Title

xxx[1], yyy[1], and zzz[1]

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

Abstract. In the article we present

1 Introduction

AJ

2 Literature Review

AJ

3 Preliminaries

3.1 The Task of Classification with Rejection

AJ

3.2 Regression

AZ, PW Regression analysis is a statistical process for estimating the relationships between variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. The estimation target is a function of the independent variables called the regression function. There are many different regression models: linear regression, polynomial regression, logistic regression, Bayesian Ridge regression, etc. In this paper we will focus mainly on two regression variants: logistic regression and polynomial one (please note that linear regression is in fact polynomial regression using polynomial of degree 2).

Logistic regression relies on the assumption that the independent variables are dichotomous (it means that their values are either 0 or 1), in other words they describe presence or absence of certain facts. In this model, the probabilities describing the possible outcomes of a single trial are modelled using a logistic function. The received answer is evaluated using maximum likelihood method. Predicted values are probabilities and are therefore restricted to (0,1) through the logistic distribution function.

Polynomial regression can be viewed as a model that uses linear regression extended by constructing polynomial features from the coefficients. This approach maintains the generally fast performance of linear methods, while allowing them to fit a much wider range of data. The predictors resulting from the polynomial expansion of the "baseline" predictors are known as interaction features, and can be used in classification problem. Just like linear regression, polynomial regression models are usually fit using the method of least squares.

3.3 Rejection Mechanism Based on Regression

AZ, PW Calculations presented in this paper were conducted on a mixed set consisting of letters and digits. Because the aim of our work was to provide classifier-based solution for classifying digits and rejecting other symbols, we decided to compare three different approaches towards this problem. All of them based on the assumption that there is no information regarding outliers during classifier training.

As the first way to deal with presented issue "one-versus-all" method was prepared. This approach requires creating a vector of classifiers constructed in a specific way. Each classifier has to be trained on a specially prepared training data, consisting of two sets: the first one (denoted as class_1) holding all training data entries for certain native class, and the second one (denoted as class_2) being the result of a subset sum operation performed on the rest of the classes except for the class used in class_1 set. Please note that both class_1 and class_2 sets should have the same size, so it is advisable to randomly choose elements when creating class_2. One classifier for each native class has to be present in a final vector. The actual classification with rejection is performed by presenting the unknown pattern to each of the classifiers from the vector. When the classifier recognizes element as a native one (belonging to class_1) then the pattern is treated as a recognized, and classified native element. In case of all classifiers rejecting such pattern (classifying it as element from class_2), it is treated as outlier and rejected. It is also worth noting that there is a possibility of more than one classifier recognizing the pattern as a native element. In such case randomly chosen class is assigned to this pattern.

The second approach uses "one-versus-one" method. Similarly to the previous one it requires preparing vector of classifiers, but this time it consists of $\binom{n}{2}$ classifiers, where n is the number of unique native classes. Each classifier is trained on data consisting of two sets: the first one (denoted as class_1) holding all training data entries for certain native class, and the second one (denoted as class_2) holding all training data entries for some other class (not the same as class_1). In the end there is one classifier for each pair of classes: 1 vs. 2, 1 vs. 3, $\ldots$, 1 vs. n, $\ldots$, (n-1) vs. n. Classification with rejection mechanism is based on presenting unknown pattern to every classifier in the vector and remembering their answers (e.g. classifier constructed for 1 vs. n classes can classify pattern as belonging to class 1 or class n). In the end those answers can be presented as a n-wide array with each element being the number of times pattern was classified as belonging to certain class. The pattern is rejected when difference between two biggest values in the result array is smaller than two. In such case it is assumed that classifiers were uncertain as to which class should this unknown element belong to. Otherwise the pattern is classified as an element from the class which had the biggest value in the result array. tutaj mona doda schemat prezentujcy opisan koncepcj

The last prepared and examined method, presented in this work, bases on the way of constructing classifiers vector used in the second approach ("one-versus-one"). The difference between those two methods lies in a rejection mechanism. In this method an unknown pattern is treated as a foreign element if its biggest value in the result array is lesser than (n-1). What it actually means is that there must be a certain class that has always been chosen by a classifier from the vector. tutaj mona doda schemat prezentujcy opisan koncepcj

3.4 Quality Evaluation

AJ

- 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.

4 Experiments

4.1 Presentation of Datasets

AJ Figure 1 presents native and foreign patterns ...

4.2 Impact on Classification

AZ, PW

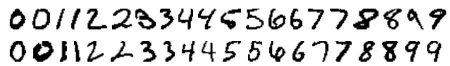

Fig. 1. ...

Table 1. Results for classification without rejection on train and test sets of native patterns in comparison with classification results with rejection mechanism.

	no rejection	with rejection

4.3 Rejection Quality

AZ, PW

5 Conclusion

AJ Proposed ...
In future ...

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.