A Multi-objective Optimisation Design Approach for Ensemble Member Combination in Binary Classification Tasks *

Victor Henrique Alves Ribeiro *
Matheus Henrique Dal Molin Ribeiro **,***
Leandro dos Santos Coelho *,*** Gilberto Reynoso-Meza *

* *Industrial and Systems Engineering Graduate Program (PPGEPS),
Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, PR,
Brazil, (e-mails: victor.henrique@pucpr.edu.br,*

leandro.coelho@pucpr.br, g.reynosomeza@pucpr.br)

** *Universidade Tecnológica Federal do Paraná (UTFPR), Pato
Branco, PR, Brazil, (e-mail: mribeiro@utfpr.edu.br)*

*** *Department of Electrical Engineering (PPGEE), Universidade
Federal do Paraná (UFPR), Curitiba, PR, Brazil*

Abstract: Machine learning algorithms are usually trained by means of optimising a single global score, and it is possible to have a poor performance when dealing with classification problems where a lower type I or type II error rate is more desirable. To deal with this situation, multi-objective optimisation techniques can be applied for finding the trade-off among the different objectives. However, such approach results in a set of non-dominated solutions, and it is important to select one preferable solution. Thus, this paper proposes the application of a holistic multi-objective optimisation design procedure for ensemble member combination. To do so, a multi-objective problem is formulated in order to maximise the class-specific recalls of a weighted majority voting ensemble. Such problem is optimised with a multi-objective evolutionary algorithm. Next, a ranking technique is used find a preferred solution based on the during the multi-criteria decision making step.

Keywords: Ensemble learning, multi-objective optimisation, multicriteria decision making, classification, machine learning.

1. INTRODUCTION

According to Rokach (2010), it is well known that the ensemble learning methodology can be used for improving prediction performance on supervised learning tasks. However, machine learning algorithms and ensemble models are usually trained by means of optimising a global prediction score. Therefore, such methods can present a lack of performance on complex problems, where a global score does not provide enough information on the quality of the model.

To deal with this situation, multi-objective optimisation (MOO) can be used. By adjusting the combination of ensemble members, it is possible to optimise different conflicting objectives and understand the trade-offs when building predictive models. With such, it is possible to generate ensembles with a preferred performance for the desired supervised learning tasks, such as classification.

* This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)*, the *Conselho Nacional de Pesquisa e Desenvolvimento (CNPq)*, and the *Fundação Araucária (FAPPR)* - Brazil - Finance Codes: PROSUC-159063/2017-0, 304066/2016-8-PQ2, 437105/2018-0-Univ, 303908/2015-7-PQ, 204893/2017-8, 404659/2016-0-Univ and PRONEX-042/2018.

The application of MOO algorithms for the optimisation of ensemble learning models has been performed by many authors, and is known as multi-objective ensemble generation (MOEG) (Gu et al., 2015). Most of the works deal with ensemble member generation, where MOO is applied in order to generate a set of diverse predictive models (Tan et al., 2014; Smith and Jin, 2014; Rosales-Pérez et al., 2014; Gu and Jin, 2014; Rosales-Pérez et al., 2017; Peimankar et al., 2017; Ribeiro and Reynoso-Meza, 2018). However, it is also possible to find works on ensemble member selection (Peimankar et al., 2017, 2018; Ribeiro and Reynoso-Meza, 2018), where a set of candidate models is pruned in order to find an optimal subset, and ensemble member combination (Wang et al., 2014; Zhang et al., 2015; Zhao et al., 2018), where weighted average and weighted voting are usually performed.

When dealing with multiple objectives, the application of MOO algorithms results in a set of non-dominated solutions, where one solution cannot be considered better than others. Thus, on the works that perform ensemble member selection and combination, a final ensemble is usually selected based on highest validation data set scores. However, on such cases, multi-criteria decision making (MCDM) techniques can aid in the selection of a preferable solution based on design preferences. Therefore, the au-

thors believe that it is possible to use a multi-objective optimisation design (MOOD) procedure Reynoso-Meza et al. (2014) for building ensembles, composed of the formulation of a multi-objective problem (MOP), the application of a MOO algorithm, and the usage of MCDM techniques.

