# Support of Contextual Classifier Ensembles Design

Janina A. Jakubczyc
Wroclaw University of Economics
ul. Komandorska 118/120,
53-345 Wrocław, Poland
Email:
janina.jakubczyc@ue.wroc.pl

Mieczysław L. Owoc
Wroclaw University of Economics
ul. Komandorska 118/120,
53-345 Wrocław, Poland
Email:
mieczyslaw.owoc@ue.wroc.pl

☐

*Abstract*— An idea of contextual classifier ensembles extends the application possibility of additional measures of quality of base and ensemble classifiers in the process of contextual ensembles design. These measures besides the obvious classifier accuracy and diversity/similarity take under consideration the complexity, interpretability and significance. The complexity (the number of used measures and multi level measure structure), the diversity of the scales of used measures and the necessity of the fusion of different measures to one assessment value are the reasons for user support in contextual classifier ensembles design using fuzzy logic and multi criteria analysis. The aim for this paper is an idea of the framework of the process of contextual ensemble design.

## I. INTRODUCTION

The contextual classifier ensemble other than, name it classical ensemble, gives the possibility to look more profoundly at problem under consideration and classifier ensemble design. The reason is context that is the criterion for diverse classifier creation. The contexts of the classification problem give the possibility to view it from different perspectives, to get more familiar with it, to choose the more appropriate contexts for each case.

This approach of classifier ensembles design extends the application possibility of additional measures of quality of the base and ensemble classifiers in the process of contextual ensemble design. These measures besides the obvious classifier accuracy and diversity/similarity take under consideration the complexity, interpretability and significance. The additional measures appear in the stages of the process of contextual ensemble design. This process and the idea of contextual classifier ensemble are introduced the section II.

The framework of contextual classifier ensembles design is included in the section III.

The complexity (the number of used measures and multi level measure structure that), the diversity of the scales of used measures and the necessity of the fusion of different measures in one assessment value for the given stage of process design are the reasons for applying fuzzy logic and multi criteria analysis. This methodology is introduced in section IV.

The first results of experiment that employs this idea contextual classifier design support is included in the section V. And finally the section VI gives some summary of the introduced idea.

## II. THE ESSENTIALS OF CONTEXTUAL CLASSIFIER ENSEMBLE

Classifier ensembles allow the different needs of a difficult problem to be handled by classifiers suited to those particular needs. Classifier ensembles provide an extra degree of freedom in the classical bias/variance tradeoff, allowing solutions that would be difficult, if not possible to reach with only a single classifier. Because of these advantages, classifier ensembles have been applied to many difficult real-world problems see: Oza and Tumer, 2008 [5], Patel and Nawathe 2013 [11].

The classifier ensemble is a set of base classifiers that together are solving the problem of the discrimination. There are three basic mechanisms of creation the classifier ensembles. The first is a mechanism of creation the base classifiers see: Dietterich, 2000 [2], Zhang and Ma, 2012 [9]. There are the following approaches:

- data manipulation - the models for different data subsets of the learning set or for different attribute subsets
- different techniques for modeling one learning data set
- the one kind of model with different parameter sets for one learning set.

Contextual classifier ensemble is a variant of classifier ensembles that is based on data

---

manipulation. Instead of random samples generation, or weighing examples according to correct classification in consecutive modeling step, contextual classifier ensemble generates attribute sets according known or discovered context from data that describe problem under consideration – compare: Jakubczyc, 2007 [3]. The most interesting situation is when the contexts are discovered in learning set. This case is a subject of this study. The context can be a single category i.e. localization with the urban and rural values that determine two contextual situations (localization=urban, localization=rural) or complex concept that is described by a set of attributes in the form of some kind of model for example decision tree. Each base classifier represents description of

classification problem for single context or for single contextual situation. In the decision tree model each path from the root to the leaf indicates the contextual situation and the whole decision tree means the context. Whether it is context or contextual situation is determined by the type of relation between contexts and learning examples that represent problem under consideration are perspective and partiality (Figure 1). The perspective relation stands for possible contextual views on a problem. The partiality relation describes only subsets of more appropriate models of the problem for contextual situations. The last possible relation means partiality and perspective relation simultaneously (not presented in the Figure 1)



