OPTIMIZATION OF A NEAREST NEIGHBORS CLASSIFIER FOR DIAGNOSIS OF CONDITION OF ROLLING BEARINGS

Maciej TABASZEWSKI
Institute of Applied Mechanics, Poznan University of Technology
Ul. Jana Pawła II 24, 60-965 Poznań
fax. 61 665 2307, email Maciej.Tabaszewski@put.poznan.pl

Summary
The paper concerns classification of technical condition state of rolling bearings. A methodology of optimization of a k-NN classifier with regard to selection of the symptom observation space has been proposed. The symptoms carrying the most information allowing identification of a class of technical condition were selected. The applied methodology enabled to develop a classifier which in the set of available data achieved the efficiency of 97.5%. In the set of considered symptoms the r.m.s. and peak values of vibration acceleration in the broad frequency band and the energy of acoustic emission pulses turned out to be the best for identification of arising fracture of a bearing outer ring.

Keywords: diagnosis of rolling bearings, classification of condition.

OPTYMALIZACJA KLASYFIKATORA NAJBLIŻSZYCH SĄSIADÓW DLA DIAGNOZOWANIA STANU ŁOŻYSK TOCZNYCH

Streszczenie
Praca dotyczy klasyfikacji stanu technicznego łożysk tocznych. Zaproponowano metodykę optymalizacji klasyfikatora k-NN ze względu na dobór symptomowej przestrzeni obserwacji. Wyłoniono symptomy, które niosą najwięcej informacji pozwalającej na identyfikację klasy stanu technicznego. Zastosowana metodyka pozwoliła opracować klasyfikator, który na zbiorze dostępnych danych osiągał efektywność rzędu 97,5%. W zbiorze rozpatrywanych symptomów najlepszymi do identyfikacji powstającego pęknięcia pierścienia zewnętrznego łożyska okazały się wartość skuteczna i szczytowa przyspieszeń drgań w szerokim pasmie częstotliwości oraz energia impulsów emisji akustycznej.

Słowa kluczowe: diagnostyka łożysk tocznych, klasyfikacja stanu.

1. INTRODUCTION

It can be estimated that about 80% of machines include rolling bearings in their structure. The bearings are, therefore, a very important element in the context of maintenance of machines. Apparently the rolling bearings cause many failures in the industry – about 80% [1]. The most often causes of premature failures of rolling bearings are [2]: poor lubrication conditions (ca. 36%), lubricant contamination (ca. 14%), assembly errors (16%), and overload. These factors may cause that the rolling bearings do not achieve their nominal durability specified by the manufacturer. According to literature data ca. 66% of rolling bearings do not achieve their nominal durability because of the above factors [2].

In order to determine the technical condition of a rolling bearing unambiguously it may be necessary to observe many diagnostic symptoms simultaneously. Building a cause and effect models is very difficult in this case. It seems a lot simpler to create a classifier in the form of a system learning on examples and to use it to determine the technical condition. Such a classifier may also be used to obtain more detailed information about the reasons of damage. It requires, however, a great number of learning examples encompassing all possible types of damage.

A classifier enabling to determine one of two classes of condition: good / defective will be considered in the paper. This seems to be sufficient, taking into account the database of the considered examples concerning failures of rolling bearings.

The database, built based on the conducted experiment, includes cases of fatigue cracking of rolling bearings' outer rings. Hence, the information good / defective with simultaneous knowledge of the cause of failure is sufficient. The main goal of the consideration is an in depth analysis of possibility of application of the k-nearest neighbor (k-NN) method for diagnosis of rolling bearings based on the available database and an attempt to obtain information on the importance of individual symptoms. The proposed methodology should allow to determine the symptoms carrying the most information about the approaching fracture of the rolling bearing's outer ring.

Among many existing methods of classification the k-NN method was chosen in the paper, because of its simplicity, which is very important, if a diagnostic system working in industrial environment is to be built in the future. Moreover, such a method seems to be sufficient to determine diagnostic symptoms being of particular importance in case of identification of the considered failure of a rolling bearing.

It is worth mentioning, that, of course, there exist the possibility to determine the most important attributes (diagnostic symptoms in this case) by applying particular measures, such as Fisher criterion or information entropy. The obtained results, however, may differ significantly dependent on the applied measure. Hence, it is worthwhile to propose other methodology enabling to find the most significant symptoms in case of application of the nearest neighbor distance classifier.

