

Title

xxx¹, yy¹, and zzz¹

¹Faculty of Mathematics and Information Science, Warsaw University of Technology
ul. Koszykowa 75, 00-662 Warsaw, Poland
{xxx,yy,zzz}@mini.pw.edu.pl

Abstract. In the article we present

1 Introduction

2 Literature Review

3 Preliminaries

3.1 Ellipsoids for Foreign Elements Rejection

In computational geometry, the smallest enclosing box problem is that of finding the oriented minimum bounding box enclosing a set of points. There are many different approaches towards solving it regardless of space dimension. These include hypercubes, diamonds, balls or ellipsoids. Comparing highlights and drawbacks of each method such as: computational complexity, ease of testing point inclusion and implementation, ellipsoids method was chosen.

MVEE

3.2 Native Elements Classification

Support Vector Machines

Random Forests

K-Nearest Neighbors

3.3 Quality Evaluation

- CC (Correctly Classified) - the number of correctly classified patterns, i.e. native patterns classified as native ones with the correct class,
- TP (True Positives) - the number of native patterns classified as native (no matter, into which native class),

- FN (False Negatives) - the number of native patterns incorrectly classified as foreign,
- FP (False Positives) - the number of foreign patterns incorrectly classified as native,
- TN (True Negatives) - the number of foreign patterns correctly classified as foreign.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

$$\text{Strict Accuracy} = \frac{CC + TN}{TP + FN + FP + TN}$$

$$\text{Native Precision} = \frac{TP}{TP + FP}$$

$$\text{Native Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Strict Native Sensitivity} = \frac{CC}{TP + FN}$$

$$\text{Fine Accuracy} = \frac{CC}{TP}$$

$$\text{Foreign Precision} = \frac{TN}{TN + FN}$$

$$\text{Foreign Sensitivity} = \frac{TN}{TN + FP}$$

$$\text{F-measure} = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

- *Strict Accuracy* is the absolute measure of the classifier's performance. It is the ratio of the number of all *correctly* classified patterns, i.e. native patterns classified to their respective classes and rejected foreign ones to the number of all patterns being processed.
- *Accuracy* is a characteristic derived from strict accuracy by ignoring the need to classify native patterns to their respective classes; in other words, it is sufficient to correctly identify whether a pattern is native or foreign one. This measure describes the ability to distinguish between native and foreign patterns.
- *Native Precision* is the ratio of the number of not rejected native patterns to the number of all not rejected patterns (i.e. all not rejected native and foreign ones). Native Precision evaluates the ability of the classifier to distinguish

00112233445566778899
00112233445566778899

Fig. 1. ...

- native patterns from foreign ones. The higher the value of this measure, the better ability to distinguish foreign elements from native ones. Native Precision does not evaluate how effective identification of native elements is.
- *Native Sensitivity* is the ratio of the number of not rejected native patterns to all native ones. This measure evaluates the ability of the classifier to identify native elements. The higher the value of Native Sensitivity, the more effective identification of native elements. Unlike the Native Precision, this measure does not evaluate the effectiveness of separation between native and foreign elements.
 - *Strict Native Sensitivity* takes only correctly classified native patterns and does not consider native patterns, which are not rejected and assigned to incorrect classes, unlike *Native Sensitivity*, where all not rejected native patterns are taken into account.
 - *Fine Accuracy* is the ratio of the number of native patterns classified to correct classes, i.e. assigned to their respective classes, to the number of all native patterns not rejected. This measure conveys how precise is correct classification of not rejected patterns.
 - *Foreign Precision* corresponds to Native Precision.
 - *Foreign Sensitivity* corresponds to Native Sensitivity.
 - Precision and Sensitivity are complementary and there exists yet another characteristic that combines them: the *F-measure*. It is there to express the balance between precision and sensitivity since these two measures affect each other. Increasing sensitivity can cause a drop in precision since, along with correctly classified elements, there might be more incorrectly classified,

4 Experiments

4.1 Presentation of Datasets

Figure 1 presents native and foreign patterns ...

4.2 Impact on Classification

4.3 Rejection Quality

5 Conclusion

Proposed ...

In future ...

Table 1. Results for classification with rejection on train and test sets of native patterns in comparison with classification results without rejection mechanism. RF - results for random forest, SVM - results for Support Vector Machines,

	no rejection			with rejection		
Basic Classifier	RF	SVM	KNN	RF	SVM	KNN
Data Set	Native Patterns, Train Set					
Fine Accuracy						
Strict Native Sensitivity						
Native Sensitivity						
Data Set	Native Patterns, Test Set					
Fine Accuracy						
Strict Native Sensitivity						
Native Sensitivity						
		—	—		—	—

Table 2. Results of classification with rejection on the set of native patterns supplemented with different sets of semi-synthetic foreign patterns....

Basic Classifier	RF SVM KNN	RF SVM KNN
Data Set	xxx	x x
Strict Accuracy		
Accuracy		
Native Precision		
Native Sensitivity		
Foreign Precision		
Foreign Sensitivity		
Native F-measure		
Foreign F-measure		
Data Set	yyy	zzz
Strict Accuracy		
Accuracy		
Native Precision		
Native Sensitivity		
Foreign Precision		
Foreign Sensitivity		
Native F-measure		
Foreign F-measure		

Acknowledgment

The research is partially supported by the

References

1. Hempstalk, K., Frank, E., Witten, I., *One-class classification by combining density and class probability estimation*, Machine Learning and Knowl. Disc. in Databases, pp. 505-519, 2008.