosky M R 2000 *IEEE Internet Computing* **4** 89

Diagnosing Skin Melanoma: Current versus Future Directions

Zdzisław S. Hippe[1], Stanisław Bajcar[2], Piotr Blajdo[1], Jan P. Grzymala-Busse[3], Jerzy W. Grzymala-Busse[3], Maksymilian Knap[1], Wiesław Paja[1] and Mariusz Wrzesień[1]

[1] *Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management, Sucharskiego 2, 35-225 Rzeszow, Poland, zhippe@wenus.wsiz.rzeszow.pl*

[2] *Regional Dermatology Center, Warzywna 3, 35-310 Rzeszow, Poland,*

[3] *Department of Electrical Engineering and Computer Science, University of Kansas Lawrence, KS 66045, USA*

(Received 18 January 2003)

Abstract: A new database containing 410 cases of *nevi pigmentosi*, in four categories: *benign nevus*, *blue nevus*, *suspicious nevus* and *melanoma malignant*, carefully verified by histopathology, is described. The database is entirely different from the base presented previously, and can be readily used for research based on the so-called constructive induction in machine learning. To achieve this, the database features a different set of thirteen descriptive attributes, with a fourteenth additional attribute computed by applying values of the remaining thirteen attributes. In addition, a new program environment for the validation of computer-assisted diagnosis of melanoma, is briefly discussed. Finally, results are presented on determining optimal coefficients for the well-known ABCD formula, useful for melanoma diagnosis.

Keywords: melanoma, TDS, machine learning in diagnosis of

1. Introduction

In recent papers [1, 2], we have presented the results of experiments on new samples relating to changes in skin melanoma, using machine learning with the idea of generating a model of learning to help identify and classify cases of skin melanoma. Skin melanoma may be a symptom of serious skin diseases, or even cancer, which has a high mortality rate. The numbers of victims of this type are rising because of the high levels of ultraviolet radiation entering the atmosphere and the increasingly thin ozone layer [3]. Anonymous data sets pertaining to cases of skin cancer have been collected by the Regional Dermatology Center in Rzeszow, Poland [4]. This data set of cases has been analyzed and expert systems based upon it have been implemented at the Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management in Rzeszow, Poland. The first version of the data set has been analyzed in a paper presented at the INFOBAZY'99 conference [1, 2]. The actual production version contains (i) new internal structures and (ii) an increased number of registered cases (from 250 to 410). Regarding (i), the data set information is stored in 13 attributes that are regularly used in dermatology for typical analysis of skin-based melanoma. In the context of these attributes, we calculate the **TDS** (Total Dermatoscopy Score) indicator [5]. The underlying idea of our experiments was to prepare our sample in both Polish and English and use specialized software algorithms in the verification of its accuracy, as well as generate learning models to help diagnose diseases. The data sets were acquired in studies taking place simultaneously in Rzeszow, Poland (University of Information Technology and Management) and the United States (University of Kansas, Lawrence, Kansas). In the following sections, the data sets have its statistical analysis and its machine learning results discussed. An earlier version of this paper was presented at the 3rd National Conference INFOBAZY'2002, Gdansk, Poland, June 24–26, 2002 [6].

2. Statistical analysis of the data sets

The attributes used in deducing diagnoses of melanoma have been broken down into 5 categories: ⟨Asymmetry⟩, ⟨Border⟩, ⟨Color⟩, ⟨Diversity⟩ and ⟨TDS⟩. The ⟨Asymmetry⟩ parameter can have the following values: symmetrical, single-axis asymmetry and dual-axis asymmetry. ⟨Border⟩ is a numerical attribute with discrete values between 0 and 8. The next two categories, ⟨Color⟩ and ⟨Diversity⟩, have symbolic values. ⟨Color⟩ can have six allowed values: black, blue, light brown, dark brown, red, and white. Likewise, ⟨Structure⟩ has five possible values: pigment dots, pigment globules, pigment network, structureless areas and branched

