Indonesian Journal of Electrical Engineering and Computer Science
Vol. 12, No. 3, December 2018, pp. 1063~1070
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v12.i3.pp1063-1070

Adaptive Data Structure Based Oversampling Algorithm for Ordinal Classification

D. Dhanalakshmi[1], Anna Saro Vijendran[2]
[1]Department of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore, India
[2]School of Computing, Sri Ramakrishna College of Arts and Science, Coimbatore, India

Article Info

Article history:

Received May 1, 2018
Revised Jun 25, 2018
Accepted Aug 21, 2018

Keywords:

Adaptive data structur
Average accuracy
Extreme learning machine for ordinal regression
Geometric mean
Kappa
Maximum mean absolute error
Multi class ordinal classification

ABSTRACT

The main objective of this research is to improve the predictive accuracy of classification in ordinal multiclass imbalanced scenario. The methodology attempts to uplift the classifier performance through synthesizing sophisticated objects of immature classes. A novel Adaptive Data Structure based Oversampling algorithm is proposed to create synthetic objects and Extreme Learning Machine for Ordinal Regression (ELMOP) classifier is adopted to validate our work. The proposed method generating new objects by analyzing the characteristics and intricacy of immature class objects. On the whole, the data set is divided into training and test data. Training data set is updated with new synthetic objects. The experimental analysis is performed on testing data set to check the efficiency of the proposed methodology by comparing it with the existing work. The performance evaluation is conducted in terms of the parameters called Mean Absolute Error, Maximum Mean Absolute Error, Geometric Mean, Kappa and Average Accuracy. The measures prove that the proposed methodology can produce authentic synthetic objects than the existing techniques. The Proposed technique can synthesize the new effective objects through evaluating the structure of immature class. It boosts the global precision and class wise precision especially preserves rank order of the classes.

Corresponding Author:

D. Dhanalakshmi,
Department of Computer Science,
Sri Ramakrishna College of Arts and Science,
Coimbatore, India.
Email: dhanadurairaj@gmail.com

1. INTRODUCTION

1.1. Background

Typical classification algorithms well behaved with appropriately balanced datasets, but many real-world applications in various disciplines exhibit imbalanced ordinal nature such as Disease prediction, Weather forecasting, Performance prediction, Rating and Financial investment etc., Most of the classification algorithms try to achieve the better performance globally. They suffer to obtain better local as well as global performance due to skewed ordinal nature of data. The main research contributions in this perspective include Algorithmic level, data level and cost sensitive approaches [1], [2] Proposed variation of smote algorithm SNOCC for imbalanced binary class which recognizes more than two seed samples to create new samples in the interior region of the bordered seed samples to simulate the even and uneven distribution of original samples [3]. Tested various classifiers performance based on cost for two class imbalanced public health dataset problem. They concluded that Bayesian classifiers work well for this problem [4]. Proposed Online version of Imbalanced SVM (OISVM) for binary email classification to improve processing speed and save storage space [5]. Proved ELM is efficient algorithm for classification problem [6]. Confirmed that, Euclidean and City block distance measures performed well in K-Nearest Neighbour Algorithm [7].

Proposed first oversampling algorithm to handle imbalanced multiclass ordinal classification problem [8]. Authors performed literature review in the context of ordinal classification to find the causes for classifier performance degradation [9], [10]. Authors proposed proposed collinear based oversampling algorithm in the safe and border line region for ordinal classification [11]. Introduced unsupervised oversampling method for ordinal regression [12]. Adopt data characteistics to identify complex objects and decide size to synthesize for such each problematic for learning as well as most responsible for performance degradation objects [13] Adopted clustering to group minority instances and synthesizing objects based on the learning complexity of the group [14] Make use of data characteristics for oversampling objects and concluded that insight the formation and group of objects in dataset collusion the prepossessing algorithms which uplift the performance. Authors [15] suggested that, synthesizing more objects in the borderline which expand the probable area of immature class.

