Classification with rejection concepts and evaluations

Conference Paper · November 2013

Marcin Luckner
Warsaw University of Technology

Classification with rejection: concepts and evaluations

Wladyslaw Homenda[1], Marcin Luckner[1] and Witold Pedrycz[2,3]

[1]Faculty of Mathematics and Information Science
Warsaw University of Technology
Plac Politechniki 1, 00-660 Warsaw, Poland
and
[2]System Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
and
[3]Department of Electrical and Computer Engineering, University of Alberta
Edmonton, Alberta, Canada T6G 2V4

Abstract. Standard classification process allocates all processed elements to given classes. Such type of classification assumes that there is are only native and no foreign elements, i.e. all processed elements are included in given classes. The quality of standard classification can be measured by two factors: numbers of correctly and incorrectly classified elements, called True Positives and False Positives. Admitting foreign elements in standard classification process increases False Positives and, in this way, deteriorates quality of classification. In this context, it is desired to reject foreign elements, i.e. to not assign them to any of given classes. Rejecting foreign elements will reduce the number of False Positives, but can also reject native elements reducing True Positives as side effect. Therefore, it is important to build well designed rejection, which will reject significant part of foreigners and only few natives. In this paper, evaluations of classification with rejection concepts are presented. Three main models: a classification without rejection, a classification with rejection, and a classification with reclassification are presented. The concepts are illustrated by flexible ensembles of binary classifiers with evaluations of each model. The proposed models can be used, in particular, as classifiers working with noised data, where recognized input is not limited to elements of known classes.

Keywords: rejection rule, binary classifiers ensemble, reclassification

1 Introduction

Pattern recognition is one of the leading subjects in the field of computer science in its both theoretical and practical aspects. For decades, it has been a subject of intense, purely theoretical research inspired by practical needs. The results have been published in prestigious scientific journals. Example applications are: recognizing printed text, manuscripts, music notation, biometric features, voice, speaker, recorded music, medical signals, images, etc.

The aspect of rejection of foreign element in pattern recognition is shallowly researched and not considered in practical applications. By foreign elements, we understand elements not belonging to recognized classes. Dissemination of technologies using pattern recognition increases the importance of identifying foreign elements. For example: in recognition of printed texts, foreign elements (blots, grease or damaged symbols) appear in a negligible scale due to regular placement of printed texts' elements (letters, numbers, punctuation marks) and due to their good separability. These features of printed texts allow employing effective segmentation methods. However, in recognition of such sources as geodetic maps or music notation, problem of foreign elements is more important. Unlike printed text, such sources contain elements placed irregularly and overlapping native elements. Such elements are hardly distinguishable by size and shape analysis. Thus, strict rules of segmentation will result in rejection of many native symbols.

Due to weak separability of foreign and native elements of recognized sources, segmentation criteria must be more tolerant than in the case of printed texts in order not to reject native elements at the stage of segmentation. In consequence, more foreign elements are subjected to stages of recognition following segmentation and should be eliminated then. The problem of analysis of foreign elements is highly important in such domains as analysis of medical signals and images or recognition of geodesic maps or music notations (printed and handwritten) and its importance will be increasing in future.

The problem of rejecting foreign elements in recognition tasks is not present frequently in research and is rather rarely present in papers on pattern recognition. Assuming that classified elements are always native ones, i.e. they belong to one of recognized classes, and ignoring rejection of foreign elements is rather a standard attempt. Alike, papers describing practical applications of pattern recognition methods ignore the problem of foreign elements, what may come from insufficient theoretical research of this subject and limited abilities of existing rejection methods. There are significant exceptions, which show that the rejection problem cannot be disregarded, c.f. [12].

The motivation of this study arises from discussion on classification with rejection option. As outlined above, up-to-date research and practice still need conducting further studies on new aspects in the domain of pattern recognition. It is expected that research in this area will overcome technological barriers and will increase effectiveness in areas mentioned above.

The paper is structured as follows. Related research and introductory remarks are presented in section 2. In section 3, various ensembles of classifications based on binary classifiers are shown. The discussion includes evaluation criteria of rejection and reclassification methods. Conclusions and directions of further research are presented in section 4.

2 Preliminaries

The problem discussed above (formally defined in equation 2) is important in practice. In Figure 1, we present a fragment of a geodesic map raising classifi-

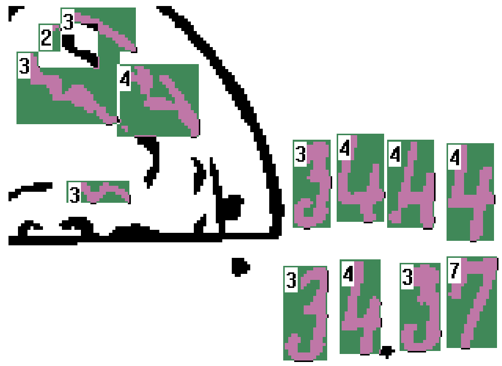

Fig. 1. Assumption that all patterns subjected to classification are digits raises classification errors for background elements unless they are rejected.

cation errors. Assumption that all patterns subjected to classification are digits raises classification errors for background elements unless they are rejected. This observation clearly shows a need for a rejection option in classification tasks. The rejection option can be described by a classification function with the rejection value in addition to computed classes. Such function may return 0 when an object should be rejected or a number of recognized class otherwise.

2.1 Related research

The theoretical foundations of rejection were created by Chow [1]. He created the optimal rejection rule that optimizes a classification error. The solution was limited to binary classifiers and presented as a method for Bayes classifiers. Recognized elements are rejected when distance to a discrimination plane is lower then a declared threshold. The optimal threshold is a compromise between a number of misclassified elements and rejected correct results.

The definition for a multi–class issue was presented by Ha [4]. The Chow's rule is calculated for all pairs of classes separated by a discrimination plane. There are also solutions for a linear multi–classification task [2]. However, most of theoretical works is limited to the binary case [10, 6].

The Chow's rule was redefined to practical applications as a rejection rule for support vector machines classifiers [7], where the probability is estimated as distance to discrimination plane.

The distance function can be also used to define a rejection rule based on dissimilarity between a class pattern and a recognized element [8].

Another practical trend is related to neural networks. The output of the neurons in feed–forward neural networks is determined by values of neurons in the output layer, where each neuron is equivalent to a recognized class. The class with a maximal value of the neurons' outputs is taken as the final classification result. This reject decision is issued when the value of the neuron is relative small [11]. In this approach, typical for fuzzy sets [3], when a new pattern cannot be classified as a single class with high certainty by a trained neural network, such a pattern is rejected in classification process, c.f. [5].

In natural way, a pattern can be rejected when two or more output neurons fire similar values. The same method can be used in all voting classifiers in a draw case [5]. Among ensembles of binary classifiers, the rejection option can be applied for ECOC (error-correcting output coding) classification systems [13].

The rejection problem has been raised in the recognition tasks of printed documents, namely recognition of music scores and geodetic maps. Rejection methods applied in such practical tasks, which in fact were created *ad hoc* or were implemented as dedicated methods of considered classifiers, increased effectiveness, and recognition rates, c.f. [9].

The above works outline highly important problem, which is solved neither in theory and research, nor in practice. Solutions based on the Chow's rule are limited to the cases of misclassifications between classes. The optimal solution assumes that rejected symbols increase the false negative factor of error for one of recognized classes. Unfortunately, noised and disrupted input patterns may be misclassified diminishing recognition efficiency.

On the other hand dedicated solutions force choice of classification methods. Moreover, there is a doubt, if methods based on rare instances can be applied in practice or if a low level neuron's response means false positive case.

