Composed Classifiers in Decision Processes

Výrostková Jana, Ocelíková Eva, Klimešová Dana*

Department of Cybernetics and Artificial Intelligence,
Technical University of Košice, Letná 9, 041 20, Košice, Slovak Republic
Jana.Vyrostkova@tuke.sk, Eva.Ocelikova@tuke.sk
* Czech University of Agriculture, Prague, Faculty of Economics and Management,
Dept. of Information Engineering, Kamycka 129, 165 21 Prague 6 – Suchdol
klimesova@pef.czu.cz
and
Czech Academy of Sciences, Institute of Information Theory and Automation
Dept. of Image Processing, Pod vodárenskou vìží 4, 182 08 Prague 8, Czech Republic
klimes@utia.cas.cz

Abstract: An important part of decision tasks is classification of objects into classes. If there is a set of input data, which class memberships are known, based on these data it is possible to take a decision on membership of new data of the same type. Nowadays many classification technologies and algorithms are developed. Increased requirements are taken on these technologies in regard to increased precision, shorter classification time and so on. One of the possible solutions how to increase fruitfulness of classification is utilization of composed classifiers. The contribution deals with a problem of creating a composed classifier with Boosting architecture, whose components are composed of classifiers working with k - NN algorithm (k - th nearest neighbour).

Keywords: classification, composed classifier, boosting, k - nearest neighbours, ID3

1 Classification

Classification is a process merge of objects into the classes. This concept could have two quite different meanings. In the first case could be the goal of classification distinguishing the existing classes in data. In the second case the goal is based on the classed object find function, which classes the new non-classed object in one of existing classes. The first case is type uncontrolled learning and it is marked as clustering. The second case is type of controlling learning that demand the existence of a set of objects to those to be known to have a membership to classes.

According to access to the classification and by characteristic properties of classifiers the classifiers can be divided into few groups and types. First main division of classifiers is possible to make by set entrance data, thus according, if system has feedback or no in learning process:

- ***controlled learning,***
- ***uncontrolled learning*.**

By the historical development:

- ***classic statistical classifiers*** - use ordinary mathematical equipment,
- ***modern classifiers*** - fuzzy practices, neural networks, etc.

Statistical classifier is also possible to divide into [3]:

- ***parametric statistical classifiers*** - assuming that conditional densities of individual classes are certain type, e.g. normal density, which statistical parameters are estimate from prototype data set,
- ***non-parametric statistical classifiers*** – not assuming any shape of densities. There are refered directly on prototype data set, i.e. densities are estimate straight from prototype set,
- ***semiparametric classifiers*** - try to make a use of the best properties former duo access and create some transition between them.

2 Composed Classifier and its Architecture

Composed classifier is a composition of component classifiers, which predictions are connecting by combining classifier, unlike contrast to simple classifiers. There are several architectures for possible combination of classifiers. Main architectures for combination of classifiers are:

- *Stacked Generalization*
- *Boosting*
- *Recursive Partitioning*

Common feature of these architectures is the fact that they trying to reduce error of classification by combined prediction of component classifiers.

Stacked Generalization is a level architecture for combinating classifiers, in which classifiers on a higher level combine prediction of classifiers immediate on the lower level.

Boosting tries to increase the precision of given classifier by creating a complementary component classifier by filtration of a training set. On recovery

resultant prediction is used to vote between existing classifier and new created component, Figure 1.

Recursive Partitioning is method for connecting of classifiers, in which domain space is divided recursively into many areas. One classifier for prediction is applied in every of these areas.

2.1 Boosting

Architecture Boosting offers progress by which improves the precision of existing classifier. A class of concepts is difficult to learn in PAC (*Probably Approximately Correct*) model, if exists an algorithm in polynomial time, which with high credibility proves to learn the arbitrary concept in class with small error.

Class of concepts is weakly for learning, if with high credibility can be arbitrary concept in class learnt with error, which is only about something smaller than accidental classification. In this work was the solving of a problem to mark up the precision of classifier k-NN i.e. k-th the nearest neighbour.

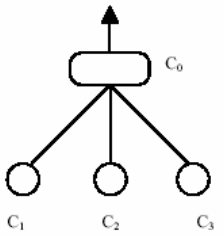

Figure 1
Architecture Boosting. C_1 is given basic classifier,
C_2 and C_3 are supplementary classifiers, C_0 is combining classifier

3 Experiments

In this work was used composed classifier based on architecture Boosting, which forms more-accurately composed classifier by combine existing basic k-NN classifier with the others supplementary classifiers. Supplementary classifiers were created with methods **CR** (***C**oarse **R**eclassification),* **DM** (***D**eliberate **M**isclassification)*, **CF** (***C**omposite **F**itness)* and **SF-F (***C**omposite **F**itness-**F**eature **S**election)* [7]. These methods verified various opinions to create accurate supplementary classifier.

Brief description of individual method:

- **CR** chooses the most accurate component,

- **DM** is a variant of CR, which increases the difference by the selection components, which has most highly precision on modified training data (data with modify marking the classes),
- **CF** chooses supplementary classifier, whose applicability is given precision resultant of composed classifier which utilize selection component,
- **CF-FS** is variation, which increase difference component by cut-down of the number of attributes included in calculation of distances.

