Journal of Theoretical and Applied Information Technology
15th October 2017. Vol.95. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A OT-*k* LABEL LEARNING CLASSIFICATION BASED ON ASSOCIATION RULES FOR MULTI-LABEL DATASETS

[1]L. KIRAN KUMAR REDDY, [2]Dr. S. PHANI KUMAR

[1] Research Scholar, GITAM University, Hyderabad, Telangana, India
[2] HOD, Department of CSE, GITAM University, Hyderabad, Telangana, India

E-mail: [1]kkl351763@gmail.com, [2]phanikumar.s@gmail.com

ABSTRACT

The real-world application has grown in need of heterogeneous data classification for almost all kind of datasets. The complexity in learning a class for a single object which is associated with multiple label sets is a key problem for multi-label datasets. Existing methods might be unfavourable for classification as each label consists of specific features characterization. This paper propose a *One-To-k* (OT-*k*) Label learning method through exploiting the labels characterization and using association rules to discover label dependencies for the classification. The main objective is to find *One-Label* which will be highly suitable for class suggestion using a OL-Prediction Table and *k*-labels to constructs a patterns of labels to deal with the multi-label database classification. The efficiency of OT-*k* is verified against other multi-label learning algorithms. The result analysis shows an improvisation in different case studies being performed.

Keywords - *Label Learning, One-To-K, Pattern, Classification, Association Rules, Multi-Label Datasets.*

1. INTRODUCTION

In the era of distributed real-world application various heterogeneous data exists which are identified as multi-label objects. These data might be from different domain such as, education, sports, multimedia, politics or medical [8],[9],[10]. The content of these data might demonstrates the several meaning and might also be related to several classes due to multi-label identification. Even high dimensional data usages in different application for information processing [11] and machine learning [12],[13],[14] facing hurdles in accurate classification. However, in all cases it has been identified that, the cause of problem is due to multi-label. We dealt with this problem through a one-to-*k* label learning classification using the association rules exist in a multi-label datasets to improvise the multi-label classifiers.

It is a challenging task in data learning and mining research to build an efficient classifiers for heterogeneous datasets having multi-label annotations. In literature, most works are being targeted to feature selection [12],[17], feature reduction[18],[19] and associative classification [4],[24] to build classifiers. A classifier which predicts the class of a data objects based on a set of training data. However, in the case of multi-label, influence of classifier construct have not explored enough which have a high impacts in the prediction of class labels and even in literature this problem is weakly being studied until now.

Feature reduction or selection methods [17],[18] are being used for multi-label classification in the past proposals [6]. Most of these proposals analyze the feature correlations between themselves and reduce those features which do not offers constructive information for the class prediction. These reduced or selected features are being used by the classifiers for training and classification to support the improvisation. But the complexity lies for the objects which are multi-labels and how to transform these objects suitable for the classification improvisation. Even though, these selection methods are doing well for some classifiers with multi-label learning [2], but for each class label it might not be optimal to its specific characteristics. For example, the text classification in a set of documents where it features words terms might be related to entertainment, politics, sports, stock etc.

The objective of this paper is to analyze and propose a new multi-label learning method based on the data objects multi-labels characterized through an association rules algorithm for improvising classification. It associates the label through a One-to-*k* Label using label density calculation and association algorithm. It emphasize to find *One-Label* initially which will be highly suitable for class suggestion among the data object classes, and in second stage, we learn *k*-label binary association among the multi-labels to construct a patterns for the classification.

The methods and algorithms are discussed in following paper is organized as follows. Section-II describes the related work performed in related to multi-label classifications, Section-III discussed the multi-label learning system which describes the problem description and One-to-*k* Label learning system, Section-IV presents the experiment evaluation and datasets and section-V presents experimental results analysis using multi-label datasets. Finally, conclusion of the paper is discussed in section-VI.

2. RELATED WORKS

The accurate classification of data is the prime focus in data mining for providing needed information [3],[5],[15],[16]. Classification mostly performed by a classifier through examining the features characteristics of an objects and assigns the trained knowledge class set [20]. For instance, a data set which consists a collection of records and each record instance have a set of attributes, where one attribute from the set will be considered for the class identification. Based on the identified class knowledge a classifier performs the classification of unobserved data objects. The objective of classification is to construct an accurate classifier to support unobserved data accurately for the real-time needs.