2.2 Classification and rejection functions

Standard classification is described as a mapping Φ from the Cartesian product of features $\mathbb{X} = \mathrm{X}_1 \times \mathrm{X}_2 \times \ldots \times \mathrm{X}_\mathrm{m}$ into the set of classes $\mathbb{C} = \{\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_\mathrm{n}\}$. For the sake of simplicity we assume that the function Φ will take indexes of classes $i = 1, \ldots, n$ as its values instead of classes themselves. Since classification is usually parameterized, we will denote classification parameters by α. Therefore, a classification function will be finally denoted as

$$\Phi_\alpha : \mathbb{X} \to \{1, 2, \ldots, n\}, \quad \Phi_\alpha(\overline{x}) = i \tag{1}$$

for given $\overline{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{X}$ and $i \in \{1, 2, \ldots, n\}$.

Therefore, classification is carried out using a classification function $\phi_\alpha(\overline{x}) = i$, which calculates an index $i \in \{1, \ldots, n\}$ of recognized class C_i based on classification parameters α.

This definition of classification always raises misclassification for foreign elements, when a classified element does not belong to any of recognized classes. Let us denote the set of feature vectors of such elements by $\mathbb{S}$. Then, for any $\overline{x} \in \mathbb{S}$ the classification function Φ_α raises error:

$$(\forall \overline{x} \in \mathbb{S})(\forall i = 1, 2, \ldots, n)\ (\Phi_\alpha(\overline{x}) = i) \Rightarrow (x \notin C_i). \tag{2}$$

Therefore, we use the following classification function to avoid misclassification of objects not belonging to any class:

$$R_{\alpha,\beta} : \mathbb{X} \to \{1, 2, \ldots, n\} \cup \{0\} \tag{3}$$

where β is a possible extra parameter affecting rejection and independent on α.

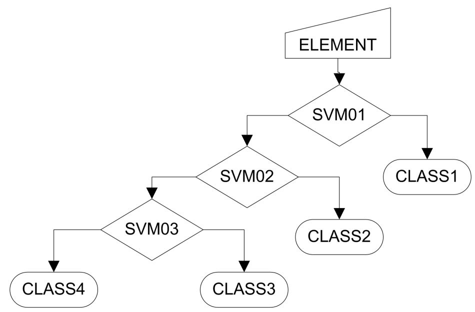

Fig. 2. Standard classification with the space of objects split into recognized classes.

In some circumstances, the learning set must be considered in classification process. In such a case, a classification function is denoted as follows:

$$\Phi_\alpha : \mathbb{X} \times \mathbb{L} \rightarrow \{1, 2, \ldots, n\}, \quad \Phi_\alpha(\overline{x}, L) = i \tag{4}$$

where: $\mathbb{L}$ is a space of possible learning sets, L is a given learning set, and the rest of notation is compatible with formula 1.

Alike, when learning set must be considered, we denote a classification function with rejection as follows:

$$R_{\alpha,\beta} : \mathbb{X} \times \mathbb{L} \rightarrow \{1, 2, \ldots, n\} \cup \{0\}$$

where: notation is compatible with formulas **??** and 4.

3 Classification with rejection

In this section, we discuss different architectures of classification with rejection based on binary classifiers, e.g. on support vector machines. In Figure 2, standard classification architecture for four classes is presented. Notice, that architecture of classification is a kind of an (unbalanced) binary tree with $n - 1$ classifiers being internal nodes and n classes of elements being leaves. This architecture can easily be extended to more than four classes. Of course, we may have different configurations of classification architecture for higher number of classes. For instance, we may have "cascade" configuration in a form of highly unbalanced three with one path of internal nodes, i.e. one path of classifiers. Other configurations may be based on different binary threes with the same numbers of internal nodes (classifiers).

3.1 Architecture of ensembles of binary classifiers with rejection

In Figure 3, a cascade configuration of classification with rejection is presented. Foreign elements are assumed to create additional class, which is classified at the beginning of the paths of binary classifiers. Then next classes are identified "step by step". Notice, that this configuration may be slightly reorganized: elements of recognized classes may be identified first and then rejected elements may be identified at the end of the path of classifiers.

A balanced tree configuration of classification with rejection outlined in Figure 5 allows for splitting the set of foreign elements for several subclasses. The number of subclasses of the class of foreign elements is equal to the number of recognized classes.

In general, in both configurations of Figures 3 and 5 rejection can be simply interpreted as standard classification with additional class(es) of foreign elements. Therefore, standard classification methods may be used in such kind of classification with rejection. Namely, supplementing the set of recognized classes with class(es) of foreign elements, we get standard classification. This method increases the number of classes by one or twice and this is the only one obstacle of processing foreign elements.

The method of distinguishing foreign elements may be used for optimization of classification without significant increment of complication. In Figure 6, a classification architecture with reclassification stage of foreign elements is presented. This configuration is based on classifiers used in basic stage of classification.

The configuration of classification with rejections employing new dedicated classifiers is shown in Figure 7. It is assumed that any class of foreign elements may include native elements of all classes besides the one corresponding to this foreign one. Therefore, every subclass of foreign elements is presented to reclassification to all but one recognized classes. It can be noticed that for this configuration, number of classifiers increases significantly.

3.2 Evaluation of rejection methods

No rejection. We assume that the set of recognized elements include classes $\mathbb{C} = \{\mathrm{C_1}, \mathrm{C_2}, \dots, \mathrm{C_n}\}$ of native elements to be recognized and the class C_0 of foreign elements to be rejected. Employing standard classification of formula 1, we get all elements classified to classes of $\mathbb{C}$. This classification is presented in Figure 2. Also all foreign elements of the class C_0 will be incorrectly assigned to the classes of $\mathbb{C}$. Therefore, we can evaluate two factors of such classification: $T(rue)P(ositives)$ and $F(alse)P(ositives)$, i.e. correctly and incorrectly classified elements. These factors are defined by the following formulas:

$$
\begin{aligned}
TP_{\Phi} &= \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \delta_{i \Phi_{\alpha}(\overline{x})} \\
FP_{\Phi} &= \sum_{i=1}^{n} \sum_{\overline{x} \notin C_i} \delta_{i \Phi_{\alpha}(\overline{x})}
\end{aligned}
\tag{5}
$$

The factor FP includes all incorrectly classified native elements of classes $\mathbb{C}$ and all foreign elements of the class C_0. As outlined in section 2.2, foreign elements of the class C_0 deteriorates classification rate. In order to improve classification, foreign elements should not be classified to any of the classes $\mathbb{C}$. Some attempts to such improvements are presented in next sections.

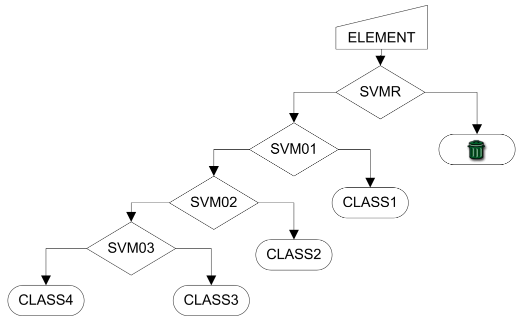

Fig. 3. Classification with rejection - cascade configuration with one class of rejected elements.

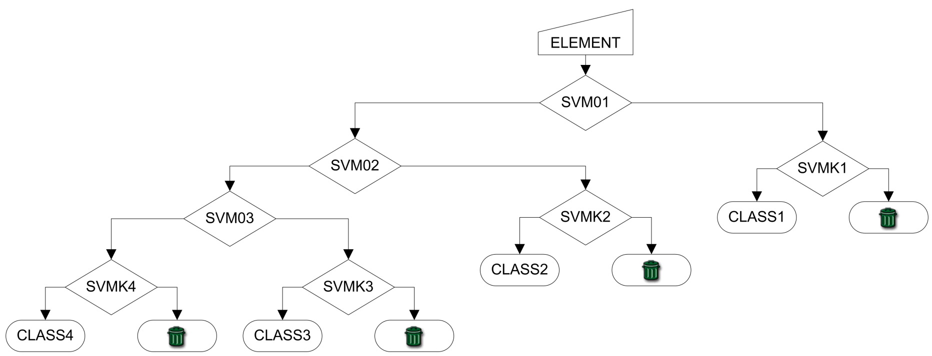

Fig. 4. Classification with rejection - cascade configuration with many classes of rejected elements.

