Medical Informatics Europe 2002
G Surján et al. (eds)
IOS Press 2002

Nearest Neighbour Classification of Otoneurological Data

Kati VIIKKI[a], Matti TAPANI[a], Martti JUHOLA[a], Ilmari PYYKKÖ[b]

[a]*Department of Computer and Information Sciences, 33014 University of Tampere, Finland*
[b]*Department of Otorhinolaryngology, Karolinska Hospital, 17176 Stockholm, Sweden*

Abstract. We studied the nearest neighbour classification of patient cases with benign positional vertigo, Menière's disease, sudden deafness, traumatic vertigo, vestibular neuritis, and vestibular schwannoma. The classification results were compared to the inference results obtained by an otoneurological expert system ONE whose inference mechanism somewhat resembles the classical nearest neighbour method. With respect to the predictive accuracy, the classification results of these two systems agreed. The best predictive accuracy for the expert system ONE was 79.7% and for the nearest neighbour method 80.5%. However, differences in the true positive rates for sudden deafness, traumatic vertigo, vestibular neuritis, and vestibular schwannoma were found. The nearest neighbour classification results will be used in the refinement of ONE's knowledge base.

1. Introduction

Diagnosis of vertiginous patients requires thorough knowledge of otoneurology; it provides challenging tasks even for specialists of the field [3]. An otoneurological expert system ONE [1] has been developed to assist the diagnostic procedure for peripheral and central diseases involving vertigo [3], and to provide a tutorial guide for medical students.

The inference mechanism of ONE [1] somewhat resembles the nearest neighbour method of pattern recognition [7]. The knowledge base of ONE contains a description or a pattern for each disease, and the inference mechanism searches the pattern that matches best to the case to be classified.

In this study, we use the nearest neighbour method to classify patient cases representing the six largest diagnostic groups of ONE's database [4]: benign positional vertigo, Menière's disease, sudden deafness, traumatic vertigo, vestibular neuritis, and vestibular schwannoma. We compare the results of the nearest neighbour classification to the inference results obtained by ONE [9].

2. Material

The data of the present study were collected in the Department of Otorhinolaryngology at Helsinki University Central Hospital. It contained 815 patient cases representing the six largest diagnostic groups in ONE's database: benign positional vertigo (n_{bp} = 146; 17.9%), Menière's disease (n_{md} = 313; 38.4%), sudden deafness (n_{sd} = 41; 5.0%), traumatic vertigo (n_{tv} = 65; 8.0%), vestibular neuritis (n_{vn} = 120; 14.7%), and vestibular schwannoma (n_{vs} = 130; 16.0%). From each diagnostic group, approximately 70% of the cases were randomly selected to the training set (n_{tr}=569). The rest 30% of the cases formed the testing set (n_{tc}=246).

3. Methods

3.1. Inference Mechanism of ONE

The knowledge base of ONE contains a description for each disease in the form of weight and fitness values. The present version covers 18 diseases and disorders that are described with the use of attributes concerning patient history, symptoms, clinical tests, and examinations [4]. To each attribute, a weight value expressing the significance of the attribute for the disease is set. Fitness values assigned to attribute values express the correspondence between the attribute values and the disease.

The inference mechanism transforms attribute values to scores based on the fitness values and weights. Let *d* be a disease and *n*(*d*) be the number of attributes associated with *d* in the knowledge base. The score *S*(*d*) for the disease *d* is calculated as

$$S(d) = \frac{\sum_{i=1}^{n(d)} x(i)w(d,i)f(d,i,j)}{\sum_{i=1}^{n(d)} x(i)w(d,i)},$$

where $x(i)$ is 1, if the value of i^{th} attribute is known for the disease *d*, otherwise 0; $w(d,i)$ is the weight value for the attribute *i*; and $f(d,i,j)$ is the fitness value for the value *j* of the attribute *i*. The fitness value varies from 0 to 1.

The diseases with the highest scores are the best fits and suggested by ONE. To handle uncertainty caused by missing values, ONE generates upper and lower bounds for the score. The lower bound is calculated using the lowest fitness values for the missing values and the upper bound using the highest fitness values, respectively. In addition to the scoring scheme, ONE uses necessary attribute values attached to certain diseases. In order to be diagnosed as having a certain disease, the case has to conform to the necessary attribute values set for the disease.

