Multi-Attributes Web Objects Classification based on Class-Attribute Relation Patterns Learning Approach

Sridhar Mourya[1]
Department of CSE
JNTUH - Hyderabad, India.

Dr. P.V.S. Srinivas[2]
Department of CSE
SNIST - Hyderabad, India.

Dr. M. Seetha[3]
Department of CSE
GNITS - Hyderabad, India.

Abstract—The amount of Web data increases with the proliferation of a variety of Web objects, primarily in the form of text, images, video, and music data files. Each of these published objects has some properties that support defining its class properties. Because of their diversity, using these attributes to learn and generate patterns for precise classification is very complicated. Even learning a set of attributes that clearly categorize the categories is very important. Existing attribute learning methods only learn attributes that are closely related to multiple similar objects, but if similar class objects have different attributes, this problem is difficult to learn and classify them. In this paper, a Multi-attributes Web Objects Classification (MA-WOC) based on Class-attribute Relation Patterns Learning Approach is being proposed, which generates a class-attribute with its multi relations patterns. The MA-WOC calculates the relationship probabilities of the attributes and the associated values of the class to understand the degree of association of the construction of classification pattern. To evaluate the effectiveness of the classifier, this will compare to an existing classifier that supports a multi-attribute data set, which shows improvisation of precision with a significant minimum Hamming loss. To evaluate the effectiveness of MA-WOC classification a comparison among the classifiers that are supported to the multi-attribute dataset are being performed to measure the accuracy and hamming loss.

Keywords—*Classification; multi-attributes; web objects; attribute learning; distinct-class relation*

I. INTRODUCTION

Web data has become a collection of heterogeneous objects for defining information, for example a particular news topic is being available in text, video and images objects form. Each of these objects can be recognized with the multi-attributes values, and also this classification can be further sub-categorized into various sectors of information such as, "politics", "sports", "education", "entertainments", etc., providing more additional attributes for each individuals [1], [2], [3], [4]. As these information content have multiple attributes, but these can be associated with these multiple classes attributes to recognize them as a distinct class. However, data processing and learning of such multi-attributes data and classifying them is a major challenge for the information providing applications [5], [6], [7]. Many solutions are suggested in the past [1-16] to learn multi-value association and categorize through learning their attributes values. But due to the diversification of these attributes the classification results are not so accurate. This classification methods are needed an accurate learning and association method to provide high classification accuracy.

In the existing classification techniques are majorly assumes that the collection of attributes pattern and being link to one class representation. A dataset collection related to academic information can be classified according to their context for different class attributes in views of the researcher, students or publishers, but these information attributes may have a multiple common or diverse relational attributes network. Thus, network data can have multiple attributes in form of a "vector-based" and multiple attribute relational in form of a "graph-based" illustrations. Moreover, in complex social network data sets such as "Twitter", "Face book", and "LinkedIn" are also typically associated with in excess of one attributes. Here, the attributes are needed to be classified in terms of user interests to the posted text data through a multi-attribute classification [8], [10].

This multi-attribute classification is a challenging task in learning and data mining research, it required to learn the effective relation among the attribute to build an efficient classifier for such heterogeneous datasets. Most of the studies in literature have to make classifier through attribute selection [13], [14] and the association classification Classifiers usually predict data object classes derived from a set of training information. However, the attributes of the classifier configuration is not sufficiently investigated to influence the attributes value of the predicted class, or even in the literature, the problem has not been explored to its extend.

The selection and reduction of feature methods [25] has been used for multi-value classification in the previous proposals. The majority of these recommendations are analyzed by mutual analysis and reduce the features that do not provide critical information for predictive classes. Improvement support can be used to train and organize these lesser or more selective features. But the complexity of the object with multiple values converts the structure of what is appropriate and what is appropriate for the classification. Although some selective methods work well for some classifier [8], [20], but multi-value learning and associating each of the value to its class may not be encouraging to their features. For example, a document with the collection of words term may have some object categories related to "entertainment", "politics", "sports", "economics", and so on, is highly complex to classify to particular class.

Classification is one of the well-liked approaches to associate both the attribute properties and relationship information. This includes node classification techniques that collects the properties of the model properties collectively and indicates the properties of the relevant equipment and the attributes of the properties and combines relationship-based classification in the existing machine learning and repetition process. Researchers have proposed a semi-supervised grouping technique to be partially related to classified networks [10], [16]. The techniques in this group all mean that data points have simply one attribute and one category representation. However, many actual data sets contain further information that can be used to advance performance.

This paper utilizes the additional information and proposes a Multi-attributes Web Objects Classification (MA-WOC) based on Class-attribute Relation Patterns Learning Approach which will be generating a class-attribute with multi its relations patterns. The method of associations of class and attributes are based on a relation probability of attribute and class association value which measure the discreteness of an attributes relation to their comparing class. The proposal is emphasizing to find a discrete class of the web objects having multiple attributes. The MA-WOC addresses the challenges through construction of the multi-attribute learning technique to solve multi-attribute web objects classification.

The following paper is organized as follows. Section-2 describes the related works performed in related to multi-attributes classifications, Section-3 discussed the proposed multi-attributes web objects classification which describes the problem multi-attribute relation learning and class-attribute relation patterns learning approach, Section-4 presents the datasets description, measures and result analysis. Section-5, conclude the conclusion of the paper.

II. RELATED WORKS

The exact classification of data is focused on making a deeper analysis of data to provide the necessary information [5], [9], [11], [12]. The classification object is classified by classifier test function and shows a learning set for trained classes [14], [15] For example, a data set containing a collection of records, and every record event has a set of attributes properties that are considered as the set of identity class categories. A classifier performs the classification of the data objects based on the established class knowledge classifier. The purpose of the classifications is to create the perfect classifier, which will provide accurate support for anonymous data classification required for real time. Supervised learning is successfully used in many learning activities to identify relevant objects. A traditional learning system not able to associate appropriately to its class due its complex multiple attributes.

X. Kong et al. [10] apply a grouping technique to handle multi-attribute classification of a single attribute's single relational network data. Converts multiple attribute problems to multiple binary-related issues for each property and captures complex attribute correlations that can exist between properties within the identical instance and correlated instances, and stacking the properties of the similar instance and related instance with the feature set.

F. Charte et al. [9] presents a multi-valued classification scheme for handling multiple data value objects. The proposal is to solve the traditional problem of high-dimensional data classification may suitable for large number of data attributes. The selection of feature selection on the basis of the data transformation and the assessment of association rules transformed based on the attributes dependence. The attribute value identifies the selection feature of the classification algorithm with multiple values. This approach can be successful for linear changes in data objects to indicate the value of addiction, but the results can be inaccurate for highly distributed data in multi-attribute data objects.

The iterative annotation of the "Multiple Relational Social Network (IMR)" [16] is a multiple attribute grouping technique for single attribute multiple relational network data. Multi-attribute problems are treated as multiple-binary related problems by learning the multi-attribute problem from multiple relationships for each attribute classifier to classifiers for each attribute of the feature set that is stacked with the attribute information of the related instance, and this technique does not capture attribute correlations [10].

X. Shi et al. [17] proposed a heterogeneous learning technique as a single attribute classification technique for multi-attribute multi-relational network data. The technique is a bug-driven model that constructs a function on each property view and tries to use two constraints to globally reduce the empirical error function, "consensus across various attribute sources" and "connected instances should have similar prediction".

M. L. Zhang et al. [2] aims at the problem of multi-attribute learning in feature selection. The author utilizes a strategy to learn property-specific features for different class property differentiation. The proposed algorithm name for multi-property learning, LIFT, implements clustering analysis for positive and spoken instances to construct clusters based on attribute by attribute. The classification knowledge base for training and testing is queried in the clustered functional group results. However, while the proposed approach shows promising directions in multi-attribute learning for classification, the importance of feature association for other features should be explored for further optimization.

Multi-attribute learning techniques learn models for each available observation of data and minimize discrepancies between different attributes of non-contributed data. Co-training [18], [19] is a multi-property semi-supervised learning algorithm that learns models for each observation of data and exchanges specific predictions to make use of complementary information available in various properties. However, the multi-attribute learning method does not model network data with relational features.

The "Multi-attribute learning techniques" observe the learning model for each available data and minimize inconsistencies between multiple attributes on unattributed data. The "Co-training methods" [18], [19] is a multi-attribute semi-supervised learning algorithm that learns models and exchanges some predictions each time they observe data, thereby utilizing the additional information available between

different attributes. However, multi-attribute learning methods do not use relational features to model network data.

Based on the above reviews and approaches, this shows the importance of multi attribute in classified areas. These areas represent the importance of attribute selection in the accuracy classification. But learning the most characteristic features for categorization is a challenging issue. To overcome these issue and limitation this will propose a new approach to classify the Multi-attributes Web Objects Classification (MA-WOC) based on Class-attribute Relation Patterns Learning Approach. The MA-WOC calculates the probability of a relationship of attribute and class associations for each feature to learn the extent of the organization to create classification patterns. The details of the proposed procedure are discussed in the following sections.

III. PROPOSED MULTI-ATRRIBUTES WEB OBJECTS CLASSIFICATION

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

A. Problem of of Multi-Attribute Relation Learning

Traditional learning systems are especially studied during the classification of supervised machine learning [21], [23], [24]. In this classification, the data object is associated with the value of the supervised learning system to determine the properties of the attribute to be used for the classification as shown in Fig. 1.
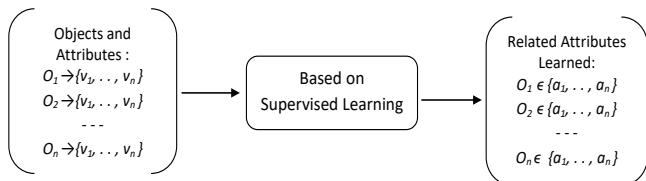

Fig. 1. Objects Attributes Learning through Traditional Supervised Learning..

Although this learning is very well suited for a single class attributes term, but complexity arises when an object has multiple attributes. In the existing traditional supervised learning [20], [22], [26] improvisations for applying multiple attribute data objects have been found. However, most of the proposed solutions are based on the functional capabilities of studying the peer dependencies or attributes and calculate the number of common events [27]. However, this type of information attribute may not be appropriate for domains that are not available. In some cases, this is determined by the association rule algorithm and the main dependencies of the correlation attribute, but this does not help to change the various sets of data attributes and other domains. The goal of this work is to create classifiers based on new associations with many attributes that can be implemented in various areas of the multidrug dataset and provide the necessary precise and quick classification.

TABLE I.   DISTINCT CLASS-ATTRIBUTES TABLE

Object Class	Attributes Associating Values
Mobile	Display, memory, camera, android, Batteries, Weight, Colour, Dual SIM, Bluetooth, etc
Scenery	Clouds, Landscapes, Lakes, waterfalls, Beach, Sunset, Fall Foliage, Fields, Mountain, Urban, forest, tree, bridges, etc.
Birds	Wings, fly, Brown Creeper, Pacific Wren, Pacific-slope Flycatcher, Red-breasted Nuthatch, Dark-eyed, etc.
Book	Information, article, book, children, story, comic, computer, dynamics, education, learning, games, social, universe, etc
Vehicle	Car, model, colour, mileage, bike, speed, make, engine, power, Displacement, fuel capacity, etc
Music	Album, sound, lyrics, record, player, melody, stereo, voice, singer, etc.

The classification depends on the choice of the object and identification its attributes. It was noted that in a particular object there are two or more attributes of data, indicating a certain level of organization among them. This mult-attribute association can be very useful for multilevel data classification. The offered approach realizes the mechanism of studying and modelling. In the first step, the probability of a relation between the attribute value and the class association is calculated, which is very suitable for the object class proposal, and in the second stage the approach generates various multi-attributes that are supported to create useful class-attribute relationship models for the different classes required for classification.

To compute the relation probability of attribute and class association value (PAvalue) for an object instance this will relates the association of attributes with a Distinct Class-Attributes Table (DCAT) defined for the objects as given in Table-I.

Let's considered a set of objects "*WO*" consists of *n* objects instances having *k* attributes vectors which represented as, "$WO = \{O_1, \ldots, O_k\}$" and its attributes as a collection of "$A = \{a_1, \ldots, a_k\}$". The objects and attributes of DCAT let be represented as, "$D = \{C_1, \ldots, C_n\}$" and its attributes as "$T = \{t_1, \ldots, t_n\}$". In order to compute a probability association value, PA_{value} of an object attributes in compared with the DCAT, this will need to learn the intersection of each record of "T" of DCAT collection is compared to compute the association frequency as "A_{freq}" using (1). The obtained "A_{freq}" is utilized to compute the "PA_{value}" of each individual objects using (2).

$$A_{freq} = \sum \left(\int_{i=1,k=1}^{n,k} (T_n \cap A_k) \right) \qquad (1)$$

$$PA_{value} = \left(\frac{A_{freq}}{n} \right) \qquad (2)$$

Where, *n* is the number of data objects attributes in "*DCAT*" of each individual objects, *k* is the number of multi-attributes of an object. The value of "PA_{value}" ranges between 0 to 1, the higher the value the closer the association to class. The computed "PA_{value}" of each object are being utilized to

construct patterns for the classification. The algorithm-1 presents the steps of association in detail.

Algorithim-1: Finding *Associated* Class for an Objects

Input :   *WO*, a single dimensional learning data
         DCAT, a two dimensional Distinct Class-Attributes Table
Output : *WO$_{Class}$* , Object Class Value
Method : **for** *i=0, i* < number of objects in *WO*
      {
         w_i = *WO[i]*;
         A[]=getValues(w$_i$);
         for *t=0, t* < number of records in *DCAT*
         {
           C_t = *DCAT[t][0]*;   *// -- class value*
           T_t *[]* = *SCT[t][0]*; *// -- Association value*
           A_{freq} = *computeAF (A[] , T$_t$ [])*; -- *(Eq-1)*
           PA_{value} = A_{freq}*/ sizeof(T$_t$ [])*;   -- *(Eq-2)*
           PA_Value[t][] = *[C$_t$][PA$_{value}$]*;
         }
         // Find the highest PA_Value associated to assign a class
         WO$_{Class}$ = *getClass (PA_Value [ ][ ])*
    }

Selecting a particular class for a multiple attribute results in a significant loss of information [10], [20]. To overcome this problem, the method will expanding the class that learns *PA_Value* with multiple attribute values to create associative patterns using association rules to minimize Hamming losses in the data classification. The process of creating a pattern is shown in Fig. 2.

Fig. 2 illustrates the value of an instance in the association to find multiple values for the construction of a sample of classification. Let's assume, the training datasets, "*D = { (d$_1$,a$_1$), (d$_2$,a$_2$), . . . , (d$_n$,a$_k$)}*", where $d_i \in D$, $v_k \subseteq A$. To find the multi-attributes which can be highly relevant to build the classifier class accuracy this approach will consider a binary relevance of each instance attributes, for example, the value "*A = {a$_1$, a$_2$, a$_3$, a$_4$, a$_5$}*" can have a binary equivalence as, "*D = {(0,1,1,1,0), ... , (1,0,1,0,1)}*". This learning mechanism will utilize all binary values, such that the list of attributes set will be generated from *D* which supports the minimum number of support count required.
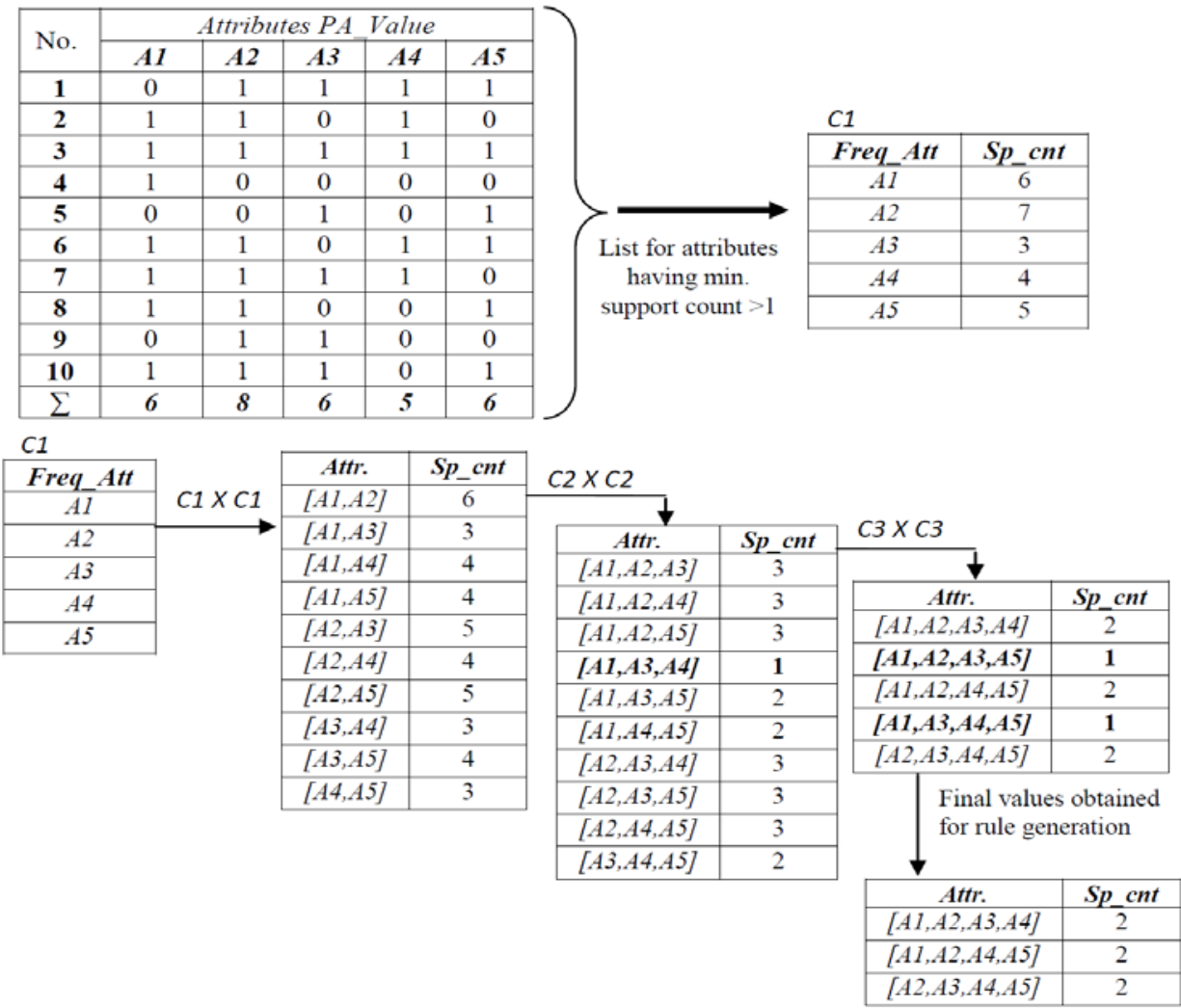

Fig. 2.   Multiple-Values Generating from Instance Value Data.

TABLE II.   PATTERN GENERATED FOR THROUGH MULTI-ATTRIBUTE ASSOCIATION

Object Class Att.	Generated Associated Multi-Values	Generated Patterns for Classifier
C_{Att}	[A1,A2,A3,A4] [A1,A2,A4,A5] [A2,A3,A4,A5]	• {A1},{A2},{A3},{A4},{A5} • {A1,A2}, {A1,A3},{A1,A4},{A1,A5} • {A1,A2,A3}, {A1,A2,A4},{A1,A2,A5} • {A1,A2,A3,A4}, {A1,A2,A4,A5},{A1,A2,A4,A5}

In this case, since the absolute rating of support is *2*, and the minimum relative support will be "*2/10 = 20%*". The list received from "C1" is configured as an item that matches the minimum support, while the rest of the items are ignored. In addition, to determine the most common and associated attribute in the resulting "*C1*", combine the outcome with "*C1* ⋈ *C1*" to build the "*C2*" attribute with two attributes and continue until you get one value of the pattern. This iteration continues until there are several attributes that satisfy the minimum support. The final multi-attribute is considered the most relevant attributes. Now through utilizing *PA_Value Class* as C and multi-attribute ingress to create classification rules for classifiers, as shown in Table II.

The generated rules will be used for accurate classification of multiple attribute objects. Classification accuracy is also supported by effective clustering of data objects. In the next section, this work will be experimentally evaluated using multiple attribute data sets to analyze the accuracy and Hamming loss compared to conventional multi-attribute classification method.

IV. EXPERIMENT EVALUATION

A. Multi-Value Datasets

The complexity of multi-attribute classification stems from a variety of real-world environments and domain applications. For the data sets related to the experimental setup, the three main application areas of multi-attribute data "multimedia classification", "text classification" and "bioinformatics" are often observed. All data sets are primarily obtained from the "MULAN" [22] data store, consist of "number of instance", "attributes", "Values" and "L_{Card}" as shown in Table III.

The "L_{Card}", represents the attribute which establish the average number of attributes per test data. The "L_{Card}" measured are discussed in [3], [17] for each dataset as, "$D = \{ (d_n, A_k) \mid 1 \le n \le k\}$" and N is the total data records are denoted as,

TABLE III.   EXPERIMENT DATASETS DISTRIBUTION

Datasets	Objects	No. of Instances	No. of distinct Attributes	No. of distinct Values	L_{Card} Value
Bibtext	text	7395	1836	159	2.402
Scenes	images	2407	294	6	1.074
Birds	audio	645	260	19	1.014

$$L_{Card} = \frac{1}{N} \sum_{i=1}^{n} \mid A_k \mid \qquad (3)$$

B. Evaluation Measures

1) Hamming Loss (HL): This is the most accepted measure of further attribution of errors in the misclassification of data attributes. The measure evaluates the incorrect classification of instances and pairs based on attributes that are independent of expected and related attributes. The performance is considered perfect if "*HL = 0*". Here,"δ is the symmetrical dissimilarity between two datasets instances", "h is the literal for the hamming loss", "N is considered as number of test datasets", "d is the each individual test data attributes" and "A is the class values that are applied to the dataset".

$$Hamming\ Loss(HL) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{A} \mid h(d_i)\delta a_i \mid \qquad (4)$$

2) Accuracy (ACC): This measures the percentages of the test attributes of *d* correctly utilized for the object classification using the *A* attributes measurements in a given data set.

$$Accuracy(ACC) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{A} \mid \frac{h(d_i) \cap a_i}{h(d_i) \cup a_i} \mid \qquad (5)$$

The proposed Multi-attributes Web Objects Classification (MA-WOC) is evaluated over the popular "*Weka Tool*" using the datasets of MULAN [22] as discussed above. This will be comparing the learned patterns with the standard multi-attribute classification methods as, "*BR - Binary Relevance*", "*LP - Attribute Powerset*", "*CLR - Calibration Attribute Ranking*" and "*RAkEL - Random-k-Attributeset*" [28], [29] to understand the precision of improvisation of the attribute selection and accuracy of the classification for different datasets.

C. Result Analysis

This section describes the experimental results analysis obtained on executing the MA-WOC and other classifiers. In utilizing the "Class-Attribute Relation learning", the approach constructs the required patterns for the classification initially. Later the learned knowledge of the MA-WOC classifier is compared with the traditional multi-label classifier methods. The results obtained for each of the data sets are presented in Table IV below.

TABLE IV.   GENERATED PATTERN PAIRS FOR DATASETS

Datasets	Labels	Associated multiple-values	Non-Associated	MA-WOC Classification Pairs
Scenes	6	3	3	8
Birds	19	13	6	38
Bibtext	159	114	45	386

TABLE V.   HAMMING LOSS MEASURE COMPARISON

Datasets	MA-WOC	BR	MA-WOC	LP	MA-WOC	CLR	MA-WOC	RAkEL
Bibtext	*0.0125*	0.0151	***0.0117***	0.0161	*0.0098*	0.0144	*0.0132*	0.0151
Scenes	***0.0841***	0.0973	*0.0951*	0.1437	*0.0994*	0.1121	*0.1012*	0.0962
Birds	*0.0462*	0.0561	*0.0599*	0.0735	*0.0452*	0.0506	***0.0437***	0.0489

TABLE VI.   ACCURACY MEASURE COMPARISON OF CLASSIFIERS

Datasets	MA-WOC	BR	MA-WOC	LP	MA-WOC	CLR	MA-WOC	RAkEL
Bibtext	***0.7204***	0.4187	*0.6437*	0.3869	*0.5015*	0.4089	*0.3854*	0.3657
Scenes	*0.799*	0.553	***0.839***	0.5893	*0.7918*	0.5265	*0.6247*	0.6841
Birds	*0.6708*	0.4666	*0.7189*	0.5295	***0.7319***	0.528	*0.727*	0.5452

Based on the Table-IV generated patterns an experimental run of a 10-fold validation classification for the test datasets, and measures the *ACC* and *HL* in compare the traditional multi-label classifiers. The Table-V presents the classifiers Hamming Loss in comparison with the "*BR,LP, CLR,* and *RAkEL*". In terms of *HL*, "the lower the loss of attributes the better the accuracy". The Table VI present the *ACC*, here "the higher the accuracy the better the improvisation".

Based on the HL and accuracy assessment results as shown in Fig. 3 and Fig. 4, the evaluation analysis found that MA-WOC showed improvement compared to the three traditional classifiers, accepted in the case of the "RAkEL" and "scene" data sets. The differences observed in both cases are very small. This difference can be considered uncertain, but it is very significant in performance compared to other situations. The improvisation of classification accuracy helps to effectively classify objects and effectively support non-categorical object clustering.
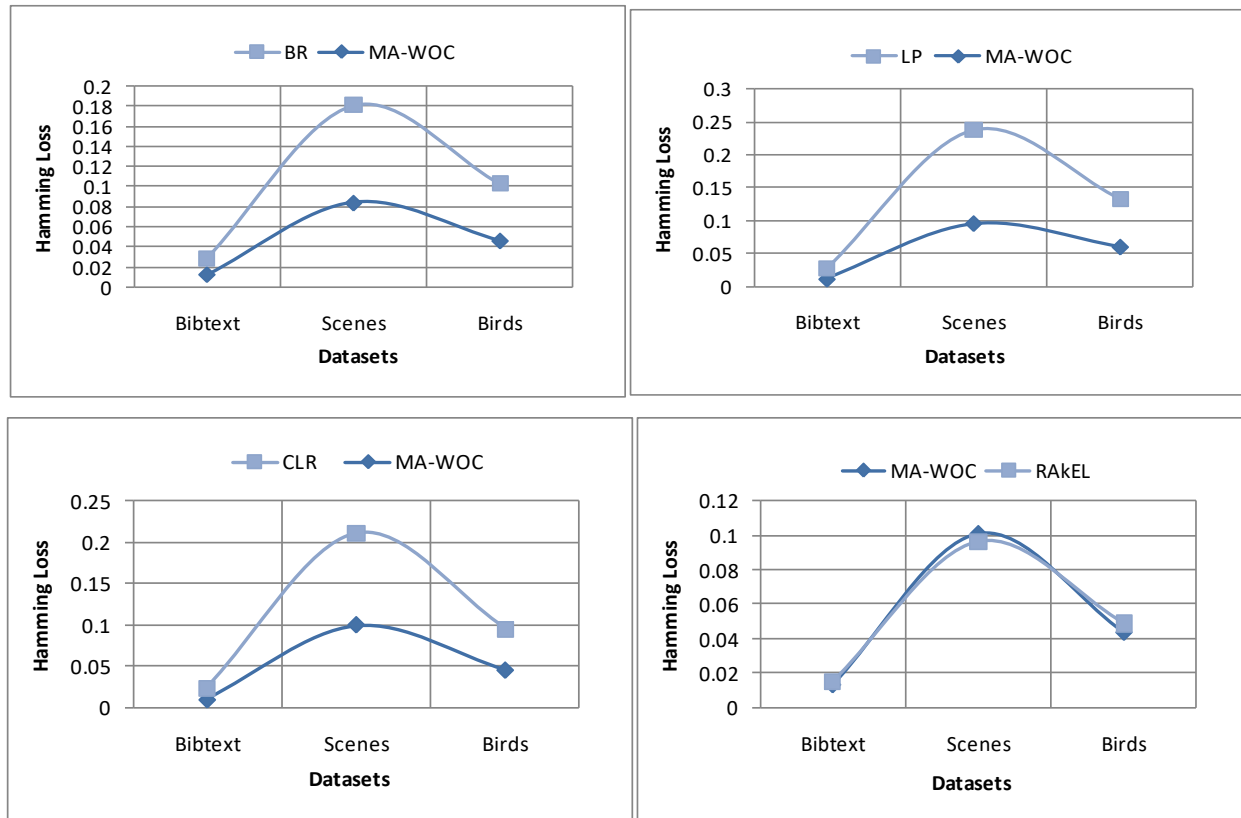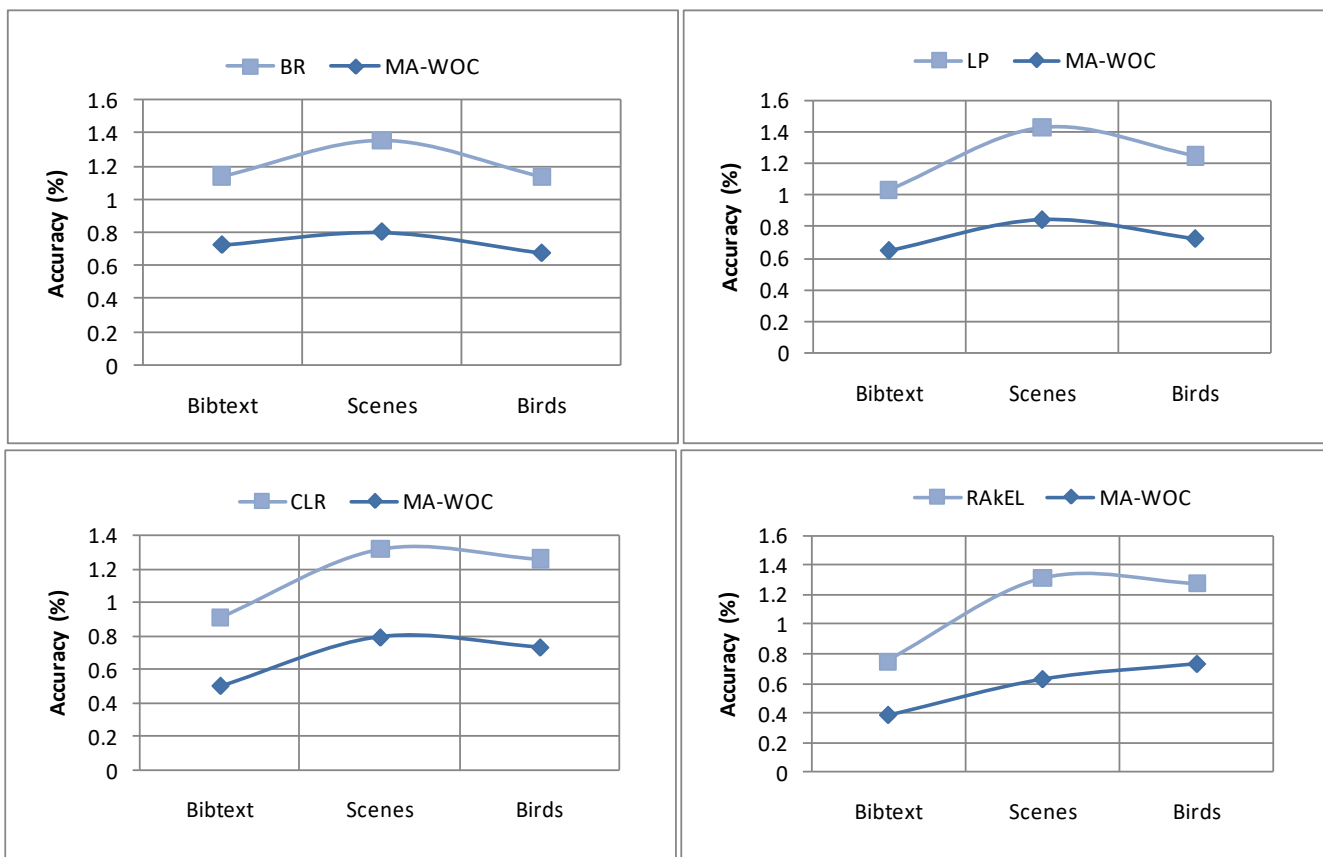