1.2. Problem Indentified from Literature Review

Most of the existing work focuses on imbalanced binary class classification, imbalanced multiclass classification and ordinal classification alone. Very few research works has been carried out for imbalanced multiclass ordinal classification. Here we deal such a complicated situation Imbalanced Multiclass Ordinal Classification. Many real time applications in various domains such as Economy, Automobile Industry, Medicine, Agriculture, BioMedicine, Human Resource Development etc., consists data in imbalanced multiclass ordinal nature. They demand effective state-of-the-art solution to tackle this scenario for improvement of predictive accuracy and minimization of error rate.

1.3. Proposed Solution

In this paper, we propose an Adaptive Data Structure Based Oversampling Algorithm. In our proposed adaptive data structure based oversampling algorithm differs from the existing work [7] it prefers only the complicated objects, compare with [12] it handles multiclass ordinal classification, deviated with [14] our algorithm handles multiclass ordinal classification and works on each patterns of minority class to analyses complexity.

2. ADAPTIVE DATA STRUCTURE BASED OVERSAMPLING ALGORITHM

2.1. Selecting Immature Class

The dataset is partitioned into training and testing group. [7], [16], [17] Different works in the literature have considered the immature classes that exhibit IR value above than 1.5.

$$\mathrm{IR}_q = \frac{\sum_{j \neq q} \mathrm{N}_j}{\mathrm{Q}.\mathrm{N}_q} \tag{1}$$

In this work using (1), IR valued is calculated for each class. Based on that value, the number of synthetic patterns to be generated for each class is calculated through (2). After (2), again IR value is calculated for each class. When (2) alters the IR value for the remaining classes as above than 1.5, new synthetic objects are created for such classes until reaches the IR new (3) value below 1.5.

$$Syn_q = \frac{\sum_{j \neq q} N_j + \left(\left(\sum_{c=1}^{c=Q} Syn_c\right) - Syn_q\right)}{(Threshold\,.Q) - N_q} \tag{2}$$

$$IR(new)_q = \frac{\sum_{j \neq q} N_j + \left(\left(\sum_{c=1}^{c=Q} Syn_c\right) - Syn_q\right)}{(N_q + Syn_q).Q} \tag{3}$$

2.2. Adapting Data Structure to Oversample

In immature class, 5 Nearest Neighbors for each object is calculated and classify each object into secure objects and insecure objects based on nearest neighbors.

$$r_j = \frac{\Delta_j}{K},\ \mathrm{j} = 1 \ldots \mathrm{Im}\,\mathrm{C_s} \tag{4}$$

Among 5 Nearest Neighbors, 4 or 5 neighbors belong to immature class it considered as safe objects, 2 or 3 means it is called as borderline object, 1 means it is rare object and for 0 it is outlier [14]. In our proposed work we treated above said object categories into 2 groups such as secure objects and insecure objects. Safe object is secure object remaining categories belong to insecure objects. For ordinal classification scenario, adjacent classes are close to each other [7]. Consideration of the above statement, for the secure object, 1 nearest neighbor of the non-immature class object belong to one of the adjacent class that safe objects is intended as ordinal borderline object. Ordinal borderline objects and insecure objects are captured for further processing. The new immature class (ImC) consists the above said two groups.

ImC={Ordinal borderline objects, insecure objects} (5)

$$\text{ImC}=\{x_1, x_2 \ldots x_j\} \tag{6}$$

2.3. Finding Adjacent Classes and Synthesizing Objects

Adjacent class patterns are very close to immature class patterns compare with nonadjacent class patterns [7]. According to that, this work finds the shortest distance for each object of both adjacent classes through immature class objects.

2.4. Graph Construction

Create graph for the immature class, q be the index of the class we want to over-sample. Create graph G_q for class $\mathrm{Im}\,\mathrm{C}_\mathrm{q}$ based on three sub graphs $G_{q-1,q}$, $G_{q,q}$ and $G_{q,q+1}$

a. Construct $G_{q-1,q}$

For every pattern in q[th] class, find its k-nearest neighbor in the q-1[th] class using the formula $N_d(X_q, X_{q-1}, k)$. Create edges.