In order to perform the ensemble member combination task, this paper proposes the formulation of a holistic MOOD procedure. First, a MOP is formulated for the optimisation of a weighted majority voting ensemble. Next, a MOO algorithm is applied in order to extract the set of non-dominated solutions. Finally, a MCDM ranking technique is applied in order to find a preferable solution that meets the decision maker's preferences.

Experiments are performed in order to validate the proposed methodology with two other ensemble learning methods. In total, thirty runs are executed on seven different binary classification data sets. Statistical tests are performed on the resulting data, indicating that the proposed technique outperforms the accuracy of other methods on three data sets, while performing similarly on other three. With such, the authors conclude that the application of a holistic MOOD procedure for ensemble member combination has promising results, and can be used in order to solve complex classification problems.

The remainder of this paper is presented as follows: Section 2 brings a background on ensemble learning, MOOD and MOEG; Section 3 introduces the proposed methodology for ensemble member combination; Section 4 explains the experiments used in this work; Section 5 shows the results and discussion; and Section 6 presents the conclusions and future works.

2. BACKGROUND

In order to build ensembles using MOO techniques, some background on ensemble learning, MOOD and MOEG is necessary. This section brings an overview on such fields of study.

2.1 Ensemble Learning

According to Rokach (2010), the supervised learning task is typically composed of three components: a data set, an learning algorithm and a predictive model, where the learning algorithm creates a predictive model for a specific training set. By combining the outputs of different predictive models, it is possible to achieve a new model that outperforms every one of its members. This is known as the ensemble methodology, and is composed of several supervised learning tasks, each one composed of a training data set, an inducer and a predictive model.

In order to build ensembles of predictive models, the existence of the following three steps is indicated in Woźniak et al. (2014): (a) member generation step; (b) member selection step; and (c) member combination step. The first one is related to creating diversity in the ensemble's predictions, which can be done by: manipulating the training sets; and/or building models with different hyperparameters or learning algorithms (Rokach, 2010). The second step is important for selecting only models that are relevant for the prediction task, since some of them may reduce the ensemble's predictive performance. It can be done

by using: ranking-based approaches and/or search-based methods (Rokach, 2010). The last step is applied in order to combine the outputs of multiple predictive models into a final and more powerful prediction. Some techniques for such include: majority voting; weighted majority voting; average; weighted average; and stacking. Such techniques are explained in details by Gunes et al. (2003).

2.2 Multi-objective Optimisation Design

When dealing with complex engineering problems, there is generally not only one objective to be optimised. Moreover, there can exist trade-offs among the multiple objectives, which makes the selection of a final solution a difficult task. A MOOD approach can deal with such complex scenarios, and is composed of the following three steps: MOP definition; MOO; and MCDM (Reynoso-Meza et al., 2014).

The MOP formulation is responsible for defining the problem's objectives, decision variables and constraints. As referred in Miettinen (2012), a MOP with m objectives can be stated as follows:

$$\max/\min_{\mathbf{x}} \mathbf{J}(\mathbf{x}) = [J_1(\mathbf{x}), \dots, J_m(\mathbf{x})] \quad (1)$$

subject to:

$$\mathbf{K}(\mathbf{x}) \leq 0 \quad (2)$$

$$\mathbf{L}(\mathbf{x}) = 0 \quad (3)$$

$$\underline{x}_i \leq x_i \leq \overline{x}_i, i \in [1, \dots, n] \quad (4)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_n]$ is defined as the decision vector with $\dim(\mathbf{x}) = n$; $\mathbf{J}(\mathbf{x})$ as the objective vector and $\mathbf{K}(\mathbf{x})$, $\mathbf{L}(\mathbf{x})$ as the inequality and equality constraint vectors, respectively; $\underline{x}_i, \overline{x}_i$ are the lower and the upper bounds in the decision space.

In the MOO step, a set composed of several solutions is retrieved by a search algorithm. Such set is the Pareto front approximation, where all solutions are considered as valid ones, since a solution which performs better than other on one objective will perform worse on other objectives. Multi-objective evolutionary algorithm (MOEA) is a family of search algorithms, which has shown to be a useful MOO tool for a wide range of engineering problems (Coello and Lamont, 2004).