Supervised learning is successfully used in many learning tasks for identifying unobserved objects. But it does not fit well in the current real-time data objects due to the multiple semantic meaning for a data object. A text document related to news might be related to politics, sports, economics, entrainments etc., builds a multi-label features complexity for the classification for the traditional supervised learning systems.

G. Tsoumakas and I. Katakis [21] identify the problem of multi-label classification and proposed a solution through data transformation and multi-label classification algorithm adaption. The data transformation deals with the problem of multi-label data transform from one to multiple labels. The proposal uses the common off-the-shelf single-label classifier which limits the classification requirement. The classification algorithm is modified to suits to the specific domains multi-label classification in a particular context it also attains a high computation complexity.

K. Dembczynski et al. [26] discussed the formalization and classification of label dependency in multi-label classification. It mostly focus on the label dependency through distinguish between the conditional and unconditional label. It was observed that multi-label classification through unconditional dependencies modelling shows good performance, where as in case of conditional dependencies it shows a low performance in comparison. X. Kong et al. [7] also explores the multi-label classification based on various type of dependencies among the objects and its labels known as PIPL. The proposal mostly focus on the heterogeneous information to facilitate the classification. The evaluation shows an improvisation in performance but its limited to heterogeneous network information datasets.

F. Charte et al. [3] presents a multi-label classification technique to deal with multi-label data objects. The proposal target to solve the traditional problem in high dimension data classification having large number of labels. A feature selection by means of instance selection through data transformation and association rules discover based on the label dependencies. The label dependencies identifies the features selection in the multi-label classification algorithm. This approach might be successful in case of linear variation in data objects to discover the label dependencies, but it might attain inaccuracy in case of high data variance in multi-label data objects.

M. Zhang and Lei Wu [1] targets the problem of multi-label learning in feature selection. It exploits the strategy to learn label-specific features for the discrimination of different class labels. An algorithm name LIFT proposed for multi-label learning, which constructs clusters based on the feature specific label by implementing an clustering analysis on positive and negative instances. The classification knowledge base for training and testing are queried from the clustered feature group results. However, the proposed approach shows a promising direction in multi-label learning for the

Journal of Theoretical and Applied Information Technology
15th October 2017. Vol.95. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classification, but the importance of an features association to other features has to be explored for the further optimization.

Based on the above reviews and approaches we understand the importance of multi-labels in the area of classification. It suggest the importance of feature selection in the accuracy classification. But learning the most characteristic features for the classification is a challenging issues. On construct of above reviews and limitation we proposed a new approach to classify the multi-label datasets using a one-to-*k* Label Learning system based on a association rules for multi-label datasets. It's learning algorithm generalized the feature dependency selection based on the need and domain requirements. The details of the system is discussed in the following section.

3. MULTI-LABEL LEARNING SYSTEM

A. Problem Description

Traditional learning systems are mostly studied over supervised machine learning systems. In these systems, data objects are associated with a label and the learning systems supervised the set of data to learn the feature characteristics which will be used for the classification as shown in Fig.1.

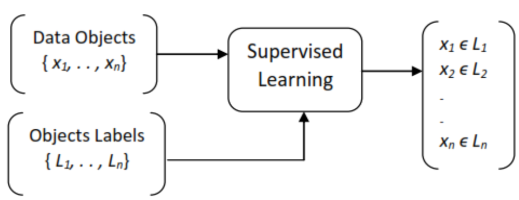

Fig.1: Traditional Supervised Learning

This learning is well adapted for single concept, but the complexity arises when the object have multi-labels. The improvisation of traditional supervised learning to adapt the multi-labels data objects are found in literature[2],[11],[17]. But most of the proposed solutions are based on the features dependence learning or associating the labels through counting their co-occurrences. However, this might be not applicable for the domains where such kind of label information are unavailable. In some cases, the fundamental dependency and correlations of label are identified using association rules algorithm, but it failed to facilitate the multiple changing datasets and its labels in different domains. We aim to construct a classifier based on new *One-to-k* (OT-*k*) Label learning system which can perform on different domain multi-labels datasets and allow us to provide an accurate and speed classification.

B. One-to-k(OT-k) Label Learning System

Classification approach depends on the accuracy of features selection and label identification. It was observed that two or more labels of a data objects in a domain suggests some levels of association among them. This association learning might be very helpful in multi-level data classification. We propose a two stage learning system to identified the *One-Label* which is highly suitable for class suggestion and in second stage we find the other *k-labels* which supports the *One-Label* class for constructing a class patterns useful for different querying classification.

Let's consider a training set *D* consists of *n* object instances having *k* labels vectors which represented as, $D = \{d_1, . . . , d_n\}$ and labels as $L = \{m_1, . ., m_k\}$. Now, the initial task of multi-label learning system is to find the *One-Label* class using the *L* vector. To do so we build an OL-Prediction Table (OPT) which consist the parental class label of domains as shown in Table-1.

Table-1: OL-Prediction Table

Class	Associating Labels
Birds	Brown Creeper, Pacific Wren, Pacific-slope Flycatcher, Red-breasted Nuthatch, Dark-eyed, etc.
Bibtex	architecture, article, book, children, community, computer, dynamics, education, elearning, games, social, social nets, etc
Medical	abnormalities, asthma, breathing, cardiac, checkups, fever, fracture, illness, kidney, mentally, radiograph, residual, stomach, swelling, throat, transplant, urinary, x-ray, zithromax, etc.
Scene	Beach, Sunset, FallFoliage, Field, Mountain, Urban, etc.

To learn *One-Label* for an instance we compute the associated label density in compared with the OPT. Label density (*LD*) will be calculated using the equation-1. The *LD* value ranges between 0 to 1, the higher the value the closer the association to class.

$$Label\ Density\ (LD) = \frac{\sum_{i=1}^{k}(l_k \in L)}{|L|} \qquad (1)$$

The method to find a instance *One-Label* Class using *OPT* and *LD* value is presented in Algorithm-1.

Algorithim-1: Finding *One-Label* Class for an instance

Input :
    D, a one dimensional training database
    OPT, a two dimensional OL-Prediction table
Output : *C*, Instance Parental Class Label
Method :
for i=0, i < number of instance in *D*
  $d_i = D[i]$;
  $L[] = getLabels(d_i)$;
    for t=0, t < number of tuples in *OPT*
      $C_t = OPT[t][0]$; // class label
      $A_t[] = OPT[t][0]$; // Association label
      $LD = computeLD(L[], A_t[])$;
      $LD\text{-}Value[t][] = [C_t][LD]$;
    end for
  // Find the highest *LD* value from the *LD-Value[ ][ ]*
  // Get the class label which has the highest *LD* value
  C = getClass (LD-Value[ ][ ]);
end for

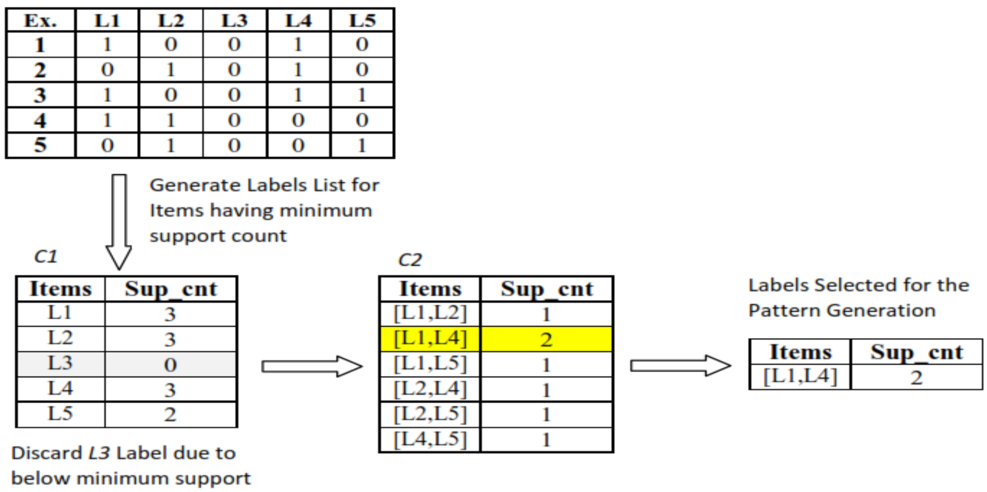

Fig.-2: k-Label Generating from Instance Label data

Selecting a single class for a multi-label instance cause a major information loss [6], [11]. To overcome this problem we extend, the *One-Label* class learning with *k-Labels* to construct a patterns through association rule to minimize the information hamming loss.

Let's assume, the training datasets $D = \{ (d_1,l_1), (d_2,l_2), \ldots, (d_n,l_k)\}$, where $d_i \in D$, $l_k \subseteq L$. To find the *k-labels* which can be highly relevant to build the class accuracy we consider a binary relevancy of the instance labels.

Let's consider an example where $L = \{l_1, l_2, l_3, l_4, l_5\}$ and $D = \{(1,0,0,1,0), (0,1,0,1,0), (1,0,0,1,1), (1,1,0,0,0), (0,1,0,0,1)\}$.

The Fig.2 illustrates the process of instance label association to find the *k*-label to construct the pattern for the classification. The learning mechanism scan all the binary labels of instance in *D* to generate a list of label sets which supports the minimum support count. Here we are using a absolute support count as 2, so that the corresponding minimum relative support will be 2/5= 40%. The list obtained as *C1* will be consist of items satisfying the minimum support and others will be discarded. Further, to identifying the most frequent and associated labels from the obtained *C1* we joins the $C1 \bowtie C1$ to generate C2 consists of 2-label itemsets. This iteration continuous till we have *k*-label satisfying the minimum support. The final *k*-labels obtained will be considered as most relevant and highly associated. Now using the *One-Label* class, *C* and the *k-label* items we generate the classification rules.

Table-2: Association Rules obtain Using OT-k Label

One-Label	***k-Label***	***Association Rules***
C_t	{L1,L4}	$\{(C_t,L1), (C_t,L4), (C_t,L1,L4)\}$

Journal of Theoretical and Applied Information Technology
15th October 2017. Vol.95. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. EXPERIMENTAL EVALAUTION AND DATASETS

To evaluate the performance of the proposal we calculates three popular measures suggested by Tsoumakas et.al. [23] for multi-label classification as, Hamming Loss (HL) and Accuracy. The datasets is used for the analysis is downloaded from MULAN [22] data repository as categorized in Table-3.

A. Evaluation Measures

1). Hamming Loss (HL) : This is most popular measure in multi-label classification. It evaluates the instance and label pair misclassification in terms of relevant and irrelevant label predicted. When, HL = 0, then the performance is perfect.

$$Hamming\ Loss\ (HL) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{L}\,|\,h(d_i)\delta l_i| \qquad (2)$$

where δ stands for the symmetric difference between two sets, N is the number of examples and L is the total number of possible class labels.

2). Accuracy (A) : It measures the percentage of correctly predicted labels among all predicted labels. Accuracy is averaged over all dataset examples as follows:

$$Accuracy\ (A) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{L}\,|\,\frac{h(d_i) \cap\ l_i}{h(d_i) \cup\ l_i}| \qquad (3)$$

B. Datasets

Multi-label classification problems appear in a wide range of real world situations and applications. The datasets that are included in the experimental setups cover three main application areas in which multi-label data is frequently observed: Text categorization, Multimedia classification and Bioinformatics. All datasets were mainly retrieved from the MULAN [22] data repository, as summarized in Table-3. It shows the domain data sets properties and their number of instances, attributes, labels and L_{Card}.

L_{Card} - means *label cardinality* which measures the average number of labels per test data. The L_{Den} measured [38],[45] for each datasets $D = \{ (d_n, L_k) | 1 \leq n \leq k\}$ are denoted as,

$$L_{Card} = \frac{1}{N}\sum_{i=1}^{n}|L_k| \qquad (4)$$

Table-3: Datasets Used for Experiment Evaluation with L_{Card}

Datasets	Domain	Instances	Attributes	Labels	L_{Card}
Birds	audio	645	260	19	1.014
Bibtext	text	7395	1836	159	2.402
Medical	text	978	1449	45	1.245
Scenes	images	2407	294	6	1.074
Yeast	Biology	2417	103	14	4.237
Genebase	Biology	662	1186	27	1.252
Enron	text	1702	1001	53	3.378

C. Algorithm Evaluated

The proposed method has been implemented in Java and integrated with Weka and Mulan [22] open-source Java libraries. We applied multi-label classification using the proposed *One-to-k Label* method on the standard multi-label classification methods are, BR, LP, CC, CLR, HOMER and RAKEL [21], [25]. The results of applying the proposed model using different multi-label classification methods on given datasets are compared to traditional multi-label learning without dividing the multi-label datasets. The experiments conducted with Weka and Mulan Libraries using the 10-fold cross-validation methodology.

5. EXPERIMENTAL RESULTS

This section presents the results of the evaluation experiments that were conducted. Initially, we learns the *One-Label* using OPT table and later discovered the *k*-Lebels using label association of datasets instances. The learned knowledge of *OT-k Label* classifier is compared with the traditional multi-label classifiers methods. The results obtained on applying on each datasets in Table-3 are presented in the following Tables.

Table-4: Number of OT-k Label Pairs Identified for the Classification

Datasets	Labels	Associated k-labels	Non-Associated	OT-k Label Classification Pairs
Birds	19	13	6	38
Bibtext	159	114	45	386
Medical	45	7	38	15
Scenes	6	3	3	8
Yeast	14	14	0	14
Genebase	27	11	16	24
Enron	53	30	23	81

Generated results are presented in Table-4. The identified results will be used for the classification with the traditional multi-label classifiers. The evaluation results comparison representing Hamming Loss (HL) and Accuracy are presented in Table-5 and Table-6 respectively and their individual comparison result is shown in Fig.3 and Fig.4.

Table-5: Classifier HL Performance (The Lower The Better)

Datasets	*OT*-k	BR	*OT*-k	LP	*OT*-k	*CC*	*OT-k*	*CLR*	*OT-k*	*HOMER*	*OT-k*	*RAKEL*
Birds	0.0462	0.0561	0.0599	0.0735	0.0452	0.0492	0.0417	0.0506	0.0627	0.0788	0.0437	0.0489
Bibtext	0.0125	0.0151	0.0117	0.0161	0.0098	0.0128	0.012	0.0144	0.0161	0.0182	0.0132	0.0151
Medical	0.0103	0.0113	0.0136	0.0135	0.0109	0.0122	0.0113	0.0116	0.0128	0.0135	0.0112	0.0115
Scenes	0.0841	0.0973	0.0951	0.1437	0.0994	0.1444	0.1094	0.1121	0.0951	0.1418	0.1012	0.0962
Yeast	0.2342	0.2454	0.251	0.2779	0.2421	0.2682	0.2151	0.2202	0.2425	0.2555	0.2258	0.2449
Genebase	0.0009	0.0011	0.0017	0.0019	0.0011	0.0017	0.0013	0.0013	0.0019	0.0021	0.0012	0.0013
Enron	0.0481	0.0508	0.0621	0.0717	0.0487	0.0524	0.0441	0.0471	0.0598	0.0584	0.0561	0.0635