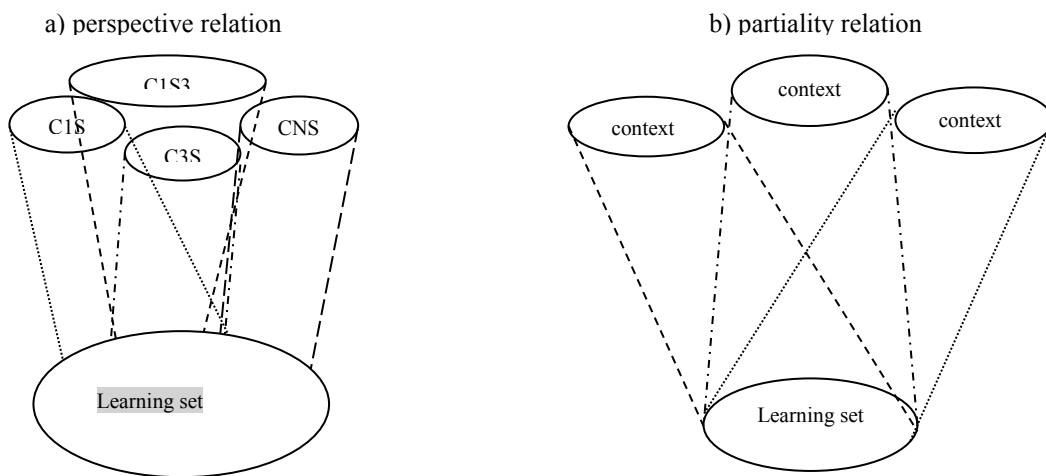a) perspective relation

b) partiality relation

Figure 1. The relations between contexts and learning set of example

The type of relation that is used in creation contextual classifier ensembles is determined by the quality of contextual models. If the level of classification accuracy is acceptable for classifiers of identified contexts there is perspective relation. In this case the contextual classifier ensemble consists of models for each discovered contexts. On the contrary, unacceptable accuracy forces the one to take under consideration only some models for contextual situations that may guarantee the requested quality for each context. In this case there is partiality relation that may not cover the whole learning set. In this case the contextual classifier ensemble consists of models for the contextual situations not for whole context (for example in Figure1 C1S1 states for contextual situation 1 from context 1 model) with acceptable accuracy. It seems acceptable because the learning set may not be representative sample according to statistic theory.

There is also possible the mix relation that joins the first two. Our interest for now is in the first one that means the base classifiers are models for whole contexts (perspective relation). For now this focus allows us to exempt from the attendance of

uncovered example by contextual classifier. It will be the subject of future research.

The selection of classifiers to the ensemble that solve the classification problem is the second mechanism. There are two problems: how many classifiers and which classifiers should be chosen.

The general answer for the first question is: "the more classifiers the better" according to the jury theorem of J.A. Marquis Condorcet - originally: Condorcet Marquis J.A.: Sur les elections for scrutiny, [in:] Histoire de l'Academie Royale des Sciences, 31-34, 1781 see: Cunnigham, 2007 [1] - but the empirical studies shows that even few classifiers may improve classification accuracy described compare: Abreu and Canuto, 2007 [10] Schiele, 2002 [7]. They have showed that ensemble of three or of five classifiers may bring much improvement too. In the case of contextual classifier ensemble the number of base classifiers in ensemble is limited to the number of identified contexts or the number of contextual situations.

The choice of the most appropriate base classifiers is generally based on two measures – applied by Kuncheva and Whitaker, 2003 [4]. The

first measure is classification accuracy that should be above 50%. The second are the measures of diversity. Intuitively it means that chosen base classifiers should be mistaken on the different cases thus they can complement one another. Because the results of empirical studies are ambiguous, i.e. there is no linear relation between the diversity and accuracy, so the increase of diversity not necessary causes the increase of classification accuracy. So there is a need for additional measures that can support the choice of base classifiers. The contextual approach to creation classifier ensembles gives such possibility. The additional perspectives may be user evaluation of identified contexts and contextual base classifiers.

There is a need for additional measures also in the next step of combining single decisions of base classifiers into one final decision (it is the third mechanism). Applied techniques are determined by continuity or discreteness of class values. Since decision tree is chosen as the classification algorithm, we are dealing with nominal values of classes. In this case, the voting schemata are accessible techniques of the combination of the decision of base classifiers in this case. In simple voting schemata each classifier has one vote. But looking at different base contextual classifiers according to their quality, significance, comprehensibility, the equally assignment of votes is difficult to accept. The none-equivalence of base classifiers should be taken into account in global assessment of contextual ensemble.

The context as a criterion for creation classifier ensembles gives the possibility to look more profoundly at generated classifiers trough different contexts and at identified contexts alone. The profound assessment of the base contextual classifier and contextual ensemble needs the framework to account for the following aspects: identification of possible assessment criteria for single contextual classifier and for contextual classifier ensemble with different level of detail and aggregation, the possibility of evaluating simultaneously qualitative and quantitative assessments, contextual classifier ranking trough fuzzy multi-criteria analysis. The general framework with methodological approach is presented in the next section. The working example is shown in further section.

### III. THE FRAMEWORK FOR CONTEXTUAL CLASSIFIER ENSEMBLE

The proposition is aimed at more reasonable creation of classifier ensemble through the different, and not random criterion to build classifier ensembles. The contextual classifiers give the possibility of more profoundly evaluation of base and ensemble classifiers. The user active participation in this process of assessment is not to overestimate. The general framework of creation contextual ensembles is introduced at Figure 2. There is one global goal to choose the most appropriate contextual classifier ensemble and one intermediate goal to choose the best base contextual classifiers.

The classification problem is difficult so there is not possible to built the one model with acceptable level of accuracy. The possible solution is classifier ensemble, in this case with the contexts as criterion for building base classifiers.
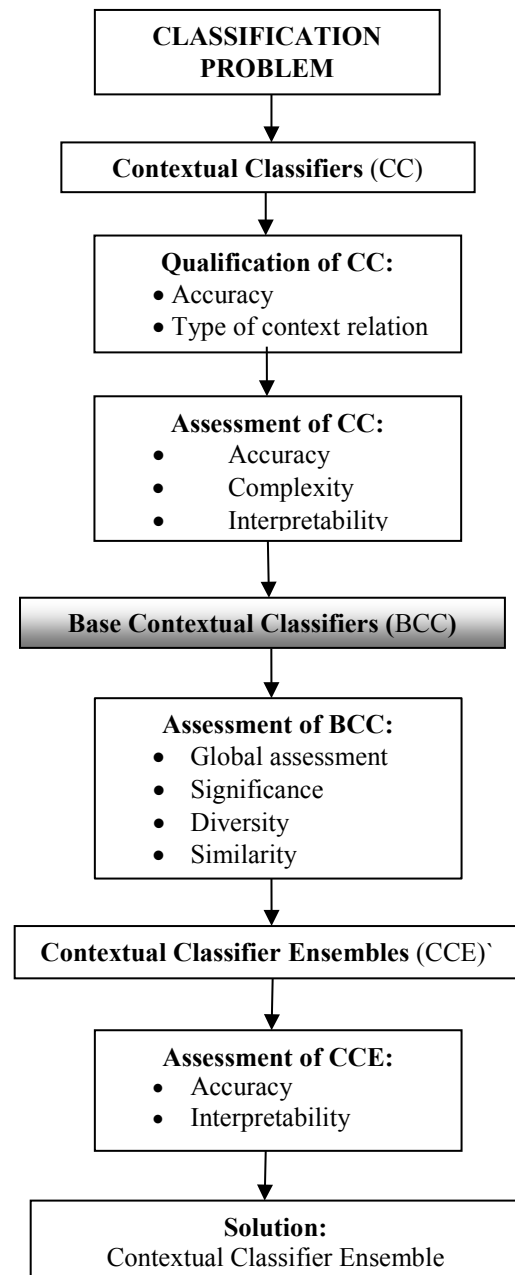


Figure 2. The general framework of creation contextual ensemble

The contextual classifiers to became the base contextual classifiers have to achieve the level of accuracy above 60% for each contextual situation (it takes place when the relation between contexts and learning set of examples is so called 'perspective relation') - presented in Figure 1 - and for each class. The contextual classifiers that have passed the qualification step are evaluated objectively according to accuracy and subjectively with support of detailed measures for complexity and interpretability (see Table 1.). To the ranking of base contextual classifiers are used two measures for each pair: similarity and diversity and user evaluation of classifier significance.

The similarity and diversity do not play a key role in process of evaluation since the number of base contextual classifiers are finite, so there can be used all possible combination of classifiers. The measures of similarity and diversity can be used to examine their quality influence.

The possible combinations of base contextual classifiers then are evaluated using objective measure of accuracy and subjective measure of interpretability. The ranking list of contextual classifier ensembles is the result of this process. The choice the most appropriate contextual classifier ensemble is up to the user. The detailed methodology used in the evaluation process is presented in the next section.

## IV. THE USED EVALUATION METHODOLOGY

The method for evaluation of base contextual classifiers and contextual classifier ensembles depends on character of the criteria and their properties (see Table 1). As we can see the criteria have the hierarchical structure, the different types of value (qualitative, quantitative), and different value scale. This imposes the use of fuzzy logic system introduced by L. Zadeh [8] as representing criteria values at all levels. The non-equivalence of identified criteria determines the use of pair wise matrices for the hierarchical structure of evaluation criteria means the need for gradual criteria aggregation at each higher levels.

The first step (the first column in Table 1) is the candidate qualification. To pass, the contextual classifier should have the level of accuracy above 60% and the context with perspective relation. The evaluation of contextual base classifier includes the three layers hierarchy of criteria and gradual way of evaluation.

The level 1 encompass four criteria. Each of them is represented as fuzzy membership function. The shape of these functions and key points are determined as result of analysis and requirement. The value of interpretability increase s when the value of 'novelty' decreases. There is assumed the 20% level of 'novelty' that does not influence the interpretability value.

Table 1. The criteria for assessment of base contextual classifier and classifier ensemble

| Contextual classifier | Base contextual classifier | | | Contextual classifier ensemble |
|---|---|---|---|---|
| | Level 1 | Level 2 | Level 3 | Level 4 |
| Accuracy  Type of context relation | Comprehensibility | Model interpretability | Base contextual classifier assessment | Accuracy  Interpretability |
| | Novelty | | | |
| | Number of descriptors | Model complexity | | |
| | Number of branches | | | |
| | | Model accuracy | | |
| | | | Similarity | |
| | | | Diversity | |
| | | | Significance | |

The value of 'comprehensibility' influences 'interpretability the another direction: the higher 'comprehensibility' value the higher the value of 'interpretability'. The range with full understanding is assumed in the interval 85-100%.

The fuzzy function of number of model descriptors has the two key points that determine the optimal range between 2 and 10 descriptors. The more descriptors indicate the more complexity of the classification model. Similarly is the matter with the 'number of branches' but with duplicated values.

The synthetic measure for base contextual classifier takes under consideration the non-

equivalence of identified criteria. The pair wise comparison matrices (see: Saaty, 1977) are applied to derive implicit weights for a given set of criteria. The user has the possibility to introduce the significance weights for pair comparison matrix.

The model accuracy (Figure 3) has three key points. The first one indicate the minimal level of accuracy that is required for classifier ensemble i.e. 60%. The next two key points indicate the most desired range of accuracy i.e. 80%-95%. The values close to 100% are unacceptable.
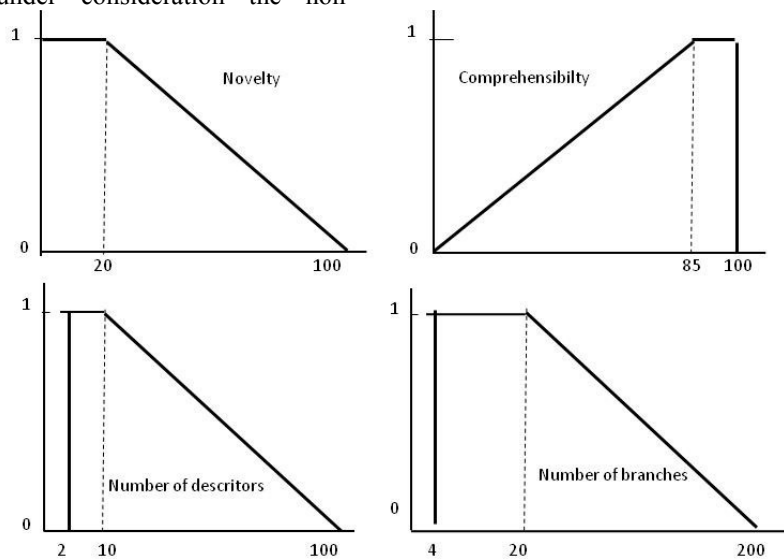