Fig. 3.   Hamming Loss Comparison.

Fig. 4.   Accuracy Percentage Comparison.

V.   CONCLUSION

In this paper, a multi-attribute Web object classification (MA-WOC) based on the class attribute relationship pattern learning method is being proposed, which utilizes attribute association between multiple attributes. The learning process initially identifies the probabilities that are well-suited for the class's suggested attributes and class-associated values, and finds another plurality of attributes that support the associated probabilities in the second step to build a class pattern that is useful for different object classifications. The proposed MA-WOC calculates the associated attribute frequency and probability associations and in compare to DCAT to understand the relationship between the attributes of the object instance and the relationship of the class association. The contribution of this proposal will be used to learn various multi-attribute data sets. Experimental evaluations show possible ways to learn multiple attributes for efficient classification using different algorithms. Statistical properties terminate availability and enhancements in multi-attribute classifications. In the future, this can be further studied for the utilization of the association properties with fuzzy and Bayesian elements to accelerate and improve multi-attribute classification.

REFERENCES

[1]   M. Elkano, M. Galar, J. Antonio Sanz, A. Fernandez, E. Barrenechea,F.o Herrera and H. Bustince, "Enhancing Multiclass Classification in FARC-HD Fuzzy Classifier: On the Synergy Between n-Dimensional Overlap Functions and Decomposition Strategies", IEEE Transactions On Fuzzy Systems, Vol. 23, No. 5, October 2015.

[2]   M. L. Zhang and Lei Wu, "LIFT: Multi-Label Learning with Label-Specific Features", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 37, No. 1, January 2015.

[3]   F. Charte, Antonio J. Rivera, María J. del Jesus, and Francisco Herrera, "LI-MLC: A Label Inference Methodology for Addressing High Dimensionality in the Label Space for MultiLabel Classification", IEEE Transactions On Neural Networks And Learning Systems, Vol. 25, No. 10, October 2014.

[4]   J. Shen, E. Zheng, Z. Cheng, C. Deng, "Assisting Attraction Classification by Harvesting Web Data", IEEE Access Volume: 5 Pages: 1600 - 1608, 2017.

[5]   T. -Y. Chan, Y.-S. Chang, "Enhancing Classification Effectiveness of Chinese News Based on Term Frequency", IEEE 7th International Symposium on Cloud and Service Computing (SC2), Pages: 124 - 131,2017.

[6]   J. Ruohonen, "Classifying Web Exploits with Topic Modeling", 28th International Workshop on Database and Expert Systems Applications (DEXA) Pages: 93 - 97, 2017.