2. *K-NN* CLASSIFIER AND ITS TESTING

Classification as one of the basic methods of data exploration is used by many researchers to solve many different problems [3, 4, 5, 6, 7, 8, 9, 10, 12].

An important group of methods of classification of state of an object are distance methods based on the assumed measure of distance in a multidimensional space of features. A variety of measures of distance are used in these methods. As a basic measure one should mention the Minkowski metric, defined as follows:

$$\rho_M\left(\boldsymbol{S}_i, \boldsymbol{S}_j\right)=\left[\sum_{k=1}^{N}\left|S_{ik}-S_{jk}\right|^{m}\right]^{\frac{1}{m}} \tag{1}$$

where $\boldsymbol{S}_i$ = [S_{i1} S_{i2} ... S_{iN}] i $\boldsymbol{S}_j$ = [S_{j1} S_{j2} ... S_{jN}] are vectors describing particular measurement record, and N is the number of features. In this paper the vector elements are the values of the measured symptoms. When m = 2, we obtain an Euclidean measure, and when m = 1, a Manhattan measure (called also *city block*). These two commonly used measures will also be used in this work.

The type of a metric is selected experimentally. In the distance methods it is essential that it is easy to implement additional weights reflecting the influence of individual symptoms on the final recognition of the condition of an object (so called stretching of axes of coordinate systems in the symptom space). As an example, in comparison with other vibration symptoms, kurtosis of vibration acceleration may theoretically be significant in recognition of the condition of a rolling bearing and it may be given a higher value of weight.

The functional algorithm of the classifier is very simple and it comes down to determination of the class of the diagnosed object based on its affiliation to the class of examples from its immediate vicinity. The immediate vicinity is determined by the selected measure of distance. Among several types of the algorithm the methods of simple and weighted voting are worth mentioning. In case of simple voting only the number of representatives of the class in the immediate vicinity of the considered object is important, and in case of weighted voting the distances from the objects being the training examples are important as well.

An important element in building the classifier is its evaluation. This evaluation is connected with division of the collected set of data into a training set and testing set. The training set is used to build the classifier. As a result of testing a system which enables to achieve acceptable errors in recognition of real classes will be created. At this stage information about both conditional attributes (here: the values of diagnostic symptoms / working parameters) and decision attributes (here: the class of state: good / defective) is used. The testing set, which is not used for training, is used to evaluate the classifier.

Ability to generalize the model is evaluated based on the recognition of testing examples. In such a way it is possible to compare the considered models, evaluate each of them, and choose the one that enables the best generalization. In practice different tests are used. Cross-validation, holdout, leave-k-out and leave-one-out [8] are worth mentioning here. Because of the small set of data (80 examples) the last method is taken into account in this paper.

To evaluate the classifier a measure of a classification error (a combined classification error):

$$\varepsilon=\frac{n_F}{n} \tag{2}$$

or accuracy of classification:

$$\eta=1-\frac{n_F}{n}=\frac{n_P}{n} \tag{3}$$

can be used, where: n_F - number of erroneous classifications, n_P – number of accurate classifications, and n – number of testing examples.

Among evaluations of classifiers used in technical diagnostics the ones which take into account different costs of erroneous classifications may be of particular importance. In case of a two-state classification two terms are introduced: negative recognition connected with the state "good object", and positive recognition connected with the state "defective object". One should remember that as a result of functioning of the classifier in most cases we obtain a number of erroneous diagnoses. If the diagnoses concern the state recognized erroneously as “good object”, we can talk about false negatives (*FN*). A state can, in turn, be classified as “defective object”, although it is really good. In such a case we can talk about false positives (*FP*). Based on the mentioned measures the following rates are created: false negative rate, false positive rate, or sensitivity and specificity. To evaluate a classifier one can also use a confusion matrix analysis [8], where different costs of an erroneous classification can be assigned to the elements outside the diagonal.

Further in the paper the combined classification error and additionally the FP to FN ratio will be used to evaluate the classifier. If the ratio approaches one it means that the classifier makes similar number of errors FP and FN. The values near zero mean that the erroneous diagnosis of a good condition dominates over the cases of the erroneous diagnosis of a defective condition.

3. CLASSIFICATION OF THE CONDITION OF ROLLING BEARINGS AND METHODOLOGY OF SYMPTOM SPACE OPTIMIZATION

The data concerning rolling bearings were obtained from an accelerated wear experiment. Type 608 rolling bearings were forced to breakdown by causing their axial overload. The database was built based on the obtained life curves (trends of changes of vibration symptoms) by selection of symptom values long before breakdown (the vectors were labeled – “god condition”) and just before breakdown (the vectors were labeled – “defective condition”). The examples acquired just before breakdown are identified as a conventional failure of a rolling bearing, as practically in the subsequent measurement steps the breakdown occurred. Additionally such an approach enables to build a classifier which in practice warns of an emergency situation.

In the end it was possible to collect 40 examples concerning the good condition and 40 ones concerning the defective condition. Unfortunately the sample size is small. Due to duration and difficulty of the experiment, however, it has been decided that at this stage such a sample must do for initial comparative analyses.

The following parameters / symptoms were taken into account in the database:

1. Bearing load (the controlled working parameter, turned on and off during the experiment).
2. Kurtosis of vibration acceleration in a broad band (to 10 kHz).
3. Peak factor of vibration acceleration in the same broad band as above.
4. R.M.S. value of vibration acceleration in the same broad band as above.
5. Peak value of vibration acceleration in the same broad band as above.
6. Count-rate of acoustic emission pulses.
7. Energy of acoustic emission pulses.

The paper considers a classifier which uses weighted voting, where weights depend on the distance from the nearest neighbors. From the comparative analyses made it resulted that another classifier *k*-NN – with simple voting – gave worse results.

Table 1. An example of nearest neighbor classifier test results for different types of the classifier

Metric	Classification error [%]	FP/FN
Manhattan	8,10	0,42
Euclidean	8,19	0,52

It results from Table 1 that for the considered set of data and non-optimized classifier irrespective of the type of the metric the classification errors equal about 8%. In terms of error structure the classifier with the Euclidean metric gives much better results (better FP/FN ratio) which means that using such a classifier we are less exposed to more expensive errors (the cost of overlooking the breakdown when it really occurred are proportional to FN and is usually much higher than the cost of stopping the machine to check it up – proportionality to FP). A question arises whether it is possible to obtain better classification results based on the available data. It seems obvious that in the first step one may optimize parameter *k* of the algorithm instead of imposing it arbitrarily. Using a cross-validation test for different *k* from the range between 3 and 12 (the range was assumed arbitrarily) it is possible to find the optimal value of this parameter for individual methods – the results are shown in Table 2.

As it can be seen from Table 2 by the appropriate choice of parameter *k* it is possible to reduce slightly the classification error for the considered classifiers.

Table 2. Optimization of *k*-parameter selection for the considered problem and all of the considered symptoms

Metric	Optimal k value	Classifier error for the optimal k [%]	FP/FN
Manhattan	11	8,02	0,53
Euclidean	3	7,12	0,44

Here it arises another possibility. The classification error may also be reduced by means of an appropriate selection of symptoms. At the same time it is possible to show which symptoms are the most important diagnostically, and which ones do not influence the quality of classification or even worsen it (introduce ambiguous diagnostic information).

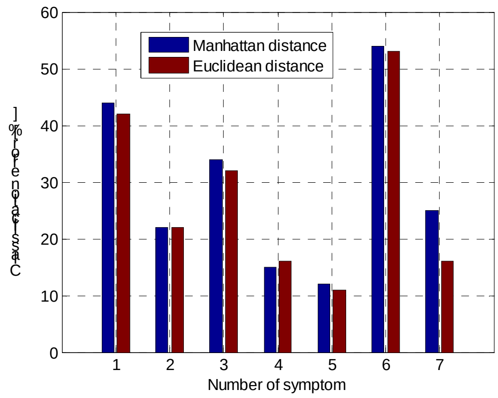

Fig.1. Classification errors for single-symptom-classifiers k -NN with weighted voting

In order to verify the method for the considered data, first, an evaluation of classification quality was made for classifying systems based only on one symptom and the optimal value of parameter k. The procedure was made for each individual symptom. The results are shown in Figures 1 and 2. For instance, an attempt at building classifiers based only on the acoustic emission count-rate (symptom 6) does not give positive results (the error reaches even 50%, which can also be achieved by making decisions by throwing a coin). Hence, it follows that the acoustic emission count-rate reaches different values both for good and defective rolling bearings and it does not carry any diagnostic information in this case.