Figure 3. The shapes of fuzzy membership functions of the level 1 criteria

The complexity of contextual classifier is determined by detailed parameters of the number of descriptors and tree branches (Figure 4). The higher value of model complexity indicate the worst classifier quality. The first and second key point indicates the range of minimal level of complexity that may differ according to number of possible descriptors. After that , the quality of model decreases when the complexity increases. For different classification problem the user recommendation of key point may differ.

The interpretability indicates to what extent user understand and unexpected the classifier according to his knowledge. In this case the fuzzy membership function has the shape of monotonically increasing – the higher the value of interpretability the higher the quality of classifier.

The next level 3 takes into account synthetic evaluation of base classifier ensemble (interpretability, complexity and accuracy) and similarity and diversity measures and significance of base contextual classifiers.

The similarity measure is used as a means to eliminate the most similar classifiers from one ensemble. On the contrary the measure of diversity gives the possibility to include in one ensemble the most different contextual classifiers.

The weights assignment to particular base contextual classifier is the task for the user with his knowledge and experiences. So the assigned weights may differ from one user to another. The pair wise matrices are applied to this operation.

The calculation of synthetic measures takes place at level 2 for interpretability, complexity, and accuracy, at level 3 for base contextual classifier

evaluation and at level 4 for contextual classifier ensemble.

All these calculation take into account non-equivalence of used criteria and identified contexts . This not-equivalence is measured by the general assessment of the base contextual classifier and the pair-relation of significance between contextual classifiers determined by the domain expert. Such approach allows for more precise choice of the most appropriate classifier ensemble for the

problem under consideration. There are not a priori settled weights. The user each time introduces the values of the pair comparison between the base contextual classifiers. The system interface gives the possibility to view the all possible values with their interpretation.
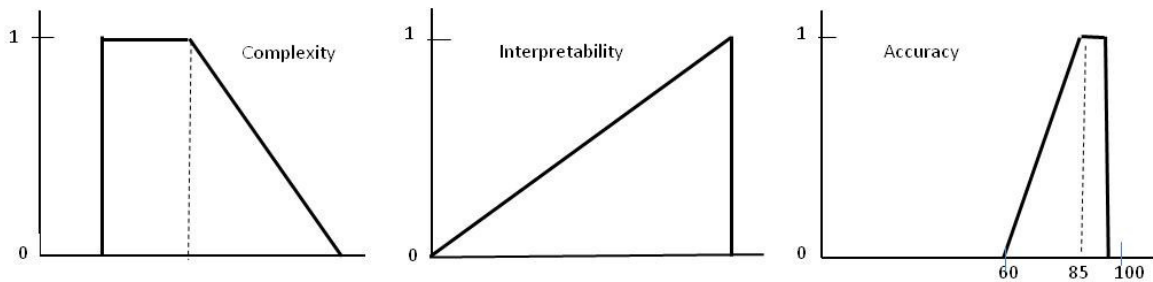


Figure 4. The shapes of fuzzy membership functions of the level 2 of criteria

The means for criteria aggregation are three following operators :

- maximum pessimism $D1 = \min(\mu_1(x_1)^{\alpha 1}, \mu_2(x_2)^{\alpha 2}, \dots, \mu_n(x_n)^{\alpha n})$
- multiplicative $D2 = \prod_{i=1}^{n} \mu_i (x_i)^{\alpha i}$
- additive $D3 = \sum_{i=1}^{n} \frac{\alpha_i * \mu_i(x_i)}{n}$

where:

$\mu_1(x_1), (\mu_2(x_2), \dots, (\mu_n(x_n)$ - membership function,

$\{x_i\}$ – numerical and nominal criteria of classifier quality

$\alpha_1, \alpha_2, \dots, \alpha_n$ – the relative significance of the criteria and in the case of global criteria assessment the relative significance of base contextual classifiers.