We use the inference results of ONE obtained in the earlier study [9]. In that study, the disease with the highest score was considered as the classification result. Neither the upper or lower bounds of the score nor the necessary attribute values were employed. The knowledge base contained descriptions for the six largest diagnostic groups. In the present study, we use the results of two scoring schemes. In the first scheme (ONE_e), the weight and fitness values were defined by experienced otoneurologists [4]. In the second scheme (ONE_d), the fitness values were calculated from the training data. For each attribute *i*, the frequency distributions were calculated within the six diagnostic groups. The fitness value $f(d,i,j)$ was calculated as the proportion of the frequency $fr(d,i,j)$ of the value *j* to the highest frequency $fr(d,i,h)$ of the distribution:

$$f(d,i,j) = \frac{fr(d,i,j)}{fr(d,i,h)}.$$

All the 170 attributes were associated with the weight 1, except those whose values were missing from all the training cases of the diagnostic group.

3.2. Nearest Neighbour Classification

Because our data are mixed having both qualitative and quantitative attributes, we wanted to use in the nearest neighbour classification a heterogeneous distance function [11] that takes into account the scales of the attributes. We chose the Heterogeneous Value Difference Metric (*HVDM*) [11] which seems to perform well with mixed real-world data [6]. *HVDM* defines the distance between two cases x and y as [11]

$$HVDM(x,y) = \sqrt{\sum_{a=1}^{A}(d_a(x,y))^2},$$

where A is the number of attributes and $d_a(x,y)$ is the distance between x and y for the a^{th} attribute.

For nominal attributes, the distance between two values x and y is defined by a simplified, normalised version [11] of the Value Difference Metric (*VDM*) [8]:

$$d_a(x,y) = \sqrt{\sum_{c=1}^{C}\left|P_{a,x,c} - P_{a,y,c}\right|^2},$$

where C is the number of output classes and $P_{a,x,c}$ is the conditional probability that the output class is c given that the a^{th} attribute has the value x.

For linear attributes, the normalised distance is the absolute difference between the values x and y divided by $4\sigma_a$, where σ_a is the standard deviation of the a^{th} attribute, i.e.,

$$d_a(x,y) = \frac{|x-y|}{4\sigma_a}.$$

The ordinal attributes were treated as nominal attributes. The distances for the nominal and the ordinal attributes, and the standard deviations for the other attributes were calculated from the training set. Classifications for the test cases were calculated on the basis of the nearest neighbour (*1NN*) and the three nearest neighbours (*3NN*) in the learning set. Three different attribute subsets were used: the subset a with all the 170 attributes ($1NN_a$ and $3NN_a$), the subset b with 38 attributes [10] ($1NN_b$ and $3NN_b$), and the subset c with five attributes [10] ($1NN_c$ and $3NN_c$).

4. Results

The true positive rate (TPR_c) was calculated for each diagnostic group as the percentage of correctly classified cases of the group:

$$TPR_c = 100\frac{tpos_c}{pos_c}\%,$$

where $tpos_c$ is the number of correctly classified cases of the group c and pos_c is the number of all cases in the group. Table 1 presents the true positive rates and their medians for the inference method of ONE and the nearest neighbour method.

Table 1: True positive rates (%) for ONE and nearest neighbour classification. e = scoring scheme defined by experts, d = scoring scheme learned from data, a = 170 attributes, b = 38 attributes, c = 5 attributes

Diagnosis	N	ONE_e	ONE_d	$1NN_a$	$1NN_b$	$1NN_c$	$3NN_a$	$3NN_b$	$3NN_c$
Benign positional vertigo	44	59.1	65.9	68.2	63.6	72.7	68.2	65.9	70.5
Menière's disease	94	62.8	94.7	92.6	92.6	81.9	94.7	95.7	86.2
Sudden deafness	12	66.7	66.7	25.0	50.0	16.7	16.7	25.0	8.3
Traumatic vertigo	20	95.0	80.0	70.0	75.0	70.0	60.0	85.0	80.0
Vestibular neuritis	37	81.1	75.7	91.9	89.2	75.7	86.5	86.5	83.8
Vestibular schwannoma	39	30.8	66.7	76.9	66.7	71.8	76.9	61.5	59.0
Median		64.8	71.2	73.5	70.9	72.3	72.6	75.5	75.3

The prediction accuracy (*ACC*) was calculated as the percentage of correctly inferred cases:

$$ACC = 100\frac{\sum_{c=1}^{C} tpos_c}{\sum_{c=1}^{C} pos_c}\%,$$

where C is the number of diagnostic groups. The prediction accuracies for ONE_e and ONE_d were 62.6% and 79.7%, respectively [9]. The $1NN_a$, $1NN_b$, $1NN_c$, $3NN_a$, $3NN_b$, and $3NN_c$ classifications produced the prediction accuracies of 80.5%, 79.3%, 73.6%, 79.3%, 79.3%, and 74.4%, respectively.