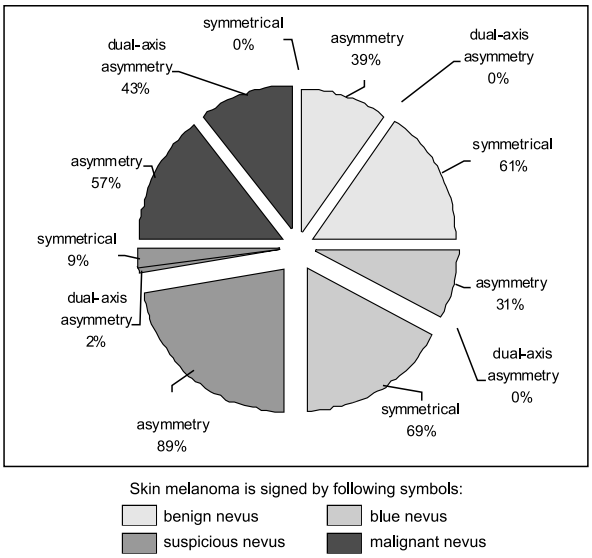

Figure 1. Appearance of the ⟨**Assymetry**⟩ attribute in each decision class

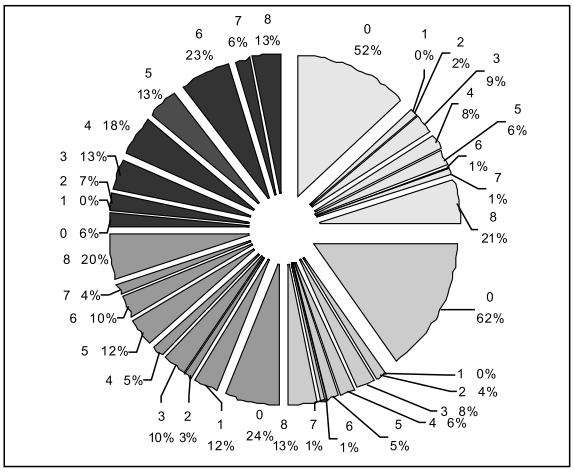

Figure 2. Appearance of the ⟨**Border**⟩ attribute in each decision class (symbols as per Figure 1)

streaks. In all these cases, the attributes pertaining to pigments and their diversity are Boolean and state the presence (1) or lack (0) of an attribute. Thus, every entry from the data set of anonymous patients is characterized by 13 attributes. For computing the fourteenth attribute, called **TDS**, the other 13 attributes are used, so that the **TDS** attribute is obtained by constructive induction [7]. The **TDS** indicator is computed using the following formula:

$$\mathbf{TDS} = 1.3 \cdot \langle \text{Asymmetry} \rangle + 0.1 \cdot \langle \text{Border} \rangle + \\ + 0.5 \cdot \langle \text{Color} \rangle + 0.5 \cdot \langle \text{Diversity} \rangle, \quad (1)$$

where the values for ⟨Asymmetry⟩ are as follows: symmetrical equals 0, single-axis asymmetry equals 1, and dual-axis asymmetry equals 2. ⟨Color⟩ represents the sum of represented pigment colors, whereas ⟨Diversity⟩ is the sum of the five represented diversity attributes. The accuracy of the calculated **TDS** plays a key role in generating a machine learning model using the concept induction system, and its correctness has been verified using an Excel spreadsheet calculation [8], which in turn allowed checking the individual work of specialist doctors. In this way,

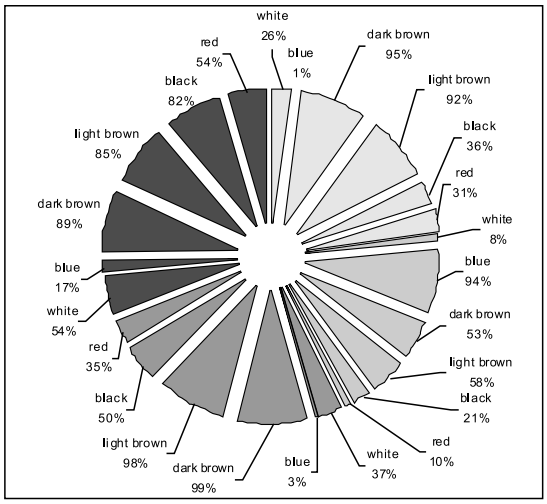

Figure 3. Appearance of the ⟨**Color**⟩ attribute in each decision class (symbols as per Figure 1)

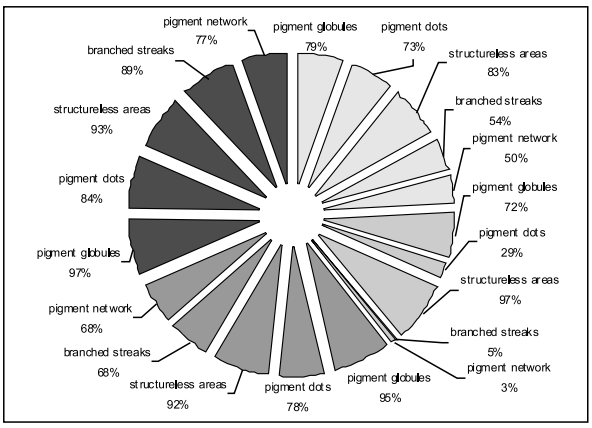

Figure 4. Appearance of the ⟨**Diversity**⟩ attribute in each decision class (symbols as per Figure 1)

we created a data set without any errors. Statistical analysis using the aforemtioned tools can be seen in Figures 1–4.

3. Using machine learning programs for modelling

The working data set was used for testing the machine learning model to help identify and diagnose changes in skin melanocytes. This field used the following program modules: ***Rule*SEEKER** (used to create rules), ***Tree*SEEKER** (generates quasi-optimal decision trees), ***Affinity*SEEKER** (seeks the similarities for diagnosing a patient with known results in the database), ***Plane*SEEKER** (searches for optimal decision planes based upon known attributes) and ***Score*SEEKER** (rates machine learning models generated by the former program modules). Due to restrictions on this paper's size, it is unfortunately not possible to elaborate on the details of all of the program modules. Instead, we shall concentrate on ***Score*SEEKER**, which is the most pertinent to the discussions of this conference, and which works (on raw data sets) on N-series data sets

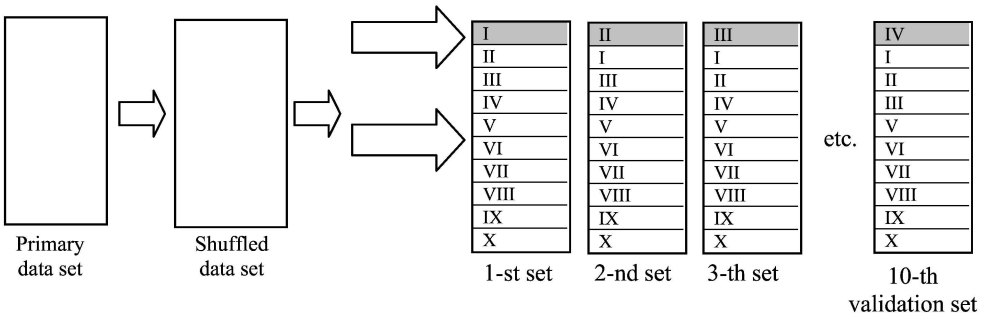

Figure 5. Illustration of the operation of the computer program system ***Score*SEEKER**. In the first step, records in the database are shuffled, and then 10 pairs of databases are created. In each of these, pairs 9/10 are used to generate a learning model and 1/10 – to test it

Table 1. Optimal coefficients for **TDS**

Attribute	Data set with 276 cases	Data set with 410 cases	
	Agglomerative discretization	Agglomerative discretization	Divisive discretization
Asymmetry	0.6	0.9	1.3
Border	0.1	0.14	0.11
Color black	0.5	0.5	0.5
Color blue	0.4	0.5	0.4
Color dark brown	0.5	0.3	0.4
Color light brown	0.5	0.4	0.5
Color red	0.4	0.5	0.5
Color white	0.4	0.5	0.4
Diversity pigment dots	0.5	0.5	0.4
Diversity pigment globules	0.6	0.5	0.5
Diversity pigment network	0.5	0.5	0.5
Diversity structureless areas	0.5	0.5	0.5
Diversity branched streaks	0.5	0.5	0.4

Table 2. Error rates in %

Data set with	Data set with 276 cases	Data set with 410 cases	
	Agglomerative discretization	Agglomerative discretization	Divisive discretization
original **TDS**	10.21	4.38	3.50
optimal **TDS**	6.04	4.51	3.63
no **TDS**	13.73	13.82	13.49

Table 3. Standard deviations in %

Data set with	Data set with 276 cases	Data set with 410 cases	
	Agglomerative discretization	Agglomerative discretization	Divisive discretization
original **TDS**	0.99	0.74	0.48
optimal **TDS**	0.84	0.78	0.49
no **TDS**	1.31	1.08	1.06

with a specific data structure (Figure 5) allowing moving current methods [9] of grading machine learning models.

