

A Novel Technique for Multi-Class Ordinal Regression-APDC

L. Mary Gladence^{1*}, M. Karthi² and T. Ravi³

¹Sathyabama University, Rajiv Gandhi Road, Jeppiaar Nagar, Chennai - 600119, Tamil Nadu, India; lgladence@gmail.com,

²Karapagam College of Engineering, Myleripalayam Village, Othakkal Mandapam Post, Coimbatore - 641032, Tamil Nadu, India; karthi.m@kce.ac.in

³Madanapalle Institute of Technology and Science, P.B. No:14, Kadiri Road, Angallu Village, Chittoor District, Madanapalle - 517325, Andhra Pradesh, India; travi675@yahoo.com

Abstract

Objectives: Ordinal regression is one which is used in Multiclass classification where there is an essential ordering among the classes. The training dataset is initially classified depending on the Random threshold values θ . Based on these values, the distance between the different class labels are predicted by one against one technique. **Method:** All Pairs Distance Calculation using one against one technique [APDC_1 AG 1] is Proposed to validate the work. But in the referred previous work, distance is calculated using adjacent classes, but here all pairs distance calculation is used to find the class label distance to all class label pairs. **Findings:** On the whole, New trained data are in the form of one dimensional representation. Here, with the knowledge of proposed work, testing data is tested with New trained data set and the results are produced. The Proposed method is seen to be ambitious when compared with previous work. Beside this, an additional set of experiments is done to study the potential quantifiability and illustratability of the proposed method when using APDC as base methodology. **Improvements:** Proposed work is analyzed with Kernel discriminant analysis, Logistic Regression, Classification via Regression, Multiclass Classifier and found APDC has attained better results according to all measures.

Keywords: All Pairs Distance Calculations, Hyper Line, Latent Space Representation, Multi-Class Ordinal Regression, One Against One Method, Ordinal Classification, Projection

1. Introduction

In, Nominal Classification features which are similar are framed under same class and it will give us the output as either “Yes” or “No”. By notifying in this manner, misclassification may occur, especially in disease prediction. To avoid this disadvantage present in Nominal Classification we are moving on to Ordinal classification. In Ordinal Classification, Ordinal class labels are assigned and object is grouped in a more convenient way, i.e. it is categorized in another class. We observe this type more in disease prediction instead of getting the result as “Yes” or “No” class label we can get it as “Severe”, “Moderate”, “slight”. So Ordinal classification avoids inconsistent data comparing to Nominal Classification.

Classification concept has been referred from¹⁻²¹. Some of the techniques used in classification are Decision Tree, Bayesian Classification, Neural Networks, Support Vector Machine, Boosting etc. Whenever the Regression problem is mapped to a real data we observe that the Classification problem are getting mapped to an unordered set of classes.

Ordinal regression involves the process which is used in the prediction of class labels. Instead of predicting the categorical class label, the relationship between the predictor variable and response variable is predicted using Ordinal Regression. This paper is mainly based on the projection of data in the latent space where each class has different order hypotheses. In order to perform classification, the distance between the class labels are predicted by one against one technique (i.e.) the class labels