Finally, it is extremely important to identify a final solution from the Pareto front when working on engineering problems. The MCDM step is applied by the decision maker in order to find such final solution. In this step, it is possible to use many different visualisation techniques, which aid the decision maker in understanding the problem's trade-offs, and ranking techniques, which order the solutions based on the decision maker's preferences (Reynoso-Meza et al., 2014).

2.3 Multi-objective Ensemble Generation

The application of MOO for generating ensembles has been the subject of study by many authors, and much work has been done on the definition of different objectives and

decision variables for the formulation of a MOP. Most of the works deal with the member generation task. However, MOO has also been applied for the tasks of ensemble member selection and combination.

The application of MOO algorithms for member generation has the objective of generating a pool of diverse base models to compose an ensemble. This can be done by the manipulation of: features; training instances; and/or model's hyperparameters. According to Gu et al. (2015), the formulation of a MOP for finding the trade-off between prediction scores and complexity is presented as the most promising such task.

For the member selection task, MOO algorithms can be used in order to prune a pool of candidate ensemble members. With such, it is possible to build a better ensemble with just a subset of the original models. In the member combination step, instead of selecting the members that will be part of the ensemble, MOO is applied in order to find optimal weights for each model in a weighted average ensemble.

Despite the application of MOO in ensemble learning, to the best of the authors' knowledge, there has not been a work that applies MCDM ranking techniques in order to select a final preferable ensemble. Thus, this work proposes the application of a holistic MOOD procedure for the ensemble member combination task.

3. MULTI-OBJECTIVE ENSEMBLE MEMBER COMBINATION

This paper proposes an MOOD procedure in order to build an ensemble of classifiers using an weighted majority voting scheme. To do so, the formulation of a MOP, the application of a MOO algorithm to extract the Pareto front approximation, and the selection of a final solution using a MCDM technique are necessary. Such steps are described below.

3.1 Multi-objective Problem Formulation

In order to find the weights for a weighted majority voting ensemble, a MOP is defined as follows:

$$\max_{\mathbf{x}} \mathbf{J}(\mathbf{x}) = [J_1(\mathbf{x}), J_2(\mathbf{x})] \quad (5)$$

subject to:

$$0 \leq x_i \leq 1, i \in [1, \dots, n] \quad (6)$$

where the decision variables x_i are the weights of each of the n base classifiers. The objectives $J_1(\mathbf{x})$ and $J_2(\mathbf{x})$ are the true positive ratio (TPR) and the true negative ratio (TNR), respectively. The first measures the proportion of actual positive values (P) that are correctly predicted as such (TP). The last measures the proportion of actual negative values (N) that are correctly predicted as such (TN). Such equations are computed as follows:

$$TPR = TP/P \quad (7)$$

$$TNR = TN/N \quad (8)$$

In order to compute such values, the MOP is performed according to the following procedure:

- (1) Get the n weights from the decision variables $\mathbf{x}$;
- (2) Predict the output on a validation data set for each of the n already trained base classifiers;
- (3) Compute the output of the ensemble using a weighted majority voting scheme;
- (4) Compare the ensemble's output to the actual outputs from the validation data set using a confusion matrix;
- (5) Compute and return the TPR and TNR objectives.

3.2 Multi-objective Optimisation

The multi-objective differential evolution with spherical pruning (spMODE-II) algorithm¹ Reynoso-Meza et al. (2014) is applied in order to extract the Pareto front approximation. Such algorithm has the following characteristics: it is based on the differential evolution (DE) algorithm (Storn and Price, 1997; Das et al., 2016) for offspring generation, which has good convergence and simplicity properties for MOO (Zhou et al., 2011); spherical pruning (Reynoso-Meza et al., 2010) is applied, where the objective space is sectioned using spherical coordinates, and one solution is selected in each spherical sector. Such mechanism promotes diversity in the approximated Pareto front by avoiding overcrowding areas; a global Physical programming (PP) (Messac, 1996) index is used to prune solutions and get a manageable size of the Pareto front approximation. This mechanism improves the pertinence of the solutions and can be used for many-objectives optimisation.

The spMODE-II algorithm is configured with: a population of 20 individuals; 100 generations, or 2000 function evaluations; a crossover rate of 0.5; and a scaling factor: 0.5.