[7]   M. P. El-Kafrawy, M. Amr Sauber, Awad Khalil, "Multi-Label classification for Mining Big Data", International Conference on Advances in Big Data Analytics, 2015.

[8]   M. L. Zhang and Zhi-Hua Zhou, "A Review on Multi-Label Learning Algorithms", IEEE Transactions On Knowledge And Data Engineering, Vol. 26, No. 8, August 2014.

[9]   F. Charte, A. J. Rivera, María J. del Jesus, and Francisco Herrera, "LI-MLC: A Label Inference Methodology for Addressing High Dimensionality in the Label Space for Multilabel Classification", IEEE Transactions On Neural Networks And Learning Systems, Vol. 25, No. 10, October 2014.

[10]   X. Kong, B. Cao, and P. S. Yu, "Multi-Label classification by mining Label and instance correlations from heterogeneous information networks", in Proceedings of the 19th ACM SIGKDD KDD'13. New York, NY, USA: ACM, pp. 614-622, 2013.

[11] P. Li, H. Wang, K. Q. Zhu, Z. Wang, and X. Wu, "Computing term similarity by large probabilistic is a knowledge", In Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management, ser. CIKM '13, New York, NY, USA, pp. 1401-1410, 2013.

[12] L. Chekina, D. Gutfreund, A. Kontorovich, L. Rokach, and B. Shapira, "Exploiting Label dependencies for improved sample complexity", Machine Learning, vol. 91, no. 1, pp. 1-42, 2013.

[13] N. Spolaor, E. A. Cherman, M. C. Monard, and H. D. Lee, "A comparison of multi-Label feature selection methods using the problem transformation approach", Electron. Notes Theoretical Comput. Sci., vol. 292, pp. 135-151, Mar. 2013.

[14] P. Mitra, C. A. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity", IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 3, pp. 301-312, Mar. 2002.

[15] C. -G. Li, X. Mei, and B.-G. Hu, "Unsupervised Ranking of Multi-Attribute Objects Based on Principal Curves", IEEE Transactions On Knowledge And Data Engineering, Vol. 27, No. 12, 2015.

[16] S. Peters et al, "Iterative annotation of multi-relational social networks", in Proc. 2010 International Conference on Advances in Social Networks Analysis and Mining. IEEE Computer Society, pp. 96-103, 2010.

[17] X. Shi et al, "Learning from Heterogeneous Sources via Gradient Boosting Consensus", in SIAM International Conference on Data Mining(SDM), pp. 224-235, 2012.

[18] L. Chekina, D. Gutfreund, A. Kontorovich, L. Rokach, and B. Shapira, "Exploiting label dependencies for improved sample complexity", Machine Learning, vol. 91, no. 1, pp. 1-42, 2013.

[19] R. S. Cabral, F. De la Torre, J. P. Costeira, and A. Bernardino, "Matrix completion for multi-Label image classification", in Advances in Neural Information Processing Systems USA: MIT Press, pp. 190-198, 2011.

[20] M. R. Boutell, X. Shen, J. Luo and C.M. Brown, "Learning multilabel scene classification", Pattern Recognit., vol. 37, no. 9, pp. 1757-1771, 2004.

[21] C. S. Ferng and Hsuan-Tien Lin, "Multi-Label Classification with Error-correcting Codes", 20th Asian Conference on Machine Learning, Journal of Machine Learning Research, 281-295, 2011.

[22] G. Tsoumakas, E. S.-Xioufis, J. Vilcek, and I. Vlahavas, "MULAN: A java library for multi-Label learning", J. Mach. Learn. Res., vol. 12, no. Jul, pp. 2411-2414, 2011.

[23] M. Wang, X. Zhou, and T.-S. Chua, "Automatic image annotation via local multi-Label classification", in Proc. 7th ACM Int. Conf. Image Video Retrieval, Niagara Falls, Canada, 2008, pp. 17-26, 2008.

[24] K. Dembczynski, W. Waegeman, W. Cheng, and E. Hullermeier, "On Label dependence in multi-Label classification", in Workshop proceedings of learning from multi-Label data. Citeseer, pp. 5-12, 2010.

[25] N. Spolaor, E. A. Cherman, M. C. Monard, and H. D. Lee, "A comparison of multi-label feature selection methods using the problem transformation approach", Electron. Notes Theoretical Comput. Sci., vol. 292, pp. 135-151, Mar. 2013.

[26] G. Tsoumakas, M.-L. Zhang, and Z.-H. Zhou, "Tutorial on learning from multi-label data", in Proc. Eur. Conf. Mach. Learn. Principles Practice Knowl. Discov. Databases, Bled, Slovenia, 2009.

[27] K. Dembczynski, W. Waegeman, W. Cheng, and E. Hullermeier, "On label dependence in multi-label classification", in Workshop proceedings of learning from multi-label data. Citeseer, pp. 5-12, 2010.

[28] J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-Label classification", In Springer Machine Learning, vol.85, no. 3, pp. 333-359, 2011.

[29] G. Tsoumakas and I. Katakis, "Multi Label classification: An overview", International Journal of Data Warehouse and Mining, vol. 3, pp. 1-13, 2007.