* Author for correspondence

distances are calculated from class (k) to other classes in the sample space and from that values the minimum distance is measured. Using these differences weights is calculated. Here all works are done in the training phase itself. The key idea of this paper is to perform the projection model as directed by class distribution obtained from the distance measures using one against one technique. The proposed work is analyzed with following standard methods such as kernel discriminant analysis, support vector machines, Logistic regression, Classification via Regression, Multiclass Classifier, J48. Wojciech Kotłowski et al.² proposed an ordinal classification along with monotonicity limitation. It varies from classification by the following factors such as Knowing background knowledge about ordering classes, ordered attributes, monotonic constraints between an object on the attributes and its class description. We can even define this by using another way called class label, i.e. Whenever input variables increases output variable should not decrease. Non-Parametric approach is one of the most useful approaches to express the Monotonicity constraints among all the existing approaches. Monotonicity constraints assumption alone is referred in this work. Here an analysis of nonparametric approach is done using statistical point of view. Shereen Fouad et al. [4] proposed Learning Using privileged Information (LUPI). Supervised learning is improved in the presence of privileged information. Here, information is available only in the training phase and not in the testing phase. This Novel Learning Methodology is expressed to incorporate privileged information in ordinal classification tasks, whenever natural order is followed. By changing the global metric in the input space, based on distance relations revealed by the privileged information is used to justify this work. Experiments show that by integrating privileged information via the proposed ordinal-based metric learning can upgrade the ordinal classification performance. M. Perez-Ortiz et al. [8], developed a method for reducing the dependency between classifiers. This method can be used as an added core to maximize the diversity, develop the essential combination of rules. This method can be used with any threshold value as a base classifier. Although, there are many classifiers due to many classes each single model differs from the remaining ones as ordinal ranks are used, i.e. differentiating each and every class with one another. This work, based on disintegrating ordinal regression problem into Simpler classification tasks in which order of each and every detail is expressed. Here, for a I class ordinal regression problem,

two binary classification problem and I-2 Ordinal ones are derived. In [12] In order to protect the data publicly available on social network a cluster based anonymization technique is developed to preserve the privacy of data by considering its utility and the knowledge of usage.

Ordinal regression has wide applications in several domains, especially in human evolution. Support Vector Machine concept is used in most of the ordinal regression methods, but it has the major disadvantage of ignoring the overall information about the data and its high computational complexity. To avoid these problem Proposed methods, consider entire data points using All Pairs Distance Calculation with One against One Method. After finding all these details, these are checked with Kernel Discriminant Learning, Support Vector Regression, Logistic Regression.

Previous work performed regression based on implicit values of each data value. Here considered the explicit values of each data value by considering the distance between the each pair using one against one method.

In our work we are going to map the input space to already defined class set called $c=\{c_1, c_2, \dots, c_l\}$. Where the order of class is in the form of $c_1 < c_2 < \dots < c_l$. eg., Severe < Moderate < slight. For example, if we are going to predict heart disease instead of predicting whether the particular person is having heart disease or not in the form of "Yes" or "No" particular person will come to know the category of disease. Based on this, he can go for treatment.

Kernel Discriminant Learning is one of the spearhead learning technique in the Machine Learning Paradigm and it has been used for Supervised Dimensionality Reduction. Kernel Discriminant Analysis has also been suitable for Ordianl Classification by the impressive constraint on the projection to be computed. Due to this it will maintain in its original or existing state and take the advantage of the ordinal information from the different classes.

Support Vector Regression uses powerful function based on statistical learning theory. It is extremely reliable and provide excellent generalization performance even though input has a complex relationship.

Projection is an important scaling factor for the ordinal regression problem. In order to handle this project the class label details with All Pairs Distance Calculation using One Against One[APDC_1 AG 1] technique. Heart Disease, automobile, Bond rate, Contact lenses, Thyroid, Pasture data sets was extracted from the uci repository to evaluate the correctness of proposed work.

The objective of this work is to reduce misclassification

error by explicitly understand the data in the dataset during the training itself. So, testing data space (i.e.) input space is mapped to already trained dataset value of considering the behaviors of that particular data set. For example, if we are going to classify the heart disease related data, then the behavior of that particular dataset is explicitly known instead of direct method using one against all method. So if test data is applied, our algorithm will understand it explicitly and reduce the misclassification error.

Multi-Class Classification: Multi-Class Classification is the process of classifying instances into more than two classes. Some of the Multi-Class Classification Techniques are Point-wise learning, pairwise learning, List-wise learning. Using Point-wise learning particular disease is found for the given data. In pairwise learning relevance ordering relationship is found. List-wise learning directly optimizes the ranking metric for each input. Proposed work uses the concept of pairwise learning approach such as one against one method.