For every pattern in q-1[th] class, find its k-nearest neighbor in the q[th] class using the formula $N_d(X_{q-1}, X_q, k)$. Create edges.

b. Construct graph $G_{q-1,q}$ with edges only those are common in:

$$N_d(X_{q-1}, X_q, k) \cap N_d(X_q, X_{q-1}, k)$$

c. Construct $G_{q,q}$

For every pattern in q[th] class, find its k nearest neighbours in the q[th] class and create edges with these neighbours.

d. Construct $G_{q,q+1}$ same like $G_{q-1,q}$

e. Find the shortest path from $G_{q-1,q}$ to $G_{q,q+1}$ via $G_{q,q}$ using Dijkstra's algorithm for each vertex in $G_{q-1,q}$

f. Select an edge from, and based on oversampling rate that should be one of the shortest path edge.

3. FURTHER CONCERNS

To clarify all the works which are done in the previous subsection, a summary of the work is given:

3.1. Pseudo Code for the Proposed ADSOS

Input: Training Dataset
Output: New Balanced Trained dataset

Phase I

1) Select the immature class to be oversampled (im) based on IR value, calculated using equation (1).
2) The number of objects to be synthesized is calculated using equation (2).
3) The new IR value is calculated using (3). Until IR value for all the classes reach less than 1.5 repeat step b to step c.

Phase II

Find the structure of the immature class using equation (4)
Secure Ordinal objects and insecure objects are deliberated for oversampling.

Phase III

Construct graph $G_{q-1,q}$, $G_{q,q}$ and $G_{q,q+1}$ as mentioned in the above section

Find shortest path from $G_{q-1,q}$ to $G_{q,q+1}$ via $G_{q,q}$

Randomly select an edge from $G_{q-1,q}$, $G_{q,q+1}$ and $G_{q,q}$

Phase IV

Synthesizing new objects

Selected edge belongs to $G_{q,q}$ apply uniform distribution for synthesizing objects

Selected edge belongs to $G_{q-1,q}$ or $G_{q,q+1}$ applies gamma distribution for synthesizing objects

Phase V

Prediction using ordinal classifier

Input: New Balanced Trained Dataset, Test dataset

Output: Predicted value

4. RESULTS AND ANALYSIS

To validate the proposed methodology some data sets Wisconsin, housing, machine, triazines, auto are derived from [18]. The rest of the datasets are extracted from UCI. Table 1 shows the description of dataset. Initially, these datasets do not represent ordinal classification, but it represents regression. To evolve this regression into ordinal classification we have considered the desired result is categorized into five or ten classes with equal frequency. Regarding the experimental setup, a holdout stratified technique was applied to divide the datasets 10 times, using 75 percent of the patterns for training and the remaining 25 percent for testing. Finally, the results are taken as the mean and standard deviation of the measures over the 10 test sets.

Table 1. Nature of Dataset

Dataset	Total no. of patterns	No. of Attributes	Total no. of classes	IR value per class
bondrate	57	37	4	**1.85**,0.19,.0.92,**2.38**
Auto	392	7	5	0.65,0.40,0.58,1.14,**7.15**
automobile	205	71	6	**12.58**,1.43,0.33,0.47,0.90,1.11
Car	1728	21	4	0.11,0.88,**5.98,6.36**
ERA1vs2345vs7vs8vs9	1000	4	5	**1.97**,0.06,**2.07,6.32,10.51**
Eucalyptus123vs4vs5	736	91	3	0.25,0.82,**2.00**
machine5	209	6	5	0.07,1.36,**2.92**,**6.04**,**4.26**
machine10	209	6	10	0.08,0.46,0.94,**3.02**,**2.50,3.80,7.70, 5.10,5.10,3.80**
triazines5	186	60	5	**5.36,3.27**,1.26,.023,0.46
wisconsin5	194	32	5	0.38,0.74,0.71,1.14,**1.87**
wisconsin10	194	32	10	0.31,0.81,0.59,1.35,0.71,1.02,1.35, **1.71,1.97,1.97**
housing5	506	13	5	1.11,0.22,0.62,**2.61,3.10**
Toy	300	2	5	1.53,0.49,0.56,0.68,**1.68**
SWD	1000	4	9	**7.56,**0.46,0.38,0.90

The Entire work is validated based on Adaptive Data Structure based Oversampling algorithm with ELMOP Classifier and these results are compared Graph based oversampling algorithm with ELMOP.

4.1. Performance Measures

This work preferred most relevant performance measures such as Mean Absolute Error (MAE), Maximum Mean Absolute Error (MMAE), Geometric Mean (GM), Cohen's Kappa and Accuracy used to validate the proposed work.

4.1.1. Mean Absolute Error

MAE is the average difference between true value and evaluated value. MAE is the essential and clear measure of average error [19].

$$MAE_q = \frac{1}{N_q} \sum_{i=1}^{N_q} \left| O(y_i) - O(\hat{y}_i) \right| \quad (7)$$

4.1.2. Maximum Mean Absolute Error

Proposed MMAE metric for ordinal classification. It displays the maximum MAE for all the classes [20].

$$\text{MMAE}=\max\{MAE_q; q \in \{1,....,Q\}\} \tag{8}$$

4.1.3. Geometric Mean

Geometric mean is one of the preferable measures for imbalanced learning. Geometric mean is defined as follows:

$$\text{GMean}=\sqrt{\frac{TP}{TP+FN}\times\frac{TN}{TN+FP}} \tag{9}$$

4.1.4. Kappa

Cohen's kappa statistic is one of the preferable measures for imbalanced multi class learning. When kappa value < 0 is indicating no coexists between actual and predicted value , 0–0.20 as slight coexists, 0.21–0.40 as fair coincide, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement.

$$\text{Kappa}=\frac{N\sum_{i=1}^{i=m}TP-\sum_{i=1}^{i=m}T_{ri}\,T_{ci}}{N^2-\sum_{i=1}^{i=m}T_{ri}\,T_{ci}} \tag{10}$$

Where N total number of patterns, T_{ri} number of rows from the confusion matrix, T_{ci} number of columns from the confusion matrix.

4.1.5. Accuracy

Accuracy is the proportion of true results, either true positive or true negative.

$$\text{Accuracy}=(TP+TN) / (TN+TP+FN+FP) \tag{11}$$

After evaluating the measures MAE, MMAE, GMean, Kappa and Accuracy Obtained Over 10 Runs for the Existing Graph Based Oversampling and Proposed ADSOS results are displayed in Table 2-Table 6.

Table 2. MAE Mean and Standard Deviations (Mean±SD)

Dataset/ Method	Graph Based Oversampling	ADSOS
bondrate	0.1733±0.1842	**0.0933±0.0326**
auto	0.3409±0.0354	**0.3407±0.0353**
automobile	0.4577±0.2724	**0.3942±0.2469**
car	0.2250±0.0200	**0.2240±0.0201**
ERA1vs2345vs7vs8vs9	0.1404±0.0440	**0.1272±0.0442**
Eucalyptus123vs4vs5	0.3097±0.0556	**0.2445±0.0730**
machine5	**0.2641±0.0462**	0.2735±0.1313
machine10	0.5849±0.1736	**0.5219±0.0978**
triazines5	0.3688±0.0265	**0.3680±0.0269**
wisconsin5	**0.3605±0.0347**	0.4489±0.0670
wisconsin10	1.1360±0.0347	**1.0271±0.0687**
housing5	0.1548±0.0427	**0.2078±0.0823**
Toy	0.5688±0.0349	**0.5466±0.0458**
SWD	0.2800±0.0097	**0.2093±0.0732**

Table 3. MMAE Mean and Standard Deviations (Mean±SD)

Dataset/ Method	Graph Based Oversampling	ADSOS
bondrate	1.0500±0.1500	**1.0000±0.0000**
auto	**0.9613±1.0029**	1.0351±0.0262

automobile	**2.2941±0.4624**	2.3000±0.4582
car	0.3207±0.0286	**0.3042±0.0286**
Dataset/ Method	Graph Based Oversampling	ADSOS
ERA1vs2345vs7vs8vs9	1.1043±0.1169	**1.0738±0.0953**
Eucalyptus123vs4vs5	0.5467±0.1000	**0.4310±0.1281**
machine5	**0.3684±0.0644**	0.3815±0.1831
machine10	1.0689±0.3172	**0.9539±0.1788**
triazines5	3.0000±0.0000	3.0000±0.0000
wisconsin5	**1.0429±0.1050**	1.1656±0.0979
wisconsin10	3.1161±0.0357	**2.7770±0.1745**
housing5	0.6140±0.3108	0.8526±0.2375
Toy	2.1574±0.1118	**2.0595±0.2284**
SWD	**1.4583±0.1559**	1.5833±0.1381

Table 4. GM Mean and Standard Deviations (Mean±SD)

Dataset/ Method	Graph Based Oversampling	ADSOS
bondrate	0.8403±0.0733	**0.8679±0.0186**
auto	0.8331±0.1224	0.8331±0.0092
automobile	0.7419±0.0476	**0.7433±0.0486**
car	0.9652±0.0030	0.9652±0.0030
ERA1vs2345vs7vs8vs9	0.8409±0.0290	**0.8536±0.0051**
Eucalyptus123vs4vs5	0.8357±0.0300	**0.8705±0.0385**
machine5	0.9657±0.0081	0.9609±0.0179
machine10	0.9753±0.0066	0.9732±0.0167
triazines5	0.6043±0.0012	0.6029±0.0029
wisconsin5	0.8535±0.0079	0.8245±0.0225
wisconsin10	0.7594±0.0764	**0.8164±0.0311**
housing5	0.9023±0.03547	0.8643±0.0419
Toy	0.7258±0.0038	**0.7289±0.0091**
SWD	0.7227±0.0084	**0.7708±0.0492**

Table 5. Kappa Mean and Standard Deviations (Mean±SD)

Dataset/ Method	Graph Based Oversampling	ADSOS
bondrate	0.6410±0.2142	**0.7205±0.0892**
auto	0.0819±0.0489	**0.0815±0.0488**
automobile	**0.3784±0.0387**	0.3670±0.0409
car	0.7000±0.0267	0.7000±0.0267
ERA1vs2345vs7vs8vs9	0.5912±0.1063	**0.6437±0.0908**
Eucalyptus123vs4vs5	0.3030±0.1251	**0.4497±0.1643**
machine5	**0.5676±0.1002**	0.5103±0.2253
machine10	0.2109±0.1284	**0.4390±0.0480**
triazines5	0.2464±0.0313	**0.2553±0.0266**
wisconsin5	0.0584±0.0271	**0.1762±0.0766**
wisconsin10	**0.6218±0.0075**	0.6026±0.0342
housing5	**0.5406±0.1142**	0.4248±0.1477
Toy	**0.2252±0.0204**	0.2167±0.0351
SWD	0.2924±0.0362	**0.4915±0.2051**

Table 6. Accuracy Mean and Standard Deviations (Mean±SD)

Dataset/ Method	Graph Based Oversampling	ADSOS
bondrate	0.8333±0.1693	**0.9066±0.0326**
auto	0.7010±0.0223	**0.7030±0.0234**
automobile	0.6750±0.1445	**0.7038±0.1392**
car	0.8875±0.0100	0.8875±0.0100
ERA1vs2345vs7vs8vs9	0.8692±0.0340	**0.8860±0.0290**
Eucalyptus123vs4vs5	0.6902±0.0556	**0.8705±0.0730**
machine5	**0.8616±0.0320**	0.8434±0.0721
machine10	**0.8302±0.0462**	0.8176±0.1114
triazines5	0.7589±0.0100	**0.7617±0.0085**
wisconsin5	**0.6599±0.0096**	0.6081±0.0363
wisconsin10	0.5238±0.0096	**0.5442±0.0372**
housing5	**0.8530±0.0365**	0.7999±0.0813
Toy	0.5866±0.0109	**0.5906±0.0179**
SWD	0.7346±0.0135	**0.8093±0.0769**

To quantify whether a statistical difference exists among the algorithms compared, t-Test is performed on the mean ranking of all the evaluation measures it is displayed in Table 7.

Table 7. t-Test on Mean Ranking of the Evaluation Measures (α=0.05)

Observations	Graph Based Oversampling	ADSOS
Mean	1.538	1.238
Variance	0.03727	0.00637
Observations	5	5
Pearson Correlation	-0.665558527	
Hypothesized Mean Difference	0	
df	4	
t Stat	2.648548478	
P(T<=t) one-tail	0.028534115	
t Critical one-tail	2.131846786	
P(T<=t) two-tail	0.057068231	
t Critical two-tail	2.776445105	

The test proves that, null hypothesis is rejected where p value is less than (α=0.05) that two algorithms performs similarly in mean ranking of the evaluation measures however ADSOS performs better than Graph Based Oversampling Method with ELMOP as classifier which is depicted in Figure 1.

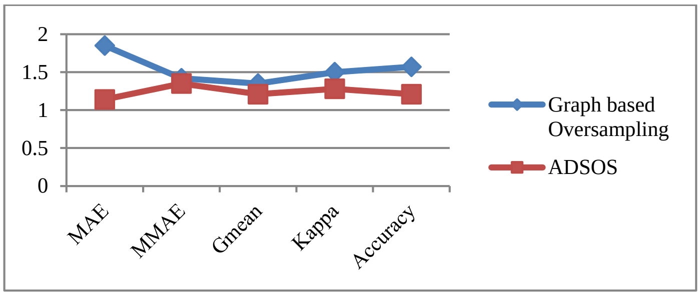

Figure 1. Mean ranking of the evaluation measures

5. CONCLUSION

In this paper we proposed the novel Adaptive Data Structure based Oversampling algorithm to prefer the useful objects for further processing. We compared our methods with existing graph based preprocessing algorithm for fourteen data sets. Our aim of this work is to compare the proposed ADSOS preprocessing algorithm with existing preprocessing algorithm for ordinal classification. With regards, we adopt any one of the ordinal classifier such as ELMOP to validate our work. Thus our proposed method only oversamples objects which have highest confidence and complicated regions. Experiments indicate that our method behaves better in terms of error rate, accuracy sensitivity.

ACKNOWLEDGEMENTS

I would like to express my special gratitude and thanks to Dr.K.Karunakaran, Principal and Secretary, Sri Ramakrishna College of Arts and Science, Coimbatore for providing excellent infrastructure and support for my Research work. I am highly indebted to my Research Guide Dr.Anna Saro Vijendran, Dean School of Computing, Sri Ramakrishna College of Arts and Science, Coimbatore for their guidance, constant supervision, support, effort, invaluably constructive criticism and friendly advice for my research work.

REFERENCES

[1] Qiong, G., Cai, Z., Zhu, L. and Huang, B., *Data Mining on Imbalanced Data Sets*, International Conference on Advanced Computer Theory and Engineering, Phuket, 20-22 December, 2008;1020-1024.
[2] Zhuoyuan Zheng , Yunpeng Cai, Ye Li, Over Sampling Method For Imbalanced Classification, *Computing and Informatics*, 2015; 34:1017–1037.
[3] Rohini R. Rao, Krishnamoorthi Makkithaya, Learning from a Class Imbalanced Public Health Dataset: a Cost-based Comparison of Classifier Performance, *International Journal of Electrical and Computer Engineering (IJECE)*, 2017;7(4): . 2215-2222.

[4] XiaoQing Gu, TongGuang Ni, Wei Wang, Online Imbalanced Support Vector Machine for Phishing Emails Filtering, *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 2014;12(6): 4306 – 4313.