4. Optimization of TDS

Both data sets, the old one with 250 training cases and additional 26 testing cases, and the new one with 410 cases, have been examined at the University of Kansas in Lawrence, Kansas. One of aims of the research conducted there was optimization of the ABCD formula to compute **TDS** or, more precisely, optimization of the 13 coefficients of Equation (1). The criterion of optimization was the minimization of the error rate in diagnosis of melanoma. The main problem is discrete optimization, which is known to be difficult and time consuming. Results of experiments on both data sets, described in [10] and [11], are presented in Tables 1, 2, and 3.

The error rate presented in Table 1 was computed using randomized ten-fold cross validation, in which, for every set of coefficients, experiments were repeated 30 times using different reshuffling of the original data set for each process of ten-fold cross validation. The optimal coefficients, presented in Table 3, were searched for using fixed (*i.e.* not randomized) ten-fold cross validation. However, the error rate of melanoma diagnosis for the final choice of optimal coefficients was verified with randomized ten-fold cross validation.

5. Conclusions

With 276 cases in the data set, there are significant differences between the original and the optimal choice of coefficients for **TDS**, at a 95% confidence level. Furthermore, with the same 95% confidence, diagnosis with **TDS** determined with any choice of coefficients (original or optimal), with 276 or 410 cases, is better than diagnosis in which **TDS** has not been used. The difference between error rates for diagnoses of melanoma using different discretization methods is not significant (with 95% confidence).

With 276 cases, the error rates for diagnosis using **TDS** are significantly higher than the rates for using **TDS** with 410 cases. This is due to better rule sets induced from the more representative data set of 410 cases. Also, with 276 cases, the error rate using optimal coefficients for **TDS** is significantly lower than the error rate using original coefficients for computing **TDS**. On the other hand, with 410 cases, the difference between the error rate for original and optimal coefficients for computing **TDS** are insignificant, always with 95% confidence.

In future experiments, with further increase in the number of registered cases, a hierarchy of importance of values may be created, based upon the described attributes. This would predictably allow more objective diagnosis and classification of cases.

Acknowledgements

This research has been partially supported by the State Committee for Research (KBN) of the Republic of Poland under grant no. 7 T11E 030 21.

References

[1] Hippe Z S 1999 *Proc. "INFOBAZY'99 – Databases for Science"*, TASK Computer Center in Gdansk, Poland, pp. 120–124

[2] Hippe Z S 1999 *TASK Quart.* **3** (4) 483

[3] Riegel D S, Friedman R J, Kopf A, Weltman R, Prideau P G, Safai B, Lebwohl M G, Elizeri Y, Torre D P and Bonford T T 1986 *J. Am. Acad. Dermatol.* **14** 857

[4] Bajcar S and Hippe Z S 1999 *Proc. Telemedicine Conference*, Lodz, Poland, pp. 175–178

[5] Braun-Falco O, Stolz W, Bilek P, Merkle T and Landthaler M 1990 *Hautartzt* **40** 131

[6] Hippe Z S, Grzymala-Busse J W, Bajcar S, Blajdo P, Knap M, Paja W and Wrzesień M 2002 *Proc. "INFOBAZY'2002 – Databases for Science"*, TASK Computer Center in Gdansk, Poland, pp. 51–55 (in Polish)

[7] Michalski R S, Bratko I and Kubat M 1998 *Machine Learning and Data Mining. Methods and Applications*, John Wiley and Sons LTD., Chichester

[8] Hippe Z S and *et al.* 2001 *Research on Methods of Computer-aided Diagnosis of Melanocytic Lesions and Melanoma*, Research Report, Grant KBN no. 7 T11E 030 21, Poland (in Polish)

[9] Weiss S and Kulikowski C A 1991 *Computer Systems that Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*, Morgan Kaufmann Publishers Inc., San Francisco, CA, pp. 17–49

[10] Grzymala-Busse J P, Grzymala-Busse J W and Hippe Z S 2001 *Proc. 25th Anniversary Annual International Computer Software and Applications Conference COMPSAC 2001*, Chicago, IL, pp. 615–620

[11] Grzymala-Busse J W, Hippe Z 2002 *Rough Sets and Current Trends in Computing* (Alpigini J J, Peters J F, Skowron A and Zhong N, Eds.), Springer-Verlag, Heidelberg, pp. 538–545