ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(36), DOI: 10.17485/ijst/2015/v8i36/88616, December 2015* ISSN (Online) : 0974-5645

Investigation of Distance, Machine and Kernel Learning Similarity Methods for Visual Search in Content based Image Retrieval

Ananthi Sheshasaayee* and C. Jasmine

PG & Research Department of Computer Science, Quaid-E-Millath Government College for Women, Chennai - 600002, Tamil Nadu, India; ananthi.research@gmail.com, jasminesamraj.research@gmail.com

Abstract

Background/Objectives: The major objective of this work is to increase the retrieval accuracy of medical images by measuring the visual similarity of Content-Based Image Retrieval (CBIR) system. **Methods/Statistical Analysis:** This paper presents an On-line Multiple Kernel Similarity (OMKS) learning framework for performing classification based on kernel-based proximity functions. In accordance with many existing methods some other related issues are also discussed and retrieval performance evaluation of the existing and proposed OMKS learning strategy is also discussed. **Findings:** Several number of the distance based learning algorithm has been proposed in recent works. It is mainly aimed for the sake of measuring the visual similarity corresponding to the images. This paper involves in providing a comprehensive review of the technical achievements of distance learning, machine learning along with kernel learning methods for conducting visual similarity search. The methods cited have limitation in their capability of the similar measurement with complicated patterns in most practical applications. Similarity of multimodal data identified through the multiple resources[28], cannot be handled. **Application/Improvements:** Evaluation of the technique proposed for CBIR is performed on a huge amount of image data sets where motivating results indicate that OMKS performs better than the state-of-the-art techniques significantly. At last, based on OMKS technology and the rise of requisitions from practical-world applications, and idealistic future research directions have been identified as suggestions[29].

Keywords: Content-Based Image Retrieval (CBIR), Machine Learning Methods, Multiple Kernel Learning (MKL), Online learning, Similarity Search

1. Introduction

Due to the advance in Internet technology, and the image capturing devices like digital cameras and image scanners being available abundantly[32], size of the digital images collections has seen manifold increase. Effective image browsing, searching and retrieval tools are needed by users from different domains, which includes medicine, publishing, fashion, architecture, remote sensing, crime prevention, and other similar domains. Development of several general purpose image retrieval systems has been implemented regarding this. Two frameworks are identified namely text-based and content-based image retrieval. One major problem with CBIR is the task of to describe the image content as highly subjective. A variety of inconsistencies between user textual queries and image annotations or descriptions are always permissible[34].

Similarity search is a progressively significant task in the case of multimedia retrieval systems and has found extensive usage in the commercial and scientific applications in various scenarios, copying and near-duplicate detecting of images and videos[1-3] especially for the purpose of Content-Based Image Retrieval (CBIR)[4,5]. For satisfying the system requirements and also the user needs with respect to adaptable multimedia retrieval, development of CBIR similarity models have been carried

** Author for correspondence*

out and are extensively employed in the before mentioned areas[31]. One of the critical challenges is that the essential properties of the data objects are collected by means of feature representations that are exploited for comparing the individual data objects based on their components[31].

The various visual similarity searches with distance based similarity metrics have been proposed in earlier work to measure the similarity between query and image database. Adaptive similarity measures employ a known ground distance function for determining the distances between the feature signatures centroids observed in the feature space[31]. The ground distance functions that are applicable, for example are those which are assessed in[6]. An extension for colour-based image retrieval, the Perceptually Modified Hausdorff Distance[7] was introduced that utilizes the information regarding both the weights and centroids corresponding to the feature signatures. Another similarity measure is the well-known Earth Mover Distance[8] has seen its origin in the computer vision domain. This similarity measure defines the cost incurred during the transformation of one of the feature signature into another one[31]. Such Distance metrics are not suitable for measuring the proximity of images, because the use of rigid functions that are fixed.

To deal with this issue, studies of active research are being done in the Distance Metric Learning (DML) algorithms[9,10]. In order to get over these disadvantages of the already available work a new OMKS learning technique is proposed. It does the ranking of images through the learning of pairwise similarity in images which are specified in multiple modalities making use of multiple kernels[28].

2. State of Art

CBIR is described as an image search mechanism developed for finding images which are mostly similar to a query given. It matches text based retrieval by making use of computable and impartial image features in the form of the search benchmarks. Essentially, CBIR events the relationship existing between two images on the basis of the similar properties of their respective visual constituents inclusive of the color, texture, shape, and also the spatial arrangement of Regions of Interest (ROIs)[30]. The non-reliance of CBIR over labels renders it perfect for huge repositories where it is not practical to manually assign the keywords and other annotations[11].

The essential features that are applied by CBIR indicate that it is also feasible to reveal which images have the similarity and also reason why they are similar in an objective, non-qualitative way[11]. In most cases, the use of formal tools is required to derive high-level semantic features[30]. The main aim of supervised learning is to calculate the value of classification results based on a set of input values. The organization or clustering of the input data is described and in unsupervised learning there is no outcome measure[11].

Supervised learning such as Support Vector Machine (SVM)[12] and Bayesian classifier[13] is proposed in recent work for learning high-level concepts from low-level image features. Among these classifiers, SVM was used as one of the good candidate to learn a CBIR system in object detection, segmentation, text classification, etc. among the various hyper-planes, the Optimal Separating Plane (OSP) helps in maximizing the margin. In Bayesian classifier Binary based classification is a commonly used learning method[13] for learning high-level concepts from low-level image features. Here the selected input images are classified into two types as indoor/outdoