# A Multiclassifier Approach for Drill Wear Prediction

Alberto Diez[1], Alberto Carrascal[1]

[1] Tecnalia Research & Innovation, Paseo Mikeletegi 7, Parque Tecnológico, 20009 Donostia-San Sebastián (Guipúzcoa, España)

**Abstract.** Classification methods have been widely used during last years in order to predict patterns and trends of interest in data. In present paper, a multiclassifier approach that combines the output of some of the most popular data mining algorithms is shown. The approach is based on voting criteria, by estimating the confidence distributions of each algorithm individually and combining them according to three different methods: confidence voting, weighted voting and majority voting. To illustrate its applicability in a real problem, the drill wear detection in machine-tool sector is addressed. In this study, the accuracy obtained by each isolated classifier is compared with the performance of the multiclassifier when characterizing the patterns of interest involved in the drilling process and predicting the drill wear. Experimental results show that, in general, false positives obtained by the classifiers can be slightly reduced by using the multiclassifier approach.

**Keywords:** Classification, multiclassifier, drill wear prediction, pattern identification.

## 1  Introduction

In the machine-tool sector, emerging industrial processes and methodologies due to technological improvements in manufacturing lines require new valid solutions when detecting failures and scheduling maintenance operations [1]. Preventive and corrective maintenance procedures have been developed and widely used during the last years; nevertheless, there yet exist important maintenance gaps to be fulfilled [2]. The research activity done in this field is focused on how to predict that something unexpected is going to happen before it really occurs, with the aim of avoiding overhead costs derived from production line breakdowns and maintenance operations. Machinery builders and vendors usually provide preventive maintenance strategies, but they are focused on normal operating conditions [3]. Benefits derived from anomaly prediction are not only oriented to reduce costs and to optimize the machine lifecycle, but also to infer new relevant knowledge about the process and the most probable cause and propagation of the problem.

The motivation regarding drill wear detection and prediction in the machine-tool processes is mainly focused on the loss of quality of the resulting holes as the drill used is close to the decline stage of its lifecycle. Into this work, new diagnosis methods regarding drill wear characterization and classification by means of a

multiclassifier approach is presented. Many authors have adopted approaches based on the combination of classifiers for resolving different problems and by using different combination rules and strategies [4]. The multiclassifier proposed in this study combines the outputs of different data mining techniques based on voting criteria and on the given label distribution, therefore more accurate predictions can be achieved. The goodness of this approach is compared with results obtained by each isolated classifier when addressing the drill wear problem, characterized by the presence of burr and roughness on holes drilled. Experiments performed consist on several drilling experiments, in which several process parameters are monitored and analyzed to extract the most relevant patterns associated to drill wear problem. It is demonstrated that the use of data mining algorithms provides a promising methodology and decision making support tool regarding drill substitution strategy.

The layout of the paper is as follows: the problem under study and its characterization is introduced in Section 2; the data mining algorithms used and the multiclassifier approach are shown in Section 3; in Section 4 experimental results are analyzed and discussed; finally, conclusions of this study are presented in Section 5.

## 2      The Drill Wear Problem

The quality of the resulting workpiece in drilling process can be estimated by means of two physic parameters: burr and roughness in the resulting holes drilled [5]. The more wear is the drill, the less quality is obtained. For this reason, over certain threshold values it is strongly recommended to replace the drill in order to assure enough quality and to avoid its breakage, which can provoke serious safety problems and maintenance costs.

The main motivation of this paper is to analyze how drill wear influences in the quality of the hole drilled and how to characterize it by means of most relevant patterns inferred from monitoring data. Several experiments regarding drilling process have been monitored in order to acquire the data needed to tackle with this study.

### 2.1      The Drilling Process

During drilling process, the intensity of the engine regulator of machine spindle and the intensity of the engine regulator of machine head are measured. The first one is related to force needed in each drilling operation, F, and the second one is the torque, T. The signals assessed show four significant points of interest that delimitate three cutting areas, as showed in Fig. 1. Those four points are the following:

1. When the drill head comes into contact with the workpiece and is introduced into the material; the intensity of the signal increases.
2. When all drill diameter is inside the workpiece.
3. When the drill head exits from the surface of the workpiece; the intensity of the signal decreases.
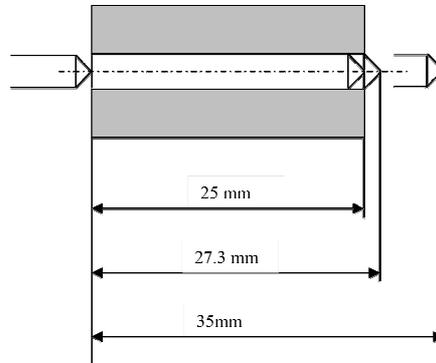4. When the drill is totally outside the workpiece.

**Fig. 1**. Cutting areas of interest during a drilling process.

Given intensity signals, F and T, obtained during each drilling operation for each critical cutting area, the variables of interest that characterize the drilling process and that can be affected by drill wear are calculated. Those variables are briefly described in Table 1.

**Table 1.** Variables measured during the drilling process.

| Attribute | Description |
|---|---|
| F23_MEAN | Advance force mean value when all the drilling tool diameter is drilling inside the workpiece |
| F23_STDE | Advance force standard deviation when all the drilling tool diameter is drilling inside the workpiece |
| F23_SKEW | Advance force bias when all the drilling tool diameter is drilling inside the workpiece |
| T23_MEAN | Torque mean value when all the drilling tool diameter is drilling inside the workpiece |
| T23_STDE | Torque standard deviation when all the drilling tool diameter is drilling inside the workpiece |
| T23_SKEW | Torque bias when all the drilling tool diameter is drilling inside the workpiece |
| F12_SLAN | Force slope at the entry point of the workpiece |
| F12_TIME | Time measured from the drilling tool makes contact with the workpiece until all the drilling tool diameter is drilling inside the workpiece, in terms of force |
| T12_SLAN | Torque slope at the entry point of the workpiece |
| T12_TIME | Time measured from the drilling tool makes contact with the workpiece until all the tool diameter is drilling inside the workpiece, in terms of torque |
| F23_A001 | Force value area under the FFT in the following rank of frequency: 50 Hz – 250 Hz |
| T23_A001 | Torque value area under the FFT in the following rank of frequency: 50 Hz – 250 Hz |
| Stde_Fin | Drilling force standard deviation when all the drilling tool diameter is drilling inside the workpiece, at the beginning of the hole |
| Fin_MEAN | Drilling force mean value when all the drilling tool diameter is drilling inside the workpiece, at the beginning of the hole |
| Stde_Fout | Drilling force standard deviation when all the drilling tool diameter is drilling inside the workpiece, at the end of the drilling process |
| Fout_MEAN | Drilling force mean value when all the drilling tool diameter is drilling inside |

| | the workpiece, at the end of the drilling process |
|---|---|
| *Delta_F* | Difference between the drilling force mean value at the beginning of the hole and when finishing the drilling process |
| *Stde_Tin* | Torque standard deviation when all the drilling tool diameter is drilling inside the workpiece, at the beginning of the hole |
| *Tin_MEAN* | Torque mean value when all the drilling tool diameter is drilling inside the workpiece, at the beginning of the hole |
| *Stde_Tout* | Torque standard deviation when all the drilling tool diameter is drilling inside the workpiece, at the end of the drilling process |
| *Tout_MEAN* | Torque mean value when all the drilling tool diameter is drilling inside the workpiece, at the end of the drilling process |
| *Delta_T* | Difference between the torque mean value at the beginning of the hole and when finishing the drilling process |
| *Roughness* | The roughness of the workpiece, measured in microns |
| *Burr* | The burr of the workpiece, measured in microns |

During different drilling experiments torque evolution, T, shows a slight variation as the number of holes drilled increases: torque measurement is hardly increased at the end of the drill. This evolution is shown in Fig. 2. At the moment the drill exits from the workpiece, force evolution, F, also shows an ascendant trend in relation with the number of holes drilled, as can be appreciated in Fig. 3.

It is assumed that drill wear effect can be easily estimated by measuring burr and roughness of holes drilled, as can be seen in Fig. 4, given the fact those measurements strongly and critically fluctuates as the number of holes drilled increases. Therefore, in this study the quality of hole drilled has been characterized in terms of burr and roughness.
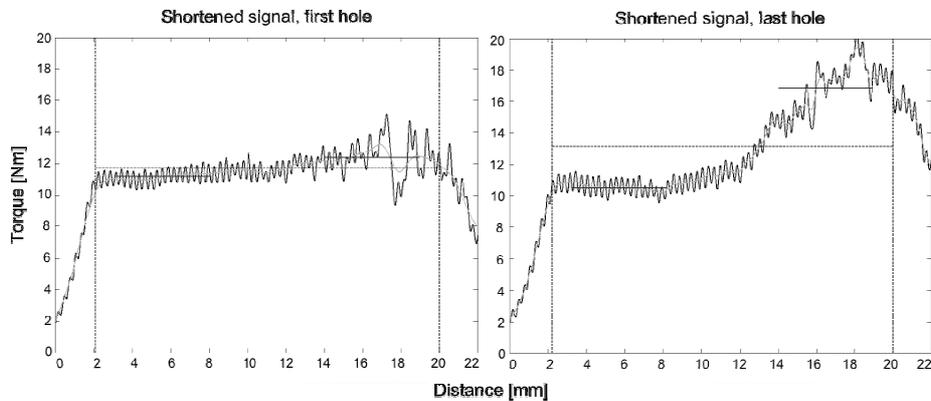


**Fig. 2.** Torque signal evolution among different drilling experiments: torque signal in the first hole drilled (left) and in the last the drilling operation performed (right).
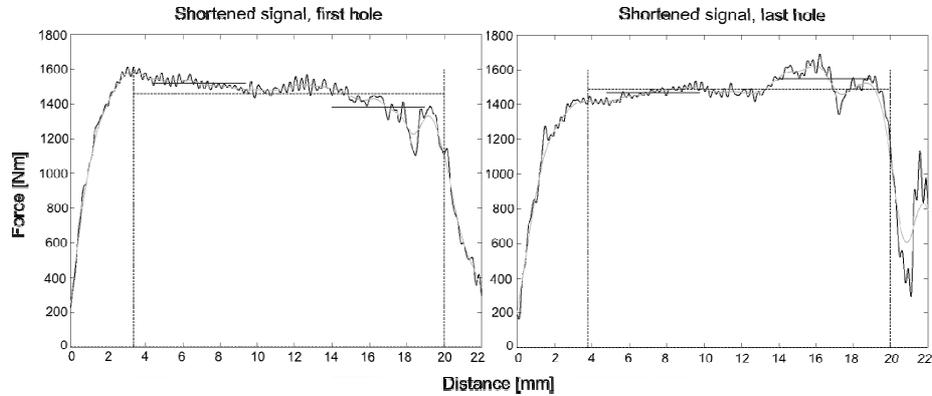
**Fig. 3**. Force signal evolution among different drilling experiments: force signal in the first hole drilled (left) and in the last drilling process performed (right).
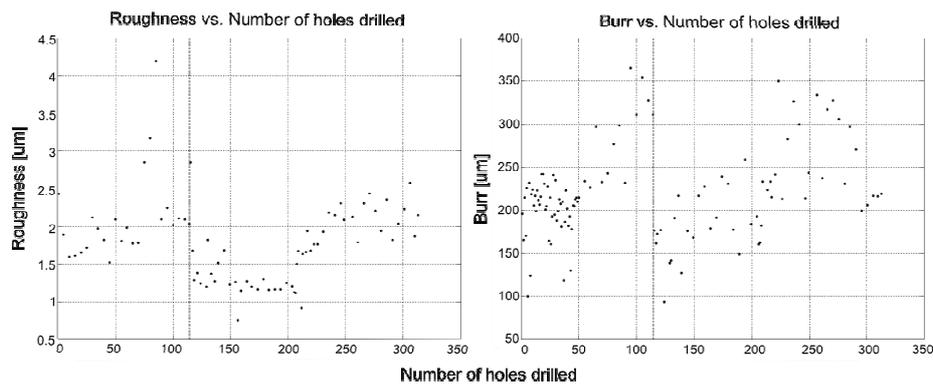


**Fig. 4**. Graphics of roughness (left) and burr (right) vs. number of drilling processes performed.

## 3    Methodology

The proposed drill wear prediction methodology accomplished the following steps: 1) filtering samples that contained incorrectly measured signals when gathering data from the drilling processes performed, 2) applying feature selection algorithm, PCA, to try to reduce search space and to prioritize problem-related attributes, 3) performing supervised classification by means of different data mining algorithms and 4) combining the information from these classifiers using different strategies, based on proposed multiclassfier approach.

### 3.1    Feature Selection

The number of predictor variables is usually huge and the amount of information provided is different depending on the variable. It is fairly common to apply feature

selection methods in order to identify patterns in data and, thus, to reduce the number of dimensions of the problem. Due to its capability of establishing any correlation among the parameter values without much loss of information, a Principal Component Analysis (PCA [6]) has been applied. Therefore, dimensionality reduction of data can be accomplished by eliminating the principal components with low variance. This means that they are not relevant with regards to the problem.
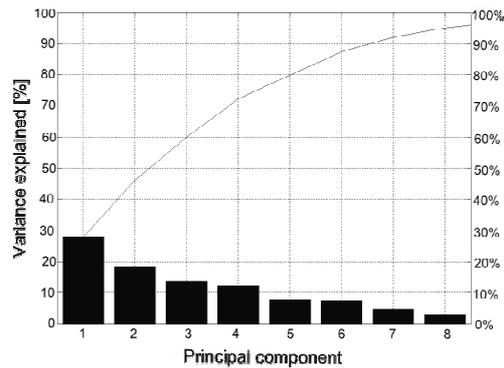


**Fig. 5.** PCA results: variance histogram.

In case under study, since the distribution in the new system of reference shows a low correlation among variables as can be seen in Fig. 5, it is not finally used. All variables provide enough information to outline the drill wear problem.

### 3.2    Supervised Classification

Classification algorithms employed in this study have in common that all of them performs a supervised classification. This is due to the nature of the data set used, in which each experiment or sample is classified according to resulting burr and roughness values. There exist other classification methods, such as non-supervised or semi-supervised classification, whose applicability is more oriented to detect anomalies when knowledge of data behaviour is limited or even inexistent [7].

The supervised classification paradigm [8] consists of a set of $N$ facts, each of them characterized by $n+1$ variables; first n variables, $X_1, X_2,..., X_n$, would be predictor variables and the variable with index $n+1$, identified as $C$, would be the *class* variable. These data can be represented in table format using the following notation:

1. $x_i^j$ is the value that the $j$-th fact takes in the $i$-th predictor variable, $i = 1, \ldots, n$ and $j = 1, \ldots, N$;
2. $c^j$ is the class that the $j$-th fact belongs to.

Facts can be also named as cases or instances and the variables are the attributes of the supervised problem. The goal is to obtain a classification model that is able to predict the value of variable $C$ when a new case is analyzed. This new case will be composed of $n$ predictor variables and an unknown value of variable $C$.

Consequently, the problem consists on correctly classifying a new fact based on previous evidences or cases.

**Table 2.** Supervised classification elements.

|   | $X_1$ | $X_2$ | $\cdots$ | $X_i$ | $\cdots$ | $X_n$ | $C$ |
|---|---|---|---|---|---|---|---|
| 1 | $x_1^1$ | $x_2^1$ | $\cdots$ | $x_i^1$ | $\cdots$ | $x_n^1$ | $C^1$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ |
| $J$ | $x_1^j$ | $x_2^j$ | $\cdots$ | $x_i^j$ | $\cdots$ | $x_n^j$ | $C^j$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ |
| $N$ | $x_1^N$ | $x_2^N$ | $\cdots$ | $x_i^N$ | $\cdots$ | $x_n^N$ | $C^N$ |

In order to accomplish this task properly, data set should be wide enough since the more cases are available about the problem the more accurate will be the classification model. Applications based on supervised classification methods are very common as, in general, any complex system can be monitored obtaining data that are characterized and used to train a model able to classify new data based on previous evidences.

### 3.3 Classification Algorithms

According to [9], six of the most popular supervised data mining algorithms have been tested in present study. They are the following:

1. Statistical learning methods. Naïve Bayes [10]: is a Bayesian algorithm built by the assumption of the conditional independence among the predictor variables, given the class value.
2. Instance-based learning methods. $k$-Nearest Neighbour [11]: lazy-based learning method that given a new case finds the closest group of $k$ cases in the training set, classifying the new case based on the predominance of a particular class in this neighborhood of size $k$.
3. Support Vector Machines, SVM [12]. This algorithm builds linear functions, $f(x)$, or hyperplanes, which separate the different classes of the training data set; thus, classification of a new case is made by testing the sign of this function.
4. Artificial Neural Networks. RBF Network [13]: typically is a three-layer feedback network composed by one unique hidden layer with radial basis functions as activation functions and the input and output layers.
5. Logic based algorithms. C4.5 algorithm [14] is an extension of ID3 algorithm [15] that builds a classification tree based on gain ratio criteria, defined as $I(X_i, C)/H(X_i)$. The algorithm also includes a pruning of induced tree, based on a test of hypothesis.
6. Rule induction methods. RIPPER algorithm [16] (Repeated Incremental Pruning Produce Error Reduction), is a rule induction algorithm that is an extension of

IREP algorithm [17] (Incremental Reduced Error Pruning) and that consists of learning rules obtained by performing a process of repeated growing and pruning.

### 3.4    Validation Method

The validation method used in present work is the 10-fold cross validation [18]. This method consists in dividing the data set $D = \left\{ (x_1^1,...,x_n^1,c^1),...,(x_1^N,...,x_n^N,c^N) \right\}$ in $f$ disjoint subsets of approximately the same size $S_1,...,S_i,...,S_f$. For each subset $S_i$ the following procedure is repeated: $S_i$ is considered the test sample, so the training sample will be composed of the $f-1$ other samples $S_1,...,S_{i-1},S_{i+1},...,S_f$, and is used to build the decision function $d$; classes from $S_i$ experiments are predicted by using $d$. This procedure is repeated for $i = 1,...,f$.

In our case, in order to validate the algorithms tested, a value of $f = 10$ is established; though there exists other values commonly used, i.e. $f =$ n. When this occurs, the method is named *leave-one-out cross validation*. Compared with other error estimation methodologies, this method provides an estimator with little bias but a lot of variance.

### 3.5    Multiclassifier Approach

Having a data set $D = \left\{ (x_1^1,...,x_n^1,c^1),...,(x_1^N,...,x_n^N,c^N) \right\}$, for each classifier mentioned in Section 3.3 a classification model is obtained and the label distribution for each test case $x = (x_1^{'},...,x_n^{'})$ is computed. Given these inputs, a multiclassifier approach is designed given three different kinds of experiments:

- First experiment: confidence voting [19]. The multiclassifier will predict the label, $y'$, of each test case, $x$, based on the sum of the label distributions obtained by each classifier:

$$\forall_x y'= \sum_{c_i \in C(i=1,...,6)} \left( dist(c_i)*I(v = y_i) \right) \ . \tag{1}$$

  In (1) $v$ is a class label, $y_i$ is the class label for the $i$th classifier tested $c_i \in C(i = 1,...,6)$, $dist(c_i)$ is the distribution that $c_i$ obtained for class label $v$ and I $(\cdot)$ is an indicator function to modify the sign of the confidence that returns the value 1 if its argument is true and -1 otherwise; a positive sign implies a correct prediction, whereas a negative sign implies a wrong prediction.
- Second experiment: weighted voting [20]. Given the class distribution, a weighting of the classifiers based on its accuracy is performed. Prediction made by the multiclassifier will be based on the sum of such weights.

$$weight(c_i,c_j) = \left\{ \begin{array}{ll} 1 & if \ dist(c_i) > dist(c_j) \\ -1 & if \ dist(c_i) < dist(c_j) \end{array} \right\} * I(v = y_i) \ ; \tag{2}$$

$$\bigvee_{x} y' = \sum_{\substack{i,j=1,\dots,6}}^{i \neq j} weight(c_i, c_j) \; . \tag{3}$$

In (2) $v$ is a class label, $y_i$ is the class label for the $i$th classifier tested $c_i \in C (i = 1,\dots,6)$ and I $(\cdot)$ is an indicator function that returns the value 1 if its argument is true and -1 otherwise; a positive sign implies a correct prediction, whereas a negative sign implies a wrong prediction.

- Third experiment: majority voting [21]. Once the list of labels for each test case is obtained by each classifier, the multiclassifier will predict the label, $y'$, of each test case, $x$, based on the majority label; the tie case is solved randomly, providing an arbitrary solution.

$$\bigvee_{x} y' = \arg\max_{v} \sum_{x_i, y_i \in D_i} I(v = y_i) \; . \tag{4}$$

In (4) $v$ is a class label, $y_i$ is the class label for the $i$th classifier tested $c_i \in C (i = 1,\dots,6)$, and I $(\cdot)$ is an indicator function that returns the value 1 if its argument is true and 0 otherwise.

The idea that is beyond these experiments is based on *Voting* concept, which is one of the simplest procedures when combining different classifier outputs in a vote-based framework: having $C_1, \dots, C_N$ the set of classification models induced by a total of $N$ different learning algorithms $L_1, \dots, L_N$ and a database $D$ with characteristic vectors, in order to classify a new instance the classifiers $C_1, \dots, C_N$ will be asked to get the class value they predicted.

Into present work voting criteria is enhanced by computing distributions of class probabilities and thus obtaining a vector of degrees of confidence for all considered class labels. Those degrees of confidence will be the input of the multiclassifier approach, which will estimate the class value considering the three different experiments mentioned above.

## 4 Experimental Results

Experiments accomplished in this study envisage a total of 313 drilling processes using two different workpieces made of the same material, aluminium 7075. They are distributed as follows: 115 drilling processes related to first workpiece and 198 drilling processes related to second workpiece. All holes have been drilled using a hard metal drill of 10 millimetres of diameter. Machining parameters considered have been the following: 200 m/min. of cutting speed and 0.3 millimetres of advance/revolution.

The same training data and the same features are used to obtain each classification model, applying the 10-fold cross-validation method. Results obtained from this analysis are illustrated in table format, containing the accuracy of each isolated classifier and the multiclassifier approach, regarding the three different distribution-based experiments performed. The meaning of each numeric value in the tables is as follows:

- *Correctly Classified Instances*: number of correctly classified instances from data.
- *Incorrectly Classified Instances*: number of incorrectly classified instances from data.
- *Precision:* estimation of the accuracy of the classifier provided that drill wear has been predicted.

$$precision = tp\,/(tp + fp)\ .\tag{5}$$

In (5) *tp* is the rate of true positives predictions and *fp* is the rate of false positives predictions.

- *Recall:* estimation of the ability of a classifier to select instances that are related to drill wear from the data set.

$$recall = tp\,/(tp + fn)\ .\tag{6}$$

In (6) *tp* is the rate of true positives predictions and *fn* is the rate of false negative predictions.

- *F-Measure:* the weighted harmonic mean of precision and recall.

$$F - Measure = \frac{2 \cdot precision \cdot recall}{(precision + recall)}\ .\tag{7}$$

### 4.1    Burr detection

A total of 104 holes have been analysed, after performing a simple sample selection process consisted in filtering samples that contained incorrectly measured signals. Discretization of continuous burr values measured has been established as follows, according to the values obtained and under the supervision of drilling process experts:

- If burr level was under 150 microns, discrete value has been set to "low"
- If burr level was between 150 and 300 microns, discrete value has been set to "medium"
- If burr level was over 300 microns, discrete value has been set to "high"

In the following tables results obtained by each isolated classifier tested are shown, based on the criteria mentioned above (see Table 3 and Table 4).

**Table 3.** Results obtained from Naïve Bayes, *k*-NN and SVM.

|                                    | Naïve Bayes  | k-NN       | SVM        |
|------------------------------------|--------------|------------|------------|
| *Correctly Classified Instances*   | 71 (68.3 %)  | 77 (74 %)  | 78 (75 %)  |
| *Incorrectly Classified Instances* | 33 (31.7 %)  | 27 (26 %)  | 26 (25 %)  |
| *Precision*                        | 0.28         | 0.31       | 0.18       |
| *Recall*                           | 0.54         | 0.36       | 0.18       |
| *F-Measure*                        | 0.37         | 0.32       | 0.18       |

**Table 4.** Results obtained from RBF Network, C4.5 and RIPPER.

|  | *RBF Network* | *C4.5* | *RIPPER* |
|---|---|---|---|
| *Correctly Classified Instances* | 85 (81.6 %) | 82 (78.9 %) | 84 (80.8 %) |
| *Incorrectly Classified Instances* | 19 (18.4 %) | 22 (21.1 %) | 20 (19.2 %) |
| *Precision* | 0.66 | 0.33 | 0.43 |
| *Recall* | 0.36 | 0.27 | 0.27 |
| *F-Measure* | 0.46 | 0.3 | 0.32 |

*RBF Network* obtained the highest average correct classification rate: 81.6%. In the case of the multiclassifier approach, as can be seen in Table 5, the burr prediction accuracy is, in general, slightly improved. *Naïve Bayes* was the classifier that presented the lowest percentage of goodness, 68.3%, much lower than percentages obtained by the three voting experiments.

**Table 5.** Final results obtained by applying the multiclassifier approach.

|  | *Confidence voting method* | *Weighted voting method* | *Majority voting method* |
|---|---|---|---|
| *Correctly Classified Instances* | 85 (81.6 %) | 86 (82.7 %) | 87 (83.6 %) |
| *Incorrectly Classified Instances* | 19 (18.4 %) | 18 (17.3 %) | 17 (16.4 %) |
| *Precision* | 0.33 | 0.4 | 0.44 |
| *Recall* | 0.27 | 0.36 | 0.36 |
| *F-Measure* | 0.3 | 0.38 | 0.4 |

Regarding the experiment related to confidence voting method, it is important to point out the information gain that implies working with class label distributions. This is really interesting from the point of view of an expert on the application field, providing an estimation of reliability for each prediction made.

## 4.2    Roughness detection

Regarding roughness detection problem, 66 holes were analysed after performing a sample selection process as it was made in the case of burr detection problem. Resulting continuous roughness values were discretized as follows, according to the values obtained and under the supervision of drilling process experts:

- If roughness values were under 2 microns, discrete value was set to "normal"
- If roughness values were equal or over 2 microns, discrete value was set to "high"

Results obtained from roughness analysis with regards to each isolated classification model are shown in the following tables taking into account previously defined criteria (see Table 6 and Table 7).

**Table 6.** Results obtained from Naïve Bayes, *k*-NN and SVM.

|  | *Naïve Bayes* | *k-NN* | *SVM* |
|---|---|---|---|
| *Correctly Classified Instances* | 47 (71.2 %) | 45 (68.2 %) | 41 (62 %) |

| | | | |
|---|---|---|---|
| *Incorrectly Classified Instances* | 19 (28.8 %) | 21 (31.8 %) | 25 (38 %) |
| *Precision* | 0.58 | 0.54 | 0.68 |
| *Recall* | 0.56 | 0.61 | 0.38 |
| *F-Measure* | 0.57 | 0.56 | 0.49 |

**Table 7.** Results obtained from RBF Network, C4.5 and RIPPER.

| | *RBF Network* | *C4.5* | *RIPPER* |
|---|---|---|---|
| *Correctly Classified Instances* | 42 (63.6 %) | 46 (69.7 %) | 43 (65.2 %) |
| *Incorrectly Classified Instances* | 24 (36.4 %) | 20 (30.3 %) | 23 (34.8 %) |
| *Precision* | 0.46 | 0.63 | 0.5 |
| *Recall* | 0.35 | 0.3 | 0.3 |
| *F-Measure* | 0.40 | 0.41 | 0.37 |

As can be seen in tables showed above, the isolated classifier that best predicted the roughness levels was *Naïve Bayes*, obtaining 71.2% of well classified samples. This value coincides with performance obtained by the multiclassifier approach when using weighted voting method (see Table 8). Performance obtained from confidence and majority voting methods was slightly higher, 74.1%. The worst percentage of correctly classified instances was obtained by *SVM* algorithm: 62%.

**Table 8.** Final results obtained by applying the multiclassifier approach.

| | *Confidence voting method* | *Weighted voting method* | *Majority voting method* |
|---|---|---|---|
| *Correctly Classified Instances* | 49 (74.1 %) | 47 (71.2 %) | 49 (74.1 %) |
| *Incorrectly Classified Instances* | 17 (25.9 %) | 19 (28.8 %) | 17 (25.9 %) |
| *Precision* | 0.69 | 0.63 | 0.69 |
| *Recall* | 0.48 | 0.38 | 0.48 |
| *F-Measure* | 0.56 | 0.46 | 0.56 |

## 5   Conclusions

From experimental results presented in this paper, it can be concluded that multiclassifer approach predictions are, in general, more accurate than predictions made by isolated classifiers. However, it is important to stress that in some cases an individual classifier is able to guarantee a good performance; therefore, a further similarity analysis among classifiers should be accomplished in order to study this interesting behaviour. A balance between the expected accuracy of the classifier model and the time and resources needed should be also established in advance, in order to estimate the usefulness of having a more accurate combination of several classification models or just one less accurate but faster and lighter model. This is strongly recommended when data sets are composed of a great number of instances or when the quality of data is poor.