2. All Pairs Distance Calculation using One Against One [APDC_1 AG 1]

Initially the untrained dataset is trained based on All Pairs Distance Calculation using One Against One[APDC_1 AG1]. "Figure 1" represents the architecture of the entire work. Initially training data set is represented in the form of two dimensional view. After representing it in the two dimensional view, data is grouped in the form of classes. Each class is separated in the form of hyperline based on threshold values (threshold value is user defined). Suppose we are considering three threshold values, there will be three hyper lines. Class Separation is calculated using equation (1). Based on the above calculation hyperlines are adjusted. To conclude hyperline into final, i.e. to find the exact class, it should go through three steps. They are minimum distance calculation, weight calculation and relevance calculation. Through weight calculation similarity and dissimilarity data are found. Since the above procedure, data point which has got similarity weight will be in the same class and which are found dissimilar are adjusted into another class. This full work is done in the training phase. Finally trained data sets (evolved using proposed method) represented by the latent space model and in the testing phase the various regression methods

like Kernel Discriminant Analysis, Logistic Regression, Classification via regression, etc are used.

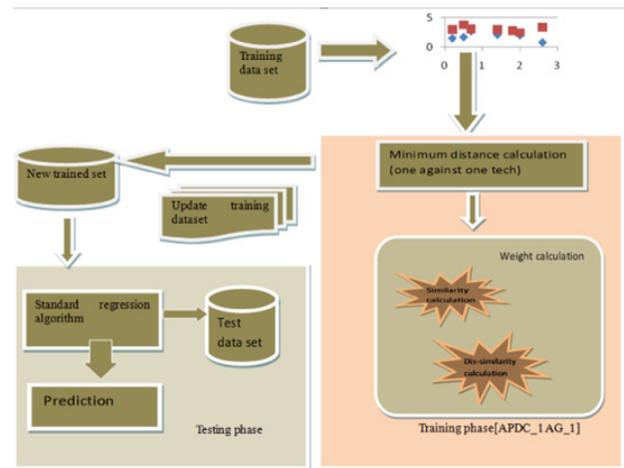


Figure 1. Architecture of APDC_1 AG 1 Method.

2.1 Latent Space Model

Latent Space Model is a two dimensional linear regression. The model proposed here is used to find exact class. For example, initially we need to split the data using hyperline (here, hyperline is used to separate the classes). In this preliminary stage based on user defined threshold value hyperlines are randomly assigned. Each data point is plotted based on the threshold values using equation (1)

$$f(x, \theta) = \begin{cases} c_1 & \text{if } p(x) \leq \theta_1 \\ c_2 & \text{if } \theta_1 p(x) \leq \theta_2 \\ \cdot \\ \cdot \\ \cdot \\ c_l & \text{if } p(x) > \theta_{l-1} \end{cases} \quad (1)$$

Where $\theta = (\theta_1, \theta_2 \dots \dots \theta_{l-1})$, c represents a class label, θ represent threshold value. The function $f(x, \theta)$ is used to relate the variable x with particular threshold value. Using this threshold value classes are separated. K is the Number of classes. C_l denotes class where $l=1,2,\dots,k$. The function $p(x)$ is used to project the data in the sample space.

2.2 Minimum Distance Calculation

The minimum distance is calculated by well defined Euclidean distance formula. Here, to find the minimum

distance one against one technique is applied. For example, assume there are 5 classes, from these five classes if we want to find second class data, minimum distance calculation should be done between class 2 to class 1, class 2 to class 3, class 2 to class 4, class 2 to class 5. Using equation (2) Minimum Distance is calculated.

$$R(x_i^{-1} - x_j^{-1}) = x_j^{(l-1)} \min |[c(x)]_i^l - c(x_j^{(l-1)})| \quad (2)$$

Where $i=1,2,\dots,r$ and $j=1,2,\dots,r$. Here, r is total instance of class 1 is used to represent the variables belonging to which class, range from 1,2,...n. The variable x_i^1 and x_j^1 denotes that instance value of each class. Here two variables x_1, x_2 distance is calculated using above mentioned equation.

2.3 Similarity and Dissimilarity Weight Calculation

In the previous step, the minimum distance is calculated by comparing All Pairs Distance and stored as $R(x_i^1, x_j^1)$. Through this weight terms is calculated using equation (3) and equation (4). Through this weight terms similarity, dissimilarity data is found (i.e.) we are reassigning the data points based on the above weight terms. Similarity and dissimilarity weight calculation is done by using the following equations.

$$\xi_{ij}^- = e^{-\frac{1}{2} \left[\frac{\omega_{\max} - (R(x_i, x_j))^2}{\sigma^2} \right]} \quad (3)$$

$$\xi_{ij}^+ = e^{-\frac{1}{2} \left[\frac{(R(x_i, x_j))^2}{\sigma^2} \right]} \quad (4)$$

Where

$$\omega_{\max} = \max \{R(x_i^l, x_j^l)\} \quad (5)$$

ω_{\max} is used to calculate maximum distance value obtained from the particular class value using the equation (2). In this work, initially we mention that the data are classified as random manner using threshold value. In this case one instance may perfectly group in to a particular class. So it doesn't need any modification. This type of instances are mentioned as ξ_{ij}^+ . In some cases initial classification may become wrong. These misclassified instances are specified as ξ_{ij}^- . Suppose x_i^1 is a instance, the similarity minimum distance must be calculated between $R(x_i^{-1}, x_{j+1}^1), R(x_i^{-1}, x_{j+2}^1), \dots, R(x_i^{-1}, x_r^1)$. From this result

if one particular instance, having the minimum distance to a particular class is mentioned as similar (ξ_{ij}^+). At the same time it may have the maximum distance (ie) it may be away from particular class which is mentioned by (ξ_{ij}^-) Gaussian constant ($[\sigma]$) Is used in both similar and dissimilar data validations. Finally calculated values are represented in sample space called latent variable representation. As per dataset, we initially represent data in a two dimensional way. But in final stage we reduce it in to single dimensional data which is represented by τ_i . Here using the equation (3) and equation (4) the similar and dissimilar values are categorized to particular class using the same conditions. In order to handle the upper and lower most data values, we handle the values of distance relation in both end cases. For each instance perfect class is checked with a distance value using $R(x_i^l, x_j^l) \leq R(x_i^l, x_{j+1}^l)$ (i.e.) $R(1,2) < R(1,3)$ in class one, using the similar value (ξ_{ij}^+) Particular instance x_i is added to a particular class. If $R(x_i^l, x_j^l) > R(x_i^l, x_{j+1}^l)$ Then, the instance is grouped by next class (i.e.) dissimilar (ξ_{ij}^-) to particular class. Then the result in τ_i Value is represented below

$$\tau_i = \emptyset[(x_i^l)] \begin{cases} c_1 + (1 - \xi_{ij}^+) \text{ if } l \in \{1, 2, \dots, k\} \text{ and} \\ R(x_i^l, x_j^l) \leq R(x_i^l, x_{j+1}^l) \\ c_1 + (1 - \xi_{ij}^-) \text{ if } l \in \{1, 2, \dots, k\} \text{ and} \\ R(x_i^l, x_j^l) > R(x_i^l, x_{j+1}^l) \end{cases} \quad (6)$$

At last, these trained data set are tested with Kernal Discriminant analysis, Logistic Regression, Classification via regression, Multiclass Classifier etc.

2.4 Further Concerns

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

Pseudo code for the proposed [APDC_1 AG 1]

Training phase (Phase I)

Input: Training dataset

Output: Trained dataset

- Compute the threshold (θ) value using Equation (1) to separate each class randomly.
- Optimal projection
 - (a) From each class min distance is calculated as per equation (2)
 - (b) From each $R(x_i, x_j)$ and from step (a) the weight terms are calculated using Equation (3),(4),(5). Here similarity and dissimilarity between the data are calculated.
 - (c) The latent space variable is predicted using equation

(d) Project the data from latent space variable τ_i .

- The regression function (Rg) is built by Considering the regression variable τ_i

Prediction (Phase II)

Input: Regression (Rg), threshold (θ), new trained dataset, test data set

Output: Predicted value

Predict the latent variable value using the Regression(Rg).

- Map the regression (Rg),to the corresponding Class and find the predicted value

3. Results and Discussion

To validate the proposed methodology some data sets like Heart Disease, Automobile, Bondrate, Contact-lenses, Thyroid, Pasture etc., are used. These datasets were extracted from uci repository. Table 1 shows the description of these, where number of patterns (N), number of attributes (K), number of classes (C) are taken. Initially, these datasets do not represent ordinal classification, but it represents regression. To evolve these regression into ordinal classification we have considered the desired result is categorized into five classes with equal frequency and these are referred from²⁻⁴.

Table 1. Characteristics of Datasets based on their classes

Dataset	N	K	C
Heart	270	13	5
Automobile	52	26	6
Bond rate	42	15	5
Contact-lenses	18	6	3
Thyroid	161	6	3
Pasture	27	23	3
Housing	39	25	3

The Entire work is validated based on All Pairs Distance Calculation using One Against One Method and these results are compared using, Kernal Discriminant Analysis, Logistic Regression, Multi Class Classifier, Classification Via Regression etc.

3.1 Performance Measures

In this work Accuracy, Mean Absolute Error, Average Mean Absolute Error, Kendall's used to validate the

Proposed work. Accuracy is stated as how well a given criteria guess the predicted attribute for the new data and it is calculated using the equation (7).

$$acc = \frac{1}{k} \sum_{i=1}^k |y_i^{\infty} - y_i| \quad (7)$$

The Mean Absolute error is close to the Mean Squared Error, only difference is, it uses absolute values instead of squaring. Here average of these absolute value is taken to consider the mean absolute and is measured using equation (8)

$$MAE = \frac{1}{k} \sum_{i=1}^k e(x_i) \quad (8)$$

Average Mean Absolute Error is measured using equation (9)

$$AMAE = \frac{1}{k} \sum_{i=1}^k (MAE_K) = \frac{1}{k} \sum_{i=1}^k \frac{1}{N_K} \sum_{i=1}^{N_K} e(x_i) \quad (9)$$

Mean Zero-one Error finds the misclassification rate and the rate of incorrect classified pattern is measured using equation (10)

$$MZE=1-acc \quad (10)$$

Association between two measured quantities are measured using Kendall's Measure. Degree of non-random correspondence between observers or similar categorical variable is calculated using equation (11)

$$[\tau]_y = \left(\sum [[c_{ij}]^{\infty} c_{ij}] \right) / \sqrt{\left(\sum [[c_{ij}]^{\infty}]^2 \sum [c_{ij}]^2 \right)} \quad (11)$$

After Evaluating measures such as Accuracy, Mean Absolute Error, Average Mean Absolute Error, Kendall's, results are displayed in Table 2 and these measures graphical representation are shown in "Figure 2"

Table 2. Measures considered for the Proposed APDC and its Methods Compared

Dataset/Measures	Accuracy	MAE	AMAE	Kendall's
Heart	0.851	0.375	0.352	0.887
Automobile	0.728	0.377	0.375	0.742
Bondrate	0.551	0.221	0.874	0.356
Contact-lenses	0.721	0.299	0.512	0.650
Thyroid	0.963	0.017	0.061	0.875
Pasture	0.665	0.254	0.322	0.817
Housing	0.703	0.371	0.391	0.627

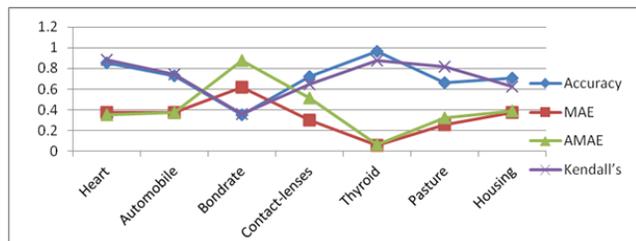


Figure 2. Mean Ranking of the Datasets considered for Proposed APDC and the Base Algorithm Used.

Table 2 shows various performance measures such as Accuracy, Mean Absolute Error, Average Mean absolute Error, Kendall’s measure of Proposed APDC_1AG 1 for different datasets. Table 3 shows about comparison between different methods like LR, J48, CVR, MCC, KDA with APDC for different datasets based on Accuracy Values. Here Bold faced data represents maximum accuracy achieved for each data set and italic font represents second highest accuracy achieved for each data set. For example, in Table 3 Heart Disease Dataset has achieved maximum accuracy in proposed APDC and second highest accuracy in Classification via Regression Method.

Table 3. Results of the accuracy for the Proposed APDC and Methods Compared

Dataset/ Method	LR	APDC	J48	CVR	MCC	KDA
Heart	0.562	0.851	0.792	0.812	0.739	0.742
Automobile	0.673	0.728	0.769	0.788	0.529	0.712
Bondrate	0.456	0.551	0.533	0.600	0.333	0.521
Contact-lenses	0.694	0.721	0.733	0.719	0.666	0.578
Thyroid	0.912	0.963	0.944	0.962	0.981	0.957
Pasture	0.678	0.665	0.778	0.879	0.667	0.883
Housing	0.832	0.703	0.712	0.810	0.789	0.752

Similarly Table 4 and Table 5 are derived for different datasets based on Mean absolute Error and Kendall’s Performance Metrics from these we can easily identify Proposed APDC gave better result while compared with other methods. In addition to this, how values varies between various measures are also compared.

Table 4. Results of the Mean Absolute Error for Proposed APDC and Methods Compared

Dataset/ Method	LR	APDC	J48	CVR	MCC	KDA
Heart	0.256	0.375	0.352	0.887	0.245	0.382
Automobile	0.343	0.377	0.139	0.091	0.139	0.268
Bondrate	0.567	0.221	0.892	0.241	0.222	0.397
Contact-lense	0.538	0.299	0.222	0.045	0.195	0.352
Thyroid	0.033	0.017	0.025	0.045	0.067	0.291
Pasture	0.357	0.254	0.225	0.149	0.186	0.362
Housing	0.275	0.371	0.391	0.627	0.321	0.256

Table 5. Results of the Kendall’s Measure for Proposed APDC and Methods Compared

Dataset/ Method	LR	APDC	J48	CVR	MCC	KDA
Heart	0.752	0.887	0.725	0.629	0.726	0.423
Automobile	0.782	0.742	0.452	0.699	0.724	0.380
Bondrate	0.353	0.356	0.181	0.239	0.082	0.314
Contact-lense	0.473	0.650	0.429	0.877	0.715	0.429
Thyroid	0.919	0.875	0.916	0.877	0.916	0.959
Pasture	0.789	0.817	0.500	0.667	0.834	0.500
Housing	0.614	0.627	0.612	0.512	0.622	0.525

However, Ordinal dataset gave better result in APDC but in some other cases it gave second best values. For efficient comparison, we compared ordinal method result such as APDC, LR, CVR with some other methods like J48, MCC. “Figure 3”, “Figure 4”, “Figure 5” shows how different dataset results are compared with various method based on Accuracy, MAE, Kendall’s vales respectively. To differentiate each method graphical representation is plotted with different colours according to datasets.