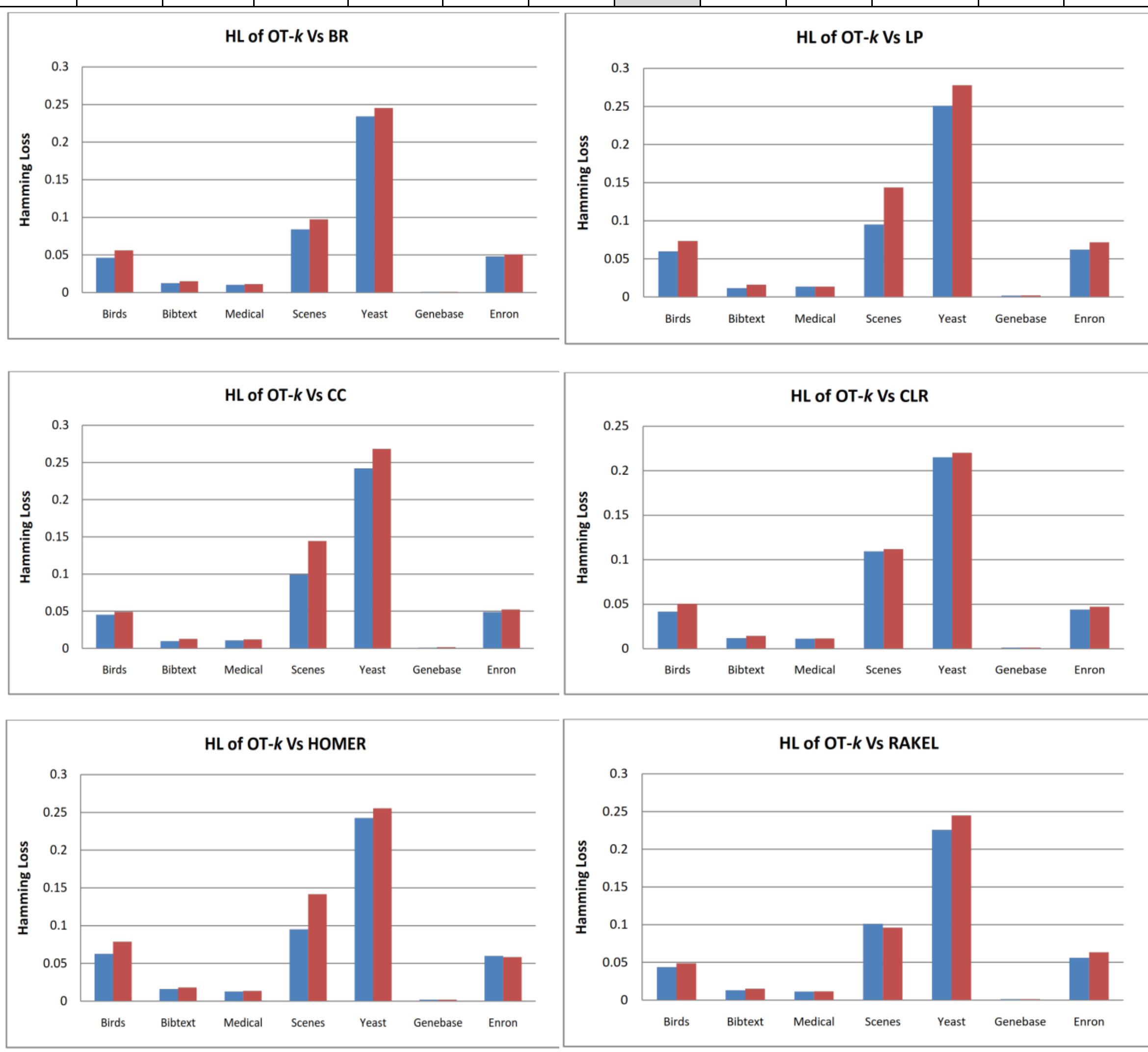

Fig. 3- Hamming Loss Individual Comparison

Table-6: Classifier Accuracy Performance
(The Higher The Better)

Datasets	*OT*-k	BR	*OT*-k	LP	*OT*-k	CC	*OT-k*	CLR	*OT-k*	HOMER	*OT-k*	RAKEL
Birds	0.6708	0.4666	0.7189	0.5295	0.7221	0.5241	0.7319	0.528	0.6731	0.4641	0.727	0.5452
Bibtext	0.7204	0.4187	0.6437	0.3869	0.6793	0.362	0.5015	0.4089	0.7182	0.3415	0.3854	0.3657
Medical	0.8573	0.7358	0.8637	0.7465	0.8624	0.7581	0.8734	0.773	0.8737	0.7741	0.856	0.7453
Scenes	0.7991	0.553	0.8391	0.5893	0.8265	0.5866	0.7918	0.5265	0.8391	0.5936	0.6247	0.6841
Yeast	0.4121	0.4395	0.4102	0.4144	0.418	0.4218	0.5198	0.5221	0.4021	0.4032	0.5021	0.5046
Genebase	0.9905	0.9826	0.9917	0.9862	0.9917	0.9862	0.9895	0.9857	0.9914	0.9845	0.9871	0.9794
Enron	0.4277	0.346	0.4851	0.4129	0.4955	0.4233	0.5434	0.4621	0.5001	0.4209	0.5317	0.4238

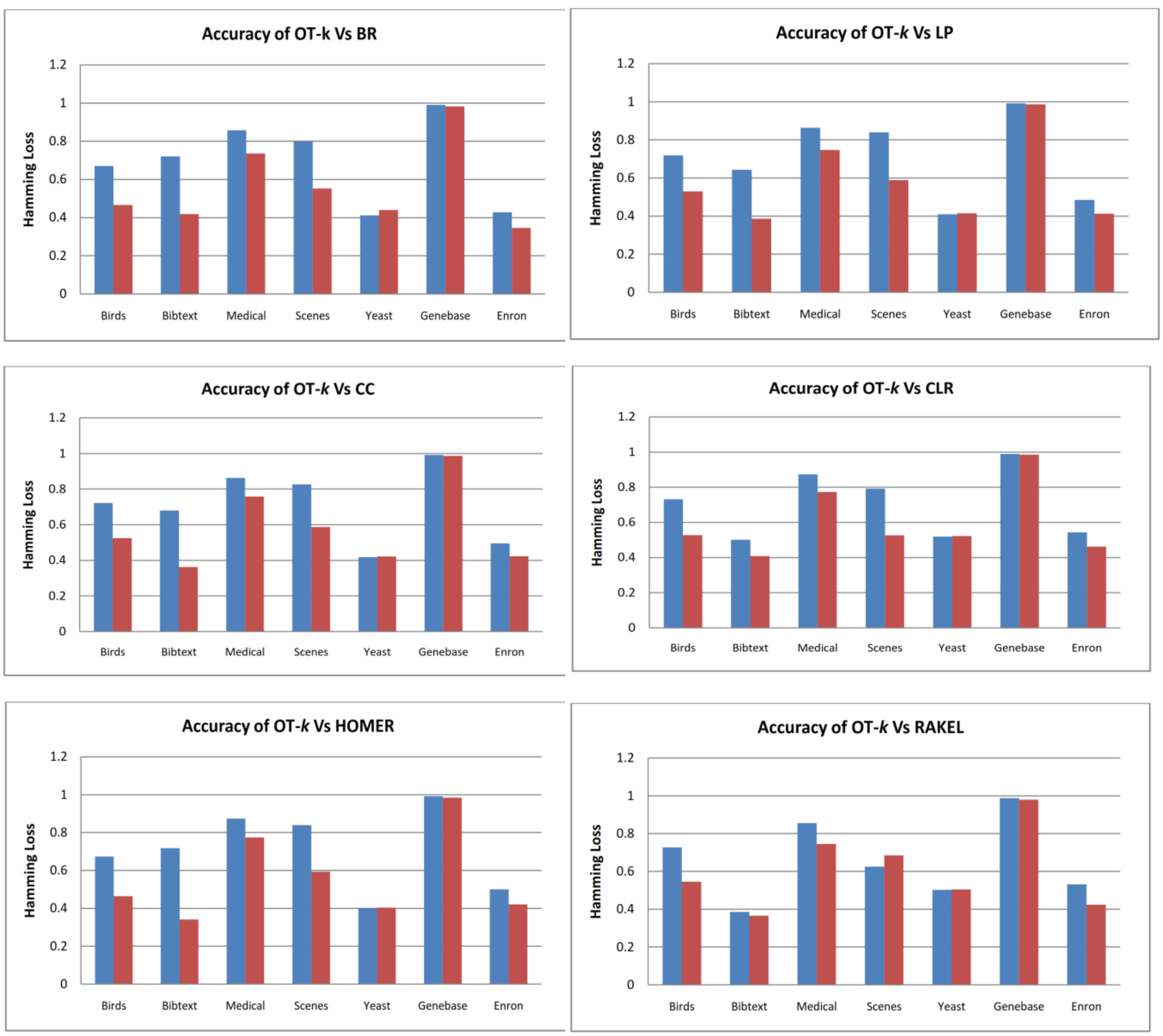

Fig. 4- Accuracy Loss Individual Comparison

Assessing results in terms of HL, we found improvements in one-third of cases. In many cases the difference in either direction is minimal. The significance of these differences is doubtful, but in any case, the improvement in very significant. But in comparison with "RAkEL" classifiers, it shows more hamming loss and low accuracy only with the "scenes" dataset as shown in Fig.3 and Fig.4. It is due to high number of misclassified labels for the data instances.

Similarly, assessing the accuracy a significant improvement is observed. Even the individual comparison result also shows OT-*k* has lower hamming loss and improvised accuracy in compared to all other classifiers.

6. CONCLUSION

In this paper, OT-*k* label learning classification is proposed using label association between *k*-labels. The learning process initially identifies the *One-Label* which is highly suitable for class suggestion and in second stage we find the other *k-labels* which supports the *One-Label* class for constructing a class patterns useful for different querying classification. To learn *One-Label* for an instance we compute the associated label density in compared with the OL-Prediction Table. The major contribution of our work is to utilizing the label for multi-label learning, which suggests a promising direction for learning from multi-label data classification. The experimental results obtained over several multi-label datasets with different classification algorithms, endorsed by the results from statistical tests, lead to the conclusion that it can be a useful approach for enhancing multi-label classification. In future, it can further investigated to exploits the label of association in addition with fuzzy and Bayes factors to get faster and enhancement in multi-label classification.

REFRENCES

[1]. M. Ling 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.

[2]. M. Ling 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.

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

[5]. 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.

[6]. 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.

[7]. 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.