Single methods of boosting for verification improvement precision were tested on data from remote sensing of the Earth by LANDSAT. The data set consists of 368,152 specimens surface of the Earth, where one of them represents area of 30 x 30 meters, representing a total of 332 sq km of land. The specimen of the Earth surface is characterized by a 7- dimensional vector. These partial components are describing the brightness of the seven spectral bands.

Figure 2
The real picture from remote sensing of the Earth

The existing basic classifier builds used solution of boosting into level 0 in composed architecture *Stacked Generalization*. As classifier of level 1 is used decision tree generated by algorithm ID3 (*Inductive Dichotomize 3*) [8] or k-NN classifier.

Algorithm ID3

Algorithm ID3 is the best known algorithm generating decision tree by method from the top to down. Finishing criterion of this algorithm is, that every subspace contains only examples of the one class. If a set of attributes is sufficient, possibly decision trees constructED by the mentioned progress, correctly classify the training examples. For classification of the new example is needed by monitoring the way from radical element of decision tree until the end node. At every interior node follows branch corresponding to the value of testing attribute. A class near the terminal node introduces prediction of the class for existing exampleS.

Classifier k-NN

The classification rule of the *k*-nearest neighbour is non-parametric statistical criterion. This algorithm designates class of unknown quantitIES sample according to base the classes to the nearest neighbour. Algorithm operates with constant numberS of attributes and it doesn't need to know statistical distribution of training set. At classification by choosing distance metric calculate distance of testing sample to all placing training samples. Then the sample is assigned from training and set into the class of the nearest neighbour.

Evaluation and Conclusion

The already mentioned method ***C**omposite **F**itness-**F**eature **S**election,* is the best for creating supplementary classifiers independently on applications combining classifierS by general results on testing database. In the three cases for methods ***C**oarse **R**eclassification*, ***D**eliberate **M**isclassifiacation* and ***C**omposite **F**itness* is better combining classifier ID3. Additionally ID3 needs smaller numberS of supplementary classifiers (1 till 3) on achievement of this precision. In the aggregate the best average precision has composed classifier with 10 supplementary classifiers combination with 5-NN classifier.

The best average accuracies for existing combining classifiers are mentioned in percentages in Table 1, where **number DK** is the number of supplementary classifiers.

	CR		DM		CF		CF-FS	
	Number DK	**Precision [%]**	**Number DK**	**Precision [%]**	**Number DK**	**Precision [%]**	**Number DK**	**Precision [%]**
ID3	1	80,21	5	79,31	1	80,51	3	83,47
5 - NN	5	78,68	10	77,39	3	79,18	10	84,40

Table 1

The best resultant accuracies for single methods with combining classifiers ID3 and k-NN

Methods for generation suitable supplementary classifiers were mentioned for existing k-NN classifier. These classifiers were combined by duo combining classifiers - ID3 and k-NN. By comparing the combined classifiers ID3 and k-NN with majority voting, on the average algorithm ID3 reaches for all methods and has better results with smaller number of supplementary classifiers.

The importance of particular properties required from supplementary classifiers (precision and difference of classification) and enhancing of precision against the classification with based classifier also by combination with small numbers of minimum supplementary classifiers was verified by testing.

On the basis of received experiences, finally we can state, the composed classifier (created by architecture Boosting) doesn’t increase the precision of the existing classifier learnt on the training set.

Acknowledgement

This work is supported by the VEGA project No. 1/2185/05 and MSM project No. 6046070904.

References

[1] Brodley, C. E.: Recursive Automatic Algorithm Selection for Inductive Learning. Ph.D. Dissertation, University of Massachusetts, Amherst, 1994

[2] Klimešová, D., Ocelíková, E.: Spatial Data and Context Information. In: Proc. of XX Congresso Basileiro de Cartografia, IX Congresso Nacional de Engenharia de Agrimensura, De 7 a 12 de Outubro de 2001, Porto Alegre, Rio Grande de Sul, Brasil, 2001

[3] Michie, D., Spiegelhalter, D. J., Taylor, C. C.: Machine Learning Neural and Statistical Classification, Ellis Horwood 1994

[4] Ocelíková, E., Krištof, J., Kočiš, M.: Using Bayes Classifier in Image Data Classification from Remote Earth Sensing. In: Proc. of the Scientific Conference 'Artificial Intelligence in Industry', High Tatras, April-22-24, 1998, pp. 225-230

[5] Ocelíková, E., Krištof, J.: Statistical Classification Techniques In: Proc. of the 34th Spring International Conference 'Modeling and Simulation of Systems MOSIS´, May 2-4, 2000. Rožnov pod Radhoštem, pp. 111-116

[6] Ocelíková, E., Krištof, J.: Classification of Multispectral Data. *Journal of Information and Organizational Sciences,* Vol. 25, Number 1, Varaždín, 2001, pp. 35-41

[7] Skalák, D. B.: Prototype Selection for Composite Nearest Neighbor Classifiers. CMPSCI Technical Report 96-89, 1997

[8] Utgoff, P. E.: Perceptron Trees: A Case Study in Hybrid Concept Representations. Connection Science 1: 377-391, 1989