Taking into account the obtained results shown in Figures 1 and 2 it can be stated that it is the best to infer about the state of a rolling bearing (considering the available set of symptoms) based on symptoms 4 and 5, i.e. on the r.m.s. and peak values of vibration acceleration in the frequency band to 10 kHz. In any case it is coherent with what is reported in the literature [11]. Moreover, classifiers based on these symptoms make similar number of errors FN and FP (the ratio equals near 1 for the Manhattan metric), i.e. the probabilities of not detecting the breakdown and of premature shut-down are similar. As a matter of fact it is not necessary an advantage in case of technical diagnostics.

As can be seen it is possible to select the best symptoms in this way. The obtained single symptom classifier, however, gives unacceptable errors of classification (over 10%). A question arises whether thanks to the made evaluation of individual symptoms it is possible to classify better the data.

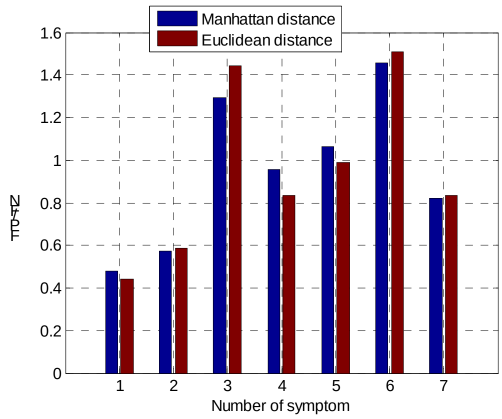

Fig.2. FP/FN ratio for single-symptom-classifiers k -NN with weighted voting

At first approach a method of weighting the individual symptoms depending on the obtained error of the single symptom classifier may be proposed. The final value of a weight should be inversely proportional to the error, enabling to emphasize the significance of the best (giving the best division of classes of condition) symptoms.

Table 3. Results obtained for a classifier based on all the symptoms, where the importance of each symptom depended on the error of the single symptom classifier

Metric	Optimal k value	Classifier error for the optimal k [%]	FP/FN
Manhattan	3	7,8	0,54
Euclidean	4	9,1	0,57

The results obtained in such a way, where the weights are used to stretch the axes of the system describing the symptom space are shown in Table 3. In Table 4, however, the weights are considered to be inverse squares of the single classifiers errors.

As can be seen from both tables, consideration of appropriate weights does not solve the problem. Generally, the obtained results are similar or worse than the results obtained when all the symptom were considered without any weights (cf. Table 2). Only the FP/FN ratio became insignificantly better.

Table 4. Results obtained for a classifier based on all the symptoms, where the importance of each symptom depended on the square error of the single symptom classifier

Metric	Optimal k value	Classifier error for the optimal k [%]	FP/FN
Manhattan	4	8,2	0,58
Euclidean	5	8,1	0,55

In order to improve the quality of state recognition another methodology may be proposed. The symptoms with the worst results of evaluation can be removed successively based on evaluation of the single symptom classifiers (cf. Figure 1). Every time, however, parameter *k* should be selected in such a way that the minimum error of the classifier is obtained. The results of the proposed procedure, where the symptoms associated with "the worst" single symptom classifiers were successively removed, are shown in Figures 3 and 4. As can be seen from Figure 3 removal of only symptom 6 from the set of considered ones already improves the situation (reduction of the error by more than 1%). Removal of a group of symptoms 6, 3, 2 (count-rate of AE pulses, peak factor of vibration acceleration and kurtosis of vibration acceleration in a broad band) and the working parameter (1 – bearing load) enables to obtain a minimum error of about 2.5%, i.e. about three times less than in case of taking all the symptoms into account. It follows that the rejected symptoms identify contradictory classes, i.e. both big and small values of a symptom concern both the good and defective conditions. Removal of further symptoms worsen the accuracy of recognition. Hence, the remaining symptoms from among the considered ones: r.m.s. value of vibration acceleration, peak value of vibration acceleration and energy of acoustic emission pulses are indispensable for correct recognition of the technical state of a rolling bearing (with minimum error of about 2.5% possible for the considered data). With an optimal selection of symptoms the classifier does not make any errors connected with premature recognition of a breakdown, and all the erroneous classifications are connected with overlooking the breakdown. Unfortunately this is a disadvantageous feature. After deeper analysis of the data, however, it can be realized that it results from the lack of information about the approaching breakdown in several examples. If such a situation, where the breakdown occurs without any clear "warnings", occurs in practice, the erroneous classification is unavoidable.