[8]. A. McCallum, "Multi-label text classification with a mixture model trained by EM", presented at the AAAI'99 Workshop on Text Learning, Orlando, FL, USA, 1999.

[9]. T. N. Rubin, A. Chambers, P. Smyth, and M. Steyvers, "Statistical topic models for multi-label document classification", Mach. Learn., vol. 88, no. 1-2, pp. 157–208, 2012.

[10]. R. E. Schapire and Y. Singer, "Boostexter: A boosting-based system for text categorization", Mach. Learn., vol. 39, no. 2/3, pp. 135–168, 2000.

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

[12]. 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 24, J. Shawe-Taylor, R. S. Zemel, P. Bartlett, F. C. N. Pereira, and K. Q. Weinberger, Eds. Cambridge, MA, USA: MIT Press, 2011, pp. 190–198.

[13]. 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.

[14]. Z. Barutcuoglu, R. E. Schapire, and O. G. Troyanskaya, "Hierarchical multi-label prediction of gene function", Bioinformatics,vol.22,no.7, pp. 830–836, 2006.

[15]. N. Cesa-Bianchi, M. Re, and G. Valentini, "Synergy of multi-label hierarchical ensembles, data fusion, and cost-sensitive methods for gene functional inference", Mach. Learn., vol. 88, nos. 1/2, pp. 209–241, 2012.

[16]. A. Elisseeff and J. Weston, "A kernel method for multi-labelled classification", in Advances in Neural Information Processing Systems 14, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA, USA: MIT Press, 2002, pp. 681–687.

[17]. 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.

[18]. 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.

[19]. C. Sung 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.

[20]. J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification", Mach. Learn.,vol.85, no. 3, pp. 333–359, 2011.

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

[22]. G. Tsoumakas, E. Spyromitros-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]. 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.

[24]. D. Sasirekha and A. Punitha, "A Comprehensive Analysis on Associative Classification in Medical Datasets", Indian Journal of Science and Technology, Vol8(33),DOI:10.17485/ijst/2015/v8i33/80081, December 2015.

[25]. J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification", in Proceedings of tECML PKDD '09: Part II. Berlin, Heidelberg: Springer- Verlag, pp. 254–269, 2009.

[26]. 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.