[5] Xiaozhi Du, Huimin Lu, Gang Liu, Software Aging Prediction based on Extreme Learning Machine, *TELKOMNIKA*,2013;11(11): 6547-6555.

[6] Mohd Abdul Talib Mat Yusoh, Saidatul Habsah Asman, Zuhaila Mat Yasin, Ahmad Farid Abidin, Classification of The NTEV Problems on the Commercial Building, *Indonesian Journal of Electrical Engineering and Computer Science*,2018;9(2): 380-386.

[7] Marıa Perez-Ortiz, Pedro Antonio Gutierrez,Cesar Hervas-Martınez and Xin Yao, Graph Based Approaches for Over-sampling in the Context of Ordinal Regression, *IEEE Transactions on Knowledge and Data Engineering*, 2015; 27(5):1233-1245.

[8] Anna Saro Vijendran, D. Dhanalakshmi, An analysis of Data Characteristics and Classifier Performance in the Context of Ordinal Classification, *International journal of Data mining and Emerging Technologies*, 2016; 7(1): 36-42.

[9] Anna Saro Vijendran, D. Dhanalakshmi, A Novel Approach in Oversampling Algorithm for Imbalanced Data Sets in the context of Ordinal Classification, *IEEE international Conference on Computational Intelligence and Computing Research*, Chennai,2016; 964-968.

[10] Anna Saro Vijendran, D. Dhanalakshmi , An Improved Sampling Algorithm for Imbalanced Data Sets in the context of Ordinal Classification, International journal of Advance Research in Science and Engineering, 2017; 6(12):201-209.

[11] Iman Nekooeimehr, Susana K. Lai-Yuen, Cluster-based Weighted Oversampling for Ordinal Regression (CWOS-Ord), *Journal Neurocomputing*, 2016; 218(C):51-60.

[12] H. He, Y. Bai, E. Garcia, S. Li, ADASYN, *Adaptive synthetic sampling approach for imbalanced learning*", in: Proceedings of the IEEE International Joint Conference on Neural Networks, IJCNN, IEEE World Congress on Computational Intelligence, 2008;1322-1328.

[13] Iman Nekooeimehr, Susana K, Adaptive semi-unsupervised weighted oversampling (A-SUWO) for imbalanced datasets, *Expert Systems with Applications*, 2016; 46:405-416.

[14] Jose A.Saez, Bartosz Krawczyk, Michal Wozniak, Analysisng the oversampling of different classes and types of examples in multi-class imbalanced dataset, *Pattern Recognition*, 2016; 57:164-178.

[15] H. Han, W.Y. Wang, and B.H. Mao, *BorderlineSMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning*, In Proceedings of the International Conference on Intelligent Computing 2005; Part I, LNCS 3644, 878–887.

[16] F. Fernandez-Navarro, C. Hervas-Martınez, and P. A. Gutierrez, A Dynamic Over-Sampling Procedure Based on Sensitivity for Multi-class Problems, *Pattern Recognit.*, 2011; 44(8):1821– 1833.

[17] Fernández A., del Jesus M.J., Herrera F., *Multi-class Imbalanced Data-Sets with Linguistic Fuzzy Rule Based Classification Systems Based on Pairwise Learning*, In: Hüllermeier E., Kruse R., Hoffmann F. (eds) Computational Intelligence for Knowledge-Based Systems Design. IPMU, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg 2010; 6178.

[18] W. Chu and Z. Ghahramani,Gaussian processes for ordinal regression,*J. Mach. Learn. Res.*, 2005; 6:1019–1041.

[19] Cort J. Willmott, Kenji Matsuura, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Research*,2005; 30: 79–82.

[20] M. Cruz-Ramırez, C. Hervas-Martınez, J. Sanchez-Monedero, and P. A.Gutierrez, Metrics to guide a multi-objective evolutionary algorithm for ordinal classification,*Neurocomputing*, 2014; 135:21–31.