3.3 Multi-criteria Decision Making

In order to select a preferred ensemble from the Pareto front approximation, the global PP (Messac, 1996) index ranks the different non-dominated solutions when given a preference matrix. Thus, a preferable solution can be automatically selected. The preferences are stated in aspiration levels using a matrix $\mathfrak{M}$, depicted in Table 1.

4. EXPERIMENTS

In order to analyse the proposed methodology, experiments are performed on seven different binary classification data sets. Also, the proposed methodology is compared to two other ensemble learning techniques, where each combination of model and data set is performed thirty times. Finally, statistical tests are performed to compare the results. Fig. 1 illustrates the experimental procedures.

4.1 Datasets

In order to test the proposed methodology, seven data sets have been selected from the UC Irvine Machine Learning Repository (Dheeru and Karra Taniskidou, 2017). For all

¹ Toolbox available at:

<https://www.mathworks.com/matlabcentral/fileexchange/47035>

Table 1. Preference matrix $\mathfrak{M}$ for ranking the non-dominated ensembles under the following ranges: highly desirable (HD), desirable (D), tolerable (T) undesirable (U) and highly undesirable (HU).

Preference Matrix											
Objective	$\leftarrow J_i^0$	HD	$\rightarrow \leftarrow J_i^1$	D	$\rightarrow \leftarrow J_i^2$	T	$\rightarrow \leftarrow J_i^3$	U	$\rightarrow \leftarrow J_i^4$	HU	$\rightarrow J_i^5$
$J_1(\mathbf{x})$	1.00		0.99		0.95		0.90		0.80		0.00
$J_2(\mathbf{x})$	1.00		0.99		0.95		0.90		0.80		0.00

Figure 1. The proposed experimental procedure, composed of the majority voting (MV), stacking (ST) and the proposed multi-objective ensemble member combination (MOEMC) models.

data sets, the rows with missing data have been removed, and problems that deal with multiple classes have been transformed to a binary classification problem. The data sets are: Cleveland (Detrano et al., 1989), Haberman, Thyroid, Ionosphere, Sonar, Statlog and Wisonsin (Mangasarian, 1990). Such data sets have 14, 3, 21, 34, 60, 13 and 9 features along with 297, 306, 3772, 351, 208, 270 and

Table 2. Data sets used in this work.

Data set	Number of features	Number of instances
Cleveland's heart disease	14	297
Haberman's survival	3	306
Thyroid disease	21	3772
Ionosphere	34	351
Connectionist bench (sonar)	60	208
Statlog (heart)	13	270
Breast cancer Wisconsin (orig.)	9	283

283 instances, respectively. Table 2 presents the number of instances and features for each problem.

4.2 Models

The proposed MOOD procedure for ensemble member combination is compared to two other ensemble techniques: simple majority voting; and stacking. According to Rokach (2010), the majority voting scheme is also known as the plurality vote or the basic ensemble method. The classification of new data is performed according to the most frequent vote from the base classifiers. The stacking methodology uses a meta-learner in order to induce which base classifiers are reliable and which are not (Rokach, 2010). In this work, a support vector machine (SVM) with linear kernel is used as the meta-learner. All ensemble techniques are composed of the following seven base classifiers, from MATLAB's Statistics and Machine Learning Toolbox: SVM (linear, Gaussian and third order polynomial) (Cortes and Vapnik, 1995); decision tree (DT) (Leo et al., 1984); and k-nearest neighbors (k-NN) (one, three and five neighbours) (Bishop, 2007).

4.3 Statistical Tests

Two statistical tests are performed in order to compare the results of the proposed method to other ensemble techniques, the Friedman's and the Wilcoxon signed-rank tests (Gregory and Foreman, 2009) with a 0.05 significance level. On the one hand, the first one is a non-parametric test used in this work for detecting if there are differences between using the three different ensemble models. On the other hand, the last one is a non-parametric test used for only two samples. In this work, it is used to compare which ensemble technique has a better predictive performance. On both tests, a small p-value indicates the rejection of the null hypothesis that there is no difference among the different groups.

5. RESULTS AND DISCUSSION

The results from the experiments can be found in Table 3. The table presents, for each of the seven data sets, the median accuracy achieved by each ensemble method. The next column presents result for the Friedman's test, where only two (ionosphere and sonar) of the seven data sets do not present statistically significant differences among the three ensembles' accuracies. Finally, the results for the paired Wilcoxon signed-rank test is presented for each combination of ensembles.

In the Cleveland, Thyroid and Statlog data sets, the Friedman's test shows that there is a difference among the three models, while the Wilcoxon signed-ranks test indicates that the results from all classifiers are different. With such, it is possible to conclude that the proposed methodology presents better results than the other two ensemble techniques. For the Haberman data set, while the first non parametric test indicates the difference among the three models, the second test indicates that the majority voting scheme has the best performance among the three models. For the Ionosphere and Sonar data sets, the first test cannot indicate that there is a difference among the three ensemble techniques. Finally, in the Wisconsin data set, the tests indicate that both the stacking model and the proposed methodology present better results than the majority voting scheme. However, there is no indicative of difference for the results of the first two models.

With such results, it is possible to observe that the proposed methodology outperforms the other models on three of the seven data sets. It is also among the best models on other three data sets. However, on one of the data sets, the proposed model is outperformed by the majority voting scheme. Thus, the proposed multi-objective optimisation design procedure for ensemble member combination presents a promising approach for building powerful ensembles.

6. CONCLUSION AND FUTURE RESEARCH

In order to build strong weighted majority voting ensembles, this paper presents a new MOOD procedure for ensemble member combination. First, the proposed method formulates a MOP to adjust the weights of multiple base classifiers. The resulting TPR and TNR of the ensemble is optimised with the spMODE-II algorithm, resulting on a set of several non-dominated ensembles. Finally, a MCDM stage is performed in order to select a preferable ensemble.

In order to validate the proposed methodology, a set of tests has been performed using seven different benchmark problems. The same set of tests is performed on two other ensemble learning methods, in order to compare the results. Finally, statistical tests are used in order to analyse the results.

The results show that the proposed methodology is able to outperform the other two models on three data sets, while being outperformed on only one data set. On other three data sets, the results show no statistical difference. Thus, it is possible to conclude that the application of a MOOD procedure for ensemble member combination is able to compete with classical ensemble learning combination methods, presenting promising results.

Future work on MOOD for ensemble member combination can deal with the application of different MOO algorithms and MCDM techniques. Also, different objectives can be used for generating a Pareto front approximation, such as performance and complexity metrics. Finally, the authors plan to apply the proposed methodology on real-world engineering problems.

ACKNOWLEDGMENT

The breast cancer database was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg.

REFERENCES

- Bishop, C.M. (2007). *Pattern Recognition and Machine Learning*. Springer.
- Coello, C.A.C. and Lamont, G.B. (2004). *Applications of multi-objective evolutionary algorithms*, volume 1. World Scientific.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297.
- Das, S., Mullick, S.S., and Suganthan, P.N. (2016). Recent advances in differential evolution—an updated survey. *Swarm and Evolutionary Computation*, 27, 1–30.
- Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J.J., Sandhu, S., Guppy, K.H., Lee, S., and Froelicher, V. (1989). International application of a new probability algorithm for the diagnosis of coronary artery disease. *American Journal of Cardiology*, 64(5), 304–310.
- Dheeru, D. and Karra Taniskidou, E. (2017). UCI machine learning repository. URL <http://archive.ics.uci.edu/ml>.
- Gregory, W. and Foreman, D. (2009). *Nonparametric statistics for non-statisticians*. John Wiley and Sons.
- Gu, S., Cheng, R., and Jin, Y. (2015). Multi-objective ensemble generation. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(5), 234–245.
- Gu, S. and Jin, Y. (2014). Generating diverse and accurate classifier ensembles using multi-objective optimization. In *IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making (MCDM)*, 9–15. IEEE.
- Gunes, V., Ménard, M., Loonis, P., and Petit-Renaud, S. (2003). Combination, cooperation and selection of classifiers: A state of the art. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(08), 1303–1324. doi:10.1142/S0218001403002897.
- Leo, B., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). *Classification and regression trees*. Wadsworth International Group.
- Mangasarian, O.L. (1990). Cancer diagnosis via linear programming. *SIAM News*, 23(5), 1–18.
- Messac, A. (1996). Physical programming: effective optimization for computational design. *AIAA Journal*, 34(1), 149–158.
- Miettinen, K. (2012). *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media.
- Peimankar, A., Weddell, S.J., Jalal, T., and Laphorn, A.C. (2017). Evolutionary multi-objective fault diagnosis of power transformers. *Swarm and Evolutionary Computation*, 36, 62–75. doi:10.1016/j.swevo.2017.03.005.

Table 3. Results for the three ensemble methods on the seven data sets.

Data set	Ensembles' median accuracy			Friedman's test (p-value)	Wilcoxon signed-ranks test (p-value)			
	ST	MV	MOEMC		ST x MV	ST x MOEMC	MV x MOEMC	
Cleveland	0.610	0.729	0.814	<0.01	<0.01	<0.01	<0.01	
Haberman	0.697	0.738	0.721	<0.01	<0.01	>0.05	<0.05	
Thyroid	0.922	0.942	0.997	<0.01	<0.01	<0.01	<0.01	
Ionosphere	0.886	0.886	0.900	>0.05	>0.05	>0.05	<0.05	
Sonar	0.857	0.857	0.833	>0.05	>0.05	>0.05	>0.05	
Statlog	0.593	0.722	0.815	<0.01	<0.01	<0.01	<0.01	
Wisconsin	0.956	0.876	0.964	<0.01	<0.01	>0.05	<0.01	

- Peimankar, A., Weddell, S.J., Jalal, T., and Lapthorn, A.C. (2018). Multi-objective ensemble forecasting with an application to power transformers. *Applied Soft Computing*, 68, 233–248. doi:10.1016/j.asoc.2018.03.042.
- Reynoso-Meza, G., Blasco, X., Sanchis, J., and Martínez, M. (2014). Controller tuning using evolutionary multi-objective optimisation: current trends and applications. *Control Engineering Practice*, 28, 58–73.
- Reynoso-Meza, G., Sanchis, J., Blasco, X., and Martínez, M. (2010). Design of continuous controllers using a multiobjective differential evolution algorithm with spherical pruning. *Applications of Evolutionary Computation*, 532–541.
- Ribeiro, V.H.A. and Reynoso-Meza, G. (2018). A multi-objective optimization design framework for ensemble generation. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 1882–1885. ACM.
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1), 1–39.
- Rosales-Pérez, A., García, S., Gonzalez, J.A., Coello, C.A.C., and Herrera, F. (2017). An evolutionary multiobjective model and instance selection for support vector machines with pareto-based ensembles. *IEEE Transactions on Evolutionary Computation*, 21(6), 863–877.
- Rosales-Pérez, A., Gonzalez, J.A., Coello, C.A.C., Escalante, H.J., and Reyes-Garcia, C.A. (2014). Multi-objective model type selection. *Neurocomputing*, 146, 83–94.
- Smith, C. and Jin, Y. (2014). Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction. *Neurocomputing*, 143, 302–311.
- Storn, R. and Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4), 341–359.
- Tan, C.J., Lim, C.P., and Cheah, Y.N. (2014). A multi-objective evolutionary algorithm-based ensemble optimizer for feature selection and classification with neural network models. *Neurocomputing*, 125, 217–228. doi:10.1016/j.neucom.2012.12.057.
- Wang, S., Minku, L.L., and Yao, X. (2014). A multi-objective ensemble method for online class imbalance learning. In *2014 International Joint Conference on Neural Networks (IJCNN)*, 3311–3318. IEEE. doi:10.1109/IJCNN.2014.6889545.
- Woźniak, M., Graña, M., and Corchado, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 16, 3–17. doi:10.1016/j.inffus.2013.04.006.
- Zhang, C., Sun, J.H., and Tan, K.C. (2015). Deep Belief Networks Ensemble with Multi-objective Optimization for Failure Diagnosis. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 32–37. IEEE. doi:10.1109/SMC.2015.19.
- Zhao, J., Jiao, L., Xia, S., Basto Fernandes, V., Yevseyeva, I., Zhou, Y., and T.M. Emmerich, M. (2018). Multiobjective sparse ensemble learning by means of evolutionary algorithms. *Decision Support Systems*, 111, 86–100. doi:10.1016/j.dss.2018.05.003.
- Zhou, A., Qu, B.Y., Li, H., Zhao, S.Z., Suganthan, P.N., and Zhang, Q. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1), 32–49.