Regarding the drill wear prediction problem, the following conclusions are drawn:

- Drill wear prediction can be characterized by the presence of burr and roughness on holes drilled with the support of data mining techniques
- When detecting burr, the multiclassifier approach experiment related to majority voting obtained the highest average correct classification rate: 83.6 %
- When detecting roughness, the multiclassifier approach experiments related to confidence values and majority vote of class labels obtained the best accuracy: 74.1 %
- In general, experimental results demonstrate that multiclassifier approach is able to slightly reduce the number of false positives obtained by most of the classifiers individually, which can lead to an incorrect drill wear substitution strategy
- By means of classification algorithms it is possible to avoid poor drilling quality due to drill wear in an efficient and automatic way

Future works in relation to this study will be oriented not only to predict burr and roughness more accurately, but also to infer relations between both parameters and how they make influence on drill wear over time, by considering different predictor variables and multiclass models. Different classification models can also be learnt for different feature recognition tasks in order to improve the accuracy of the multiclassifier, which will combine predictions made for each feature separately. The study must be also extended to different application sectors and domains.

## 6    References

1. C. Emmanouilidis, et al.: Flexible software for condition monitoring, incorporating novelty detection and diagnostics, Computers in Industry, vol. 57, pp. 516-527, 2006
2. Cassady, C.R., K. Schneider and P. Yu: Impact of Maintenance Resource Limitations on Manufacturing System Productivity. Proceedings of the Industrial Engineering Research 2002 Conference
3. A. Muller, et al.: Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system, Reliability Engineering & System Safety, vol. 93, pp. 234-253, 2008
4. M. van Erp, L. Vuurpijl, L. Schomaker: An overview and comparison of voting methods for pattern recognition. In Proc. of the 8th IWFHR, pp. 195-200, 2002
5. Ferreiro, S., Arana, R., Aizpurua, G., Aramendi, G., Arnaiz, A., and Sierra, B.: Data Mining for Burr Detection (in the Drilling Process). Proceedings of the 10th international Work-Conference on Artificial Neural Network: Salamanca, Spain, June 10 - 12, 2009
6. Jolliffe I.T. Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., Springer, NY, XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4, 2002
7. Carrascal, A., Diez, A., Azpeitia, A.: Unsupervised Methods for Anomalies Detection through Intelligent Monitoring Systems, HAIS09 (Hybrid Artificial Intelligence Systems), pp. 137-144, Salamanca, 2009
8. S.B. Kotsiantis: Supervised Machine Learning. A Review of Classification Techniques, Informatics 31, 249-268, 2007
9. Xindong Wu, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey McLachlan, Angus Ng, Bing Liu, Philip Yu, Zhi-Hua Zhou, Michael Steinbach, David Hand, and Dan Steinberg: Top 10 algorithms in data mining. Knowledge and Information Systems, 14(1): 1-37, January 2008
10. Rish, Irina: An empirical study of the naive Bayes classifier. IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, 2001

11. D. Aha, D. Kibler: Instance-based learning algorithms. Machine Learning. 6: 37-66, 1991

12. Cristianini, N. and Shawe-Taylor, J.: An Introduction to Support Vector Machines. Cambridge University Press, 2000

13. Powell, M. J. D.: Radial basis functions for multivariable interpolation: a review. In Algortithms for Approximation on Functions and Data, edited by J. C. Mason and M. G. Cox, 143-167. Oxford University Press, Oxford, 1987

14. Quinlan: C4.5. Programs for Machine Learning. Morgan Kaufmann, 1993

15. Quinlan J. R.: Induction of decision trees. Machine Learning, 1(1): 81-106, 1986

16. W.W. Cohen: Fast effective rule induction. Proceedings of the Twelfth Internacional Conference on Machine Learning, 1995

17. J. Fürnkranz, G. Widner: Incremental Reduced Error Pruning. Proceedings of the Eleventh Internacional Conference on Machine Learning, 1994

18. Stone M.: Cross-validatory choice and assessment of statistical predictions (with discussion). Journal of the Royal Statistical Society, Series B 36: 111-147, 1974

19. Kittler, J., Hated, M., Duin, R. P. W. & Matas, J.: On Combining Classifiers. IEEE Transactions PAMI, 20, 226-239, 1998

20. Verma, B., P. Gader and W. Chen: Fusion of multiple handwritten word recognition techniques. Patt. Recog. Lett., 22: 991-998, 2001

21. Dietterich, T. G.: Machine learning research: Four current directions. AI Magazine, 18(4): 97-136, 1997