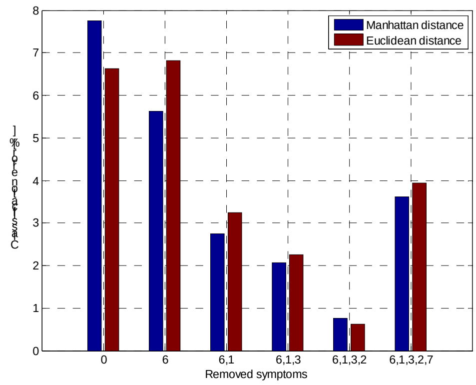

Fig. 3. Changes of classification error in consecutive steps of rejecting useless symptoms (0 means that no symptom was rejected)

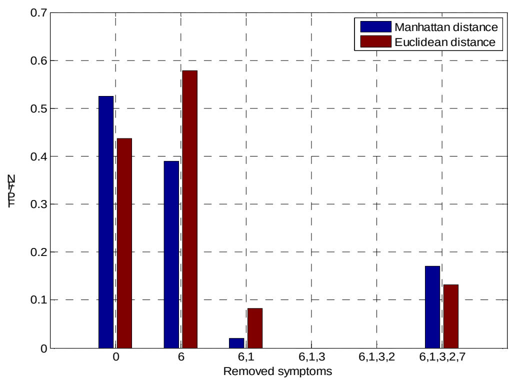

Fig. 4. Changes of the FP/FN ratio in consecutive steps of rejecting useless symptoms (0 means that no symptom / parameter was rejected)

4. CONCLUSIONS

As a conclusion it can be stated that the described methodology applied to the considered data enables to select symptoms carrying the most diagnostic information (are of the greatest significance in recognition of a condition class). Among the collected symptoms the r.m.s. and peak values of vibration acceleration in the broad frequency band (to 10kHz), and energy of acoustic emission pulses are the most significant diagnostically for the considered rolling bearings. It is coherent with some literature data. However, negative influence of consideration of kurtosis of vibration acceleration or even the working load is surprising. This may result from specificity of the process of damaging the rolling bearings. The usage of the optimal set of symptoms for diagnosis of type 608 rolling bearings (under conditions of the conducted experiment) enables to achieve a small error of recognition (about 2.5%) minimizing the possibility of false recognition of the defective state. Unfortunately erroneous recognition of good condition for the considered symptoms can take place, if the symptoms do not react to the deterioration of the technical condition. Optimization of the *k*-NN neighbors classifier with regard to the classifier parameter, type of the metric, or symptom selection will not enable then to avoid the erroneous identification of the class of condition.

REFERENCES

[1] Materials of AS INSTRUMENT POLSKA available at http://www.asinstrument.com.pl/
[2] S. Szymaniec, *Diagnostics of motor rolling bearings in conditions their industrial operating* (in polish) Maszyny Elektryczne Nr 74/2006
[3] Cholewa W., Kiciński J., *Diagnostyka techniczna, odwrotne modele diagnostyczne*, Gliwice, Wydawnictwo Politechniki Śląskiej 1997.
[4] Schurmann J., *Pattern Classification. A Unified View of Statistical and Neural Approaches*, New York, Wiley 1996.
[5] Demuth H., Beale M., *Neural Network Toolbox. For Use with MATLAB, User's Guide ver. 4*, The MathWorks 2001.
[6] Zanasi A., Brebbia C. A., Ebecken N. F. E. E, Melli P. (red.), *Data Mining III*, Sunthampton, Boston, WIT Press 2002.
[7] Kornacki J., Ćwik J., *Statystyczne systemy uczące się*, Warszawa, WNT 2005.
[8] Korbicz J., Kościelny J. M., Kowalczuk Z., Cholewa W. (red.), *Diagnostyka procesów. Modele. Metody sztucznej inteligencji. Zastosowania*, Warszawa, WNT 2002.
[9] Therrien Ch. W., *Decision Estimation and Classification*, New York, Wiley 1989.
[10] Szczepaniak P., *Obliczenia inteligentne, szybkie przekształcenia i klasyfikatory*, Warszawa, Akademicka Oficyna Wydawnicza EXIT 2004.
[11] Cempel C., *Vibroacoustic condition monitoring*, Ellis Horwood, 1991, p. 212.
[12] Thair Nu Phyu, *Survey of Classification Techniques in Data Mining*, Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I, IMECS 2009, March 18 - 20, 2009, Hong Kong.

Maciej TABASZEWSKI, D.Sc. Ph.D. Eng. is an assistant professor at the Institute of Applied Mechanics, Poznan Univer-sity of Technology.
His current research interest include condition monitoring, data mining and methods of forecasting.