The global assessment of contextual classifier and contextual classifier ensemble is determined by the three criteria: maximum pessimism (D1) multiplicative (D2), additive (D3) to observe the relation between them. D2 and D3 have the propriety of little value compensation of one criterion by the increasing the other criteria. The conformity of the obtained results for D1, D2, D3 increases the confidence of study results. All of them gives the global value in the range of 0 and 1 (the higher the better).

### V.WORKING EXAMPLE

The experiment was conducted on the problem of the banks client behavior and the estimation of clients behavior. The classification task was to predict whether the client is active or non-active.

This problem is very important for each bank management If the client is non-active there is an indication to take pro-active action to the client to keep him as an stable bank client for the future and do not let him switch the bank for example. The bank provided 24 000 examples in learning set. The Bank is providing 24 000 representation data. There were nine discovered context that were the basis for contextual classifier design.

Our experiment aimed at two issues. The first is to verify the usability of the proposal of the framework and applied methodology. It proved to be appropriate and flexible. The flexibility concerns the automatic or manual design of contextual classifier ensemble. Because of the subjectivity of some used measures, as for example interpretability and comprehension, assumption in automatic process for these measures default values was not the best idea.

The second aspect refers to view the classification problem according to discovered contexts and knowledge/experience of the user. Generally the designed contextual ensemble by the various users were different but their quality quite similar. They differ from the weights assignment to significance of contexts and contextual models. Each of users built own fuzzy rule set and used the possibility to design few contextual ensemble to choose the best one. It results of knowledge and experience of the users.

## VI. Summary

The context as criterion for creation classifier ensembles gives the possibility to consider more profoundly classification problem under consideration. The user activation seems to be important too, although the interactive problem solving do not forces the user to join the process. But from our current experiments with the user the solutions were more reasonable. The applied approach proved to be adequate. The applied methodology showed to be appropriate too. Additionally the users fast get familiar with it.

We observed how the level of diversity changes the quality of classifier ensemble, how the knowledge and experiences influence the solutions. But it is too few experiments to conclude more general conclusions. Our work is going on.

## References

[1] Cunnigham P., Ensaml;e Techniques Technical Report UCD-CSI-2007-5 April 2, 2007

[2] Dietterich T. G.: Ensemble methods in Machine Learning. [in:] The Proceedings of 1th International Workshop on Multiple Classifier Systems, s. 1-15, 2000

[3] Jakubczyc J. A., Contextual Classifier Ensembles. [in:] Abramowicz W. (ed.): Business Information Systems, LNCS 4439, Springer 2007.

[4] Kuncheva L.I., Whitaker C.J.: Measures of diversity in classifier ensembles, Machine Learning 51, s. 181-207, 2003.

[5] Oza N. C. and Tumer K., Classifier Ensamles: Select Real World Applications. Information Fusion Vol 9 Issue 1 2008, Elseviewer

[6] Saaty T. Scaling Method for Priorities in Hierarchical Structures //J. of Mathematical Psychology. 1977. Vol. 15. 3. p. 234 – 281.

[7] Schiele B.: How many Classifiers Do I Need? IEEE Pattern Recognition, vol.2, 2002

[8] Zadeh L. A.: Is there a need for fuzzy logic?, Information Sciences an International Journal 178 (2008) 2751-2779, www.elsevier.com/locatelinks

[9] Zhang C., Ma Y.(eds.): Ensemble Machine Learning: Methods and Application, Springer Science+Business Media LLC 2012

[10] Abreu M. C. C., Canuto A. M. P.: Using Fuzzy, Neural and Fuzzy-Neural Combination Methods in Ensembles with different Levels of Diversity, in: Proceeding ICANN'07 Proceedings of the 17th international conference on Artificial neural networks, Springer-Verlag Berlin, Heidelberg 2007, p. 349-359

[11] Patel S. P., Patel M. P., Nawathe A. N.: A review on Ensemble of Classifier Using Artificial Neural Networks as Base Classifier, International Journal of Computer Science and Mobile Applications, vol.1 Issue 4, October 2013, p 7-16