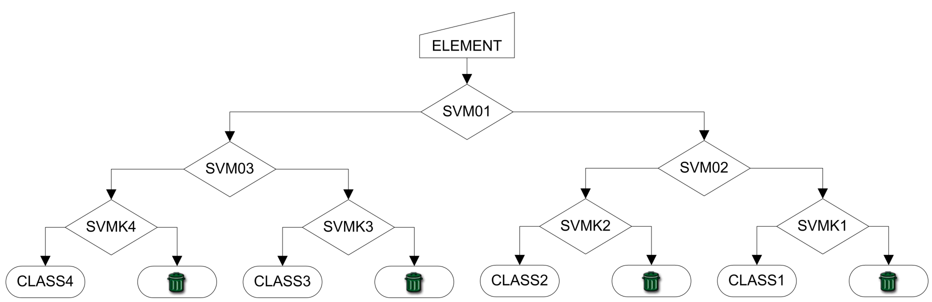

Fig. 5. Classification with rejection - balanced tree configuration with many classes of rejected elements.

Rejection. The first attempt to improve the classification presented in Figure 2 relies on rejecting foreign elements at the beginning of classification process. Configurations of the simple classification with rejecting foreign elements is shown in Figure 3 and then in Figures 4 and 5. Now, these three configurations can be characterized by four factors: TP, FP, TN and FN. These factors are computed in formulas 6. In this formulas sgn denotes the signum function, i.e. $\mathrm{sgn} : \mathbb{N} \to \{0, 1\}$, $\mathrm{sgn}(x) = 0 \Leftrightarrow x = 0$ and $\overline{\mathrm{sgn}} = 1 - \mathrm{sgn}$.

$$\begin{aligned}
TP_R &= \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \delta_{i\Phi_\alpha(\overline{x})} \cdot \mathrm{sgn}(R(\overline{x})) \\
FP_R &= \sum_{i=1}^{n} \sum_{\overline{x} \notin C_i} \delta_{i\Phi_\alpha(\overline{x})} \cdot \mathrm{sgn}(R(\overline{x})) \\
FN_R &= \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \delta_{i\Phi_\alpha(\overline{x})} \cdot \overline{\mathrm{sgn}}(R(\overline{x})) \\
TN_R &= \sum_{\overline{x} \in C_0} \overline{\mathrm{sgn}}(R(\overline{x}))\big)
\end{aligned} \tag{6}$$

Factors TP, FP are explained above. The factor TN counts all rejected foreign elements, while the factor FN counts all elements (native and foreign) incorrectly classified to classes of $\mathbb{C}$. Notice that these classifications' configurations decrease the factor FP by rejecting those foreign elements, which are incorrectly assigned to classes of $\mathbb{C}$ in the process of classification without rejection. Unfortunately, the factor TP may also be decreased by incorrect rejection of native elements. An answer to the question about possible benefits of rejection depends of quality of rejecting. If rejecting accuracy is high, then deterioration of correct classifications is smaller than improvements in incorrect classifications. In practice, rejection of foreign elements overheads incorrect rejection of native elements. Moreover, other two factors allow for better characterization of classification. In addition, it is worth to mention, that construction and usage of such classifiers do not create difficulties comparing to standard classification without rejection.

Rejection and reclassification in cascade configuration. Architectures of classification with rejection may allow for improving classification factors. In Figure 6, rejected elements are subjected for reclassification. This solution is bounded to architecture and configuration of classification. Reclassification solutions discussed in this paper are based on cascade configuration, c.f. Figure 6. Below, the configuration is analyzed in details.

In Figure 6, elements rejected at the stage of class C_1 are subjected again to classification at the stage of class C_2 and further classes. Then, elements rejected at the stage of class C_2 are subjected again to classification at the stage of class C_3 and further classes. And so on.

Assume that we have n classes, $n-1$ binary classifiers carrying the function Φ_α (i.e. SVM01, SVM02, ...) and n binary classifiers carrying the function $R_{\alpha,\beta}$ (i.e. SVMK1, SVMK2, ...). After the first reclassification, we have:

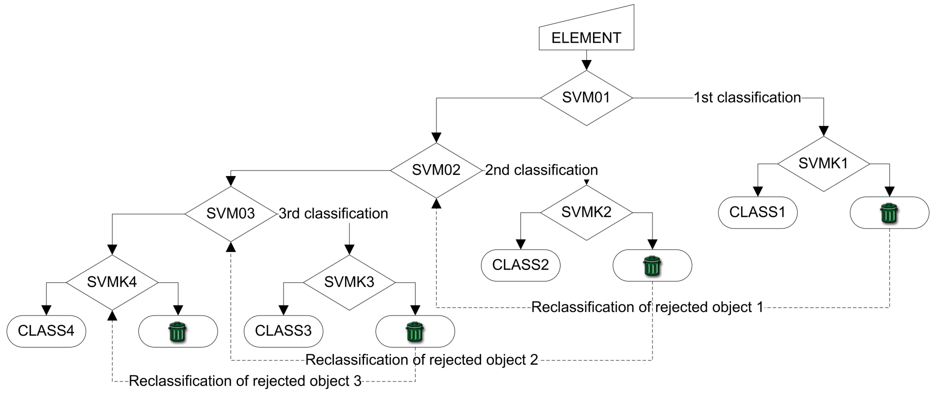

Fig. 6. Cascade configuration of classification with rejection. Reclassification with existing classifiers.

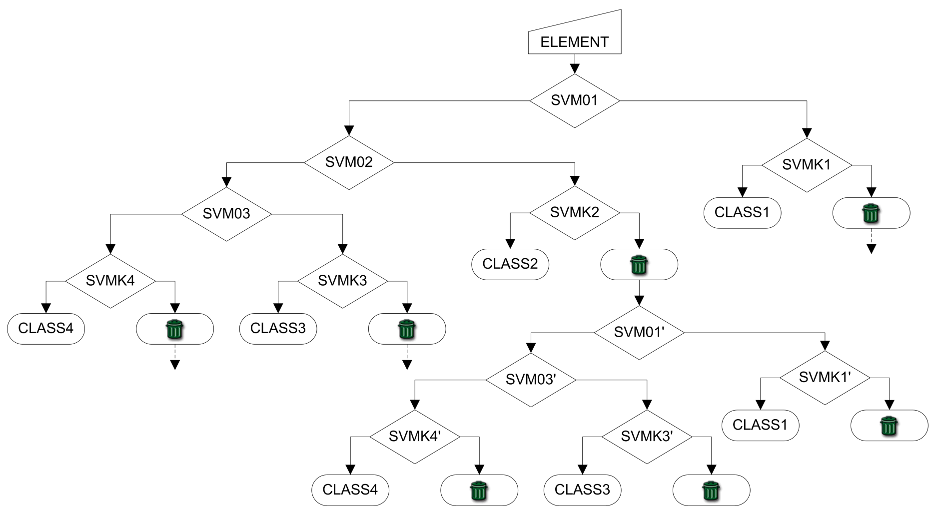

Fig. 7. Expanded cascade tree configuration of classification with rejection. Reclassification of rejected elements with new dedicated classifiers.

$$
\begin{aligned}
TP^1_{RR} = & \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x})) + \\
& \sum_{i=2}^{n} \sum_{\overline{x} \in C_i} \delta_{1\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R_1(\overline{x})) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x})) \\
& = TP_R + \sum_{i=2}^{n} \sum_{\overline{x} \in C_i} \delta_{1\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R_1(\overline{x})) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x}))
\end{aligned}
$$

and after the first and second reclassification we have:

$$\begin{aligned}
TP_{RR}^2 = & \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x})) + \\
& \sum_{i=2}^{n} \sum_{\overline{x} \in C_i} \delta_{1\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R^1(\overline{x})) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x})) + \\
& \sum_{i=3}^{n} \sum_{\overline{x} \in C_i} \delta_{1\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R^1(\overline{x})) \cdot \delta_{2\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R^2(\overline{x})) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x})) \\
= & \, TP_{RR}^1 + \\
& \sum_{i=3}^{n} \sum_{\overline{x} \in C_i} \delta_{1\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R^1(\overline{x})) \cdot \delta_{2\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}(R^2(\overline{x})) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}(R(\overline{x}))
\end{aligned}$$

where $R^1(\overline{x})$ and $R^2(\overline{x})$ are rejecting classifiers corresponding to classes C_1 and C_2, respectively. Below, $R^j(\overline{x})$ denotes the rejecting classifier corresponding to the class C_j.

Finally and by analogy:

$$\begin{aligned}
TP_{RR} &= TP_R + \sum_{k=2}^{n} \Bigg(\sum_{i=k}^{n} \sum_{\overline{x} \in C_i} \Big(\prod_{j=1}^{i-1} \delta_{j\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}\big(R^j(\overline{x})\big) \Big) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}\big(R(\overline{x})\big) \\
&= \sum_{k=1}^{n} \Bigg(\sum_{i=k}^{n} \sum_{\overline{x} \in C_i} \Big(\prod_{j=1}^{i-1} \delta_{j\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}\big(R^j(\overline{x})\big) \Big) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}\big(R(\overline{x})\big) \Bigg) \\
FP_{RR} &= FP_R + \sum_{k=2}^{n} \Bigg(\sum_{i=k}^{n} \sum_{\overline{x} \notin C_i} \Big(\prod_{j=1}^{i-1} \delta_{j\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}\big(R^j(\overline{x})\big) \Big) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}\big(R(\overline{x})\big) \Bigg) \\
&= \sum_{k=1}^{n} \Bigg(\sum_{i=k}^{n} \sum_{\overline{x} \notin C_i} \Big(\prod_{j=1}^{i-1} \delta_{j\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}\big(R^j(\overline{x})\big) \Big) \cdot \delta_{i\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}\big(R(\overline{x})\big) \Bigg) \\
FN_{RR} &= FN_R - \\
& \quad \sum_{i=1}^{n} \sum_{\overline{x} \in C_i} \sum_{k=i+1}^{n} \Big(\prod_{j=i}^{k-1} \delta_{j\Phi_\alpha(\overline{x})} \cdot \overline{\operatorname{sgn}}\big(R^j(\overline{x})\big) \Big) \cdot \delta_{k\Phi_\alpha(\overline{x})} \cdot \operatorname{sgn}\big(R(\overline{x})\big) \\
TN_{RR} &= TN_R - \sum_{\overline{x} \in C_0} \Bigg(\sum_{i=2}^{n} \Big(\prod_{j=1}^{i-1} \delta_{j\Phi_\alpha \overline{x}} \cdot \overline{\operatorname{sgn}}(R^j(\overline{x})) \Big) \delta_{i\Phi_\alpha \overline{x}} \cdot \operatorname{sgn}(R(\overline{x})) \Bigg)
\end{aligned} \tag{7}$$

The above formulas are applied to reclassification using the existing classifiers. There is no need for additional classifiers. However, it is possible and desired to increase quality of classification by applying ones. Such configuration is presented in Figure 7. Elements rejected at the first stage of classification are then presented to additional classifiers dedicated to given set of rejections. Structures of these additional classifiers are directly based on the main classification. They may be interpreted as main classifiers trained at dedicated learning sets. As it is outlined in Figure 7, new dedicated classifiers exclude one indicated class. These classifiers may be obtained from the main one by its training at indicated sets of rejected elements. Therefore, formulas 7 can be applied to describe factors TP, FP, FN and TN for this configuration. Adaptation of formulas 7 relies on considering dedicated learning sets, c.f. also formulas **??**. Namely, instances of $R(\overline{x})$ should be replaced with $R(\overline{x}, L)$, where L is respective learning set.