5. Discussion

The nearest neighbour technique was used to classify otoneurological data on the basis of the nearest (*1NN*) and the three nearest neighbours (*3NN*). In the classification, three attribute subsets were used. The best true positive rate (*TPR*) obtained by *1NN* varied from 50.0% (sudden deafness) to 92.6% (Menière's disease). Benign positional vertigo was classified best with the five attributes (72.7%). The 38 attributes produced the best *TPR* for sudden deafness (50.0%) and traumatic vertigo (75.0%). For Menière's disease, the subsets of 38 and 170 attributes yielded the best result (92.6%). The largest attribute set with 170 attributes produced the best results for vestibular neuritis (91.9%) and vestibular schwannoma (76.9%). The above phenomena relate to the problem of attribute or feature subset selection [2,5]: The largest attribute set is not necessarily the optimal one. The search of the three nearest neighbours increased the true positive rate for Menière's disease, whereas for sudden deafness, it decreased *TPR*.

With respect to the prediction accuracy, the classification results of the nearest neighbour technique and the inference mechanism of ONE agreed. The best predictive accuracy for the expert system ONE was 79.7% and for the nearest neighbour method 80.5%. The better true positive rates for vestibular neuritis and vestibular schwannoma obtained by $1NN_a$ suggest that the refinement of their weight and fitness values might enhance ONE's inference capability. The possibilites given by ONE's scoring scheme can be seen in the true positive rates for sudden deafness and traumatic vertigo.

Future work will include a thorough analysis of the nearest neighbour classification results. These results will be used to refine the knowledge base of ONE.

Acknowledgements

The authors thank Yrjö Auramo, Erna Kentala, and Erkki Mäkinen for their contribution to the study. The work of the first author was funded by Tampere Graduate School in Information Science and Engineering (TISE) and by a grant from Emil Aaltonen Foundation.

References

[1] Y. Auramo *et al.*, An expert system for the computer-aided diagnosis of dizziness and vertigo, Medical Informatics **18** (1993) 293-305

[2] M. Dash and H. Liu, Feature selection for classification, Intelligent Data Analysis **1** (1997) 131-156

[3] E. Kentala, A neurotologic expert system for vertigo and characteristics of six otologic diseases involving vertigo. Academic Dissertation, Department of Otorhinolaryngology, University of Helsinki, Finland (1996)

[4] E. Kentala, Characteristics of six otologic diseases involving vertigo, The American Journal of Otology **17** (1996) 883-892

[5] R. Kohavi and G.H. John, Wrappers for feature subset selection, Artificial Intelligence **97** (1997) 273-324

[6] J. Laurikkala and M. Juhola, Nearest neighbour classification with heterogeneous proximity functions, In A. Hasman *et al.* (eds.), Medical Infobahn for Europe: Proceedings of MIE2000 and GMDS2000, Studies in Health Technology and Informatics, Vol. 77. IOS Press, Amsterdam (2000) pp. 753-757

[7] T.M. Mitchell, Machine Learning. McGraw-Hill, New York (1997)

[8] C. Stanfill and D. Waltz, Toward memory-based reasoning, Communications of the ACM **29** (1986) 1213-1228

[9] K. Viikki and M. Juhola, Refining the knowledge base of an otoneurological expert system, In J. Crespo *et al.* (eds.), Medical Data Analysis, Lecture Notes in Computer Science, Vol. 2199. Springer-Verlag, Berlin Heidelberg (2001) pp. 276-281

[10] K. Viikki *et al.*, Decision tree induction in the diagnosis of otoneurological diseases, Medical Informatics & The Internet in Medicine **24** (1999) 277-289

[11] D.R. Wilson and T.R. Martinez, Improved heterogeneous distance functions, Journal of Artificial Intelligence Research **6** (1997) 1-34