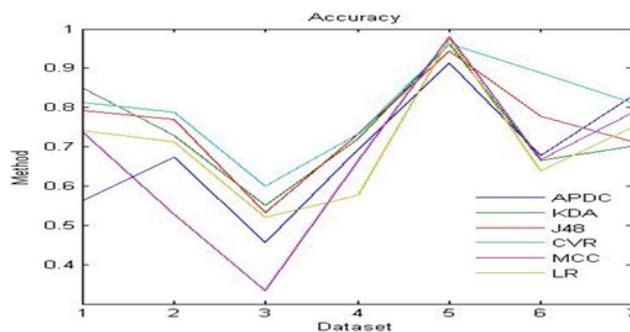


Figure 3. Graphical Representation of Accuracy for Proposed APDC and Base Algorithm Used.

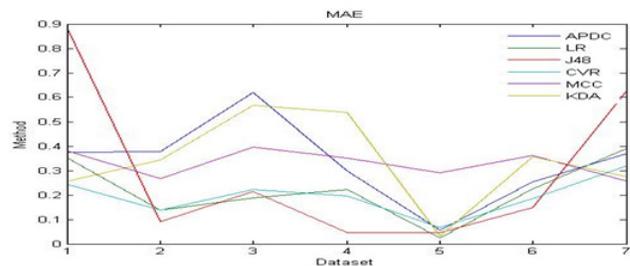


Figure 4. Graphical Representation of Mean Absolute Error for Proposed APDC and the Base Algorithm Used.

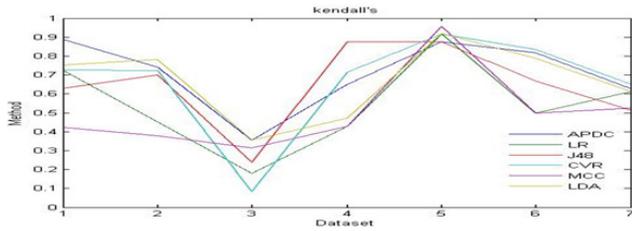


Figure 5. Graphical Representation of Kendall's for Proposed APDC and the Base Algorithm Used.

4. Conclusions

The Methodology proposed here is based on different classification tasks which are performed earlier. All Pairs Distance Calculation using One Against One Method is proposed here. One against one method checks all the possible pairs of combination to check whether randomly set data belongs to particular class or not. Here reformulation of data point is taken in to account, since we are rearranging the hyperline till we get similarity vectors. Proposed methodology uses data set like Heart Disease, Automobile, Bondrate, Thyroid, Contact-Lenses, Pasture, Housing etc. to validate the proposed method and it has been found that it is superior when comparing with other classifiers. Kernel Discriminant Analysis, Logistics regression, J48, CVR, MCC were applied to appraise this conclusion. In addition to this, the superiority of the proposal for All Pairs Distance Calculation Using One Against One has been confirmed while discussing with ordinal regression.

5. Acknowledgement

Thanks to Sathyabama University for their support in Publishing this work. Also we would like to express our great appreciation to all the members of our laboratory for their technical insight & simulating ideas, which is greatly contributed to the success of our research.