Intuitively, quality of classification depends on the rate of correctly classified elements. In this case, it depends on factors TP and TN, which should be maximized. On the other hand, minimization of factors FP and FN increases quality of classification. As it is seen in the above formulas, reclassification increases TP and decreases FN factors, what is desired. However, reclassification increases FP and decreases TN factors, what is undesirable. Therefore, quality of reclassification depends on significant overhead of desired factors over undesired ones. Fortunately, configurations presented in Figures 6 and 7 naturally lead to such overhead. Moreover, experiments confirm this advantage of classification with rejection. Since this paper is focused on concepts and algorithms, we do not expand the experimental part of this study.

4 Conclusions

In this paper, concepts and algorithms of classification with rejection of foreign elements are discussed. The discussion is based on three models of classification: classification without rejection, classifier with rejection and classification with reclassification of rejected elements. All models of classification use hierarchical ensembles of binary classifiers. The study if focused on evaluation of quality of different classification models using cascade configuration of classification.

Evaluation of classification models shows how rejection and reclassification affects quality of classification. Comparing to standard classification, i.e. classification without foreign elements, rejection of foreign elements lowers the number of False Positives, but may also decrease the number of T(rue)P(ositives). Reclassification can be used to reduce negative influence of rejections by reclassifying rejected elements. Such reclassification hunts up for F(alse)N(egatives), i.e. native elements incorrectly rejected, increasing quality of classification. Reclassification also may affect quality negatively. In the paper it is shown, how to evaluate and compare quality of different models of classification. It is indicated that positive influence of rejection and classification overheads deterioration.

Future research and practical directions would include studies on different architectures of classifications based on binary classifiers, discussion on many classes classifiers, reinforcement of basic classification with rejection to design reclassification models.

Acknowledgments

The research is supported by the National Science Center,
grant No 2012/07/B/ST6/01501, decision no UMO-2012/07/B/ST6/01501.

References

1. Chow, C.: On optimum recognition error and reject tradeoff. Information Theory, IEEE Transactions on 16(1), 41–46 (1970)
2. Fumera, G., Roli, F.: Analysis of error-reject trade-off in linearly combined multiple classifiers. Pattern Recognition 37, 1245–1265 (2004)
3. Graves, D., Pedrycz, W.: Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study (2010)
4. Ha, T.M.: Optimum tradeoff between class-selective rejection error and average number of classes. Engineering Applications of Artificial Intelligence 10(6), 525–529 (1997)
5. Ishibuchi H., Nakashima T., M.T.: Voting in fuzzy rule-based systems for pattern classification problems. Fuzzy Sets and Systems 103, 223–238 (1999)
6. Jigang Xie and Zhengding Qiu and Jie Wu: Bootstrap Methods for Reject Rules of Fisher LDA.Pattern Recognition, 425-428 (2006)
7. Li, M., Sethi, I.K.: Confidence-based classifier design. Pattern Recogn. 39(7), 1230–1240 (2006)
8. Lou, Z., Liu, K., Yang, J.Y., Suen, C.Y.: Rejection criteria and pairwise discrimination of handwritten numerals based on structural features. Pattern Anal. Appl. 2(3), 228–238 (1999)
9. Luckner, M.: Comparison of hierarchical svm structures in letters recognition task. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J. (eds.) Computational Intelligence: Methods and Applications, pp. 291–302. Challenging Problems of Science, Academic Publishing House EXIT, Warsaw (2008)
10. Mascarilla, L., Frélicot, C.: Reject strategies driven combination of pattern classifiers. Pattern Anal. Appl. 5(2), 234–243 (2002)
11. Pierce, S.G., Worden, K., Manson, G.: Evaluation of neural network performance and generalisation using thresholding functions. Neural Comput. Appl. 16, 109–124 (2007)
12. Pillai, I., Fumera, G., Roli, F.: A classification approach with a reject option for multi-label problems. In: Maino, G., Foresti, G.L. (eds.) ICIAP (1). Lecture Notes in Computer Science, vol. 6978, pp. 98–107. Springer (2011)
13. Simeone, P., Marrocco, C., Tortorella, F.: Design of reject rules for ecoc classification systems. Pattern Recognition 45(2), 863–875 (2012)