6. References

- Gu B, Sheng VS, Tay KY, Romano W, Li S. Incremental Support Vector Learning for Ordinal Regression. *IEEE Transactions on Neural Networks*. 2007; 6(1):1403-16.
- Kotlowski W, Slowinski R. On Nonparametric Ordinal Classification with Monotonicity Constraints. *IEEE Transactions on Knowledge and Data Engineering*. 2012; 25(11):2576-79.
- Krishnapuram B, Carin L, Figueiredo MA, Hartemink AJ. Sparse multinomial logistic regression: Fast algorithms and generalization bounds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005; 27(6):957-68.
- Fouad S, Tino P. Ordinal-Based Metric Learning for Learning Using Privileged Information. *Dallas, TX: IEEE International Joint Conference on Neural Networks, (ICJNN)*. 2013; p. 1-8.
- Hastie TJ, Tibshirani RJ. Nonparametric regression and classification. Part II: Nonparametric classification. Springer-Verlag: New York: From Statistics to Neural Networks: Theory and Pattern Recognition Applications, Cherkassky V, Friedman JH and Wechsler H, Eds. 1996; p. 70-82.
- Wu T-F, Lin C-J, Weng RC. Probability estimates for multiclass classification by pairwise coupling. *Journal of Machine Learning Research*. 2004; 5:975-1005.
- Tino P, Dorffner G. Predicting the Future of Discrete Sequences from Fractal Representations of the Past. *Machine Learning*. 2001; 45(2):187-17.
- Perez-Ortiz M, Gutierrez PA, Hervás-Martínez C. Projection Based Ensemble Learning For Ordinal Regression. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*. 2014; 44(5):681-94.
- Yarnold PR. Maximum-Accuracy Multiple Regression Analysis: Influence of Registration on Overall Satisfaction Ratings of Emergency Room Patients. *Optimal Data Analysis*. 2013; 2(1):72-75.
- Baccianella S, Esuli A, Sebastiani F. Evaluation Measures for Ordinal Regression. *Pisa: 2009 Ninth International Conference on Intelligent Systems Design and Applications, ISDA'09*. 2009; p. 283-87.
- Liu Y, Zheng YF. One-Against-All Multi-Class SVM Classification Using Reliability Measures. *Proceedings IEEE International Joint Conference on Neural Networks, ICJNN'05*. 2005; 2:849-54.
- Kaveri VV, Maheswari V. Cluster Based Anonymization for Privacy Preservation in Social Network Data Community. *Journal of Theoretical and Applied Information Technology*. 2015; 73(2):269-74.
- Polat K, Gunes S. A new feature selection method on classification of medical datasets: Kernel F-score feature selection. *Journal of Expert Systems with Applications*. 2009; 36(7):10367-73.
- Chu W, Keerthi SS. Support Vector Ordinal Regression. *Neural Computation*. 2007; 19(3):792-15.
- Jinila B, Komathy K. Cluster oriented ID based multi-signature scheme for traffic congestion warning in Vehicular Ad hoc Networks. *Springer International Publishing: Switzerland: Emerging ICT Bridging the future Proceedings of the 49th Annual Convention of the Computer Society of India*. 2015; p. 337-45.
- Baccianella S, Esuli A, Sebastiani F. Evaluation measures for ordinal regression. *Pisa, Italy: Proceedings of the Ninth International Conference on Intelligent Systems Design and Applications (ISDA 09)*. 2009; p. 283-87.
- Mary Gladence L, Ravi K, Karthi M. Heart Disease Prediction using Naive Bayes Classifier -Sequential Pattern Min-

- ing. International Journal of Applied Engineering Research (IJAER). 2014; pp.
18. Gladence LM, Ravi K, Karthi M. An Enhanced Method For Detecting Congestive Heart Failure -Automatic Classifier. Ramanathapuram: IEEE International Conference on Advanced Communication Control and Computing Technologies, ICACCCT-2014. 2014; p. 586-90.
 19. Subhashini R, Jawahar Senthil Kumar V. A Framework for Efficient Information Retrieval using NLP Techniques. Springer-Verlag Berlin Heidelberg: Proceedings of the International Conferences on Advances in Communication Network and Computing, CNC 2011. 2011; p. 391-93.
 20. Rajalakshmi V, Anandha Mala GS. Anonymization by Data Relocation Using Sub-clustering for Privacy Preserving Data Mining. Indian Journal of Science and Technology. 2014 Jan; 7(7):975-80.
 21. Shabana AP, Samuel SJ. An analysis and accuracy prediction of heart disease with association rule and other data mining techniques. Journal of Theoretical and Applied Information Technology. 2015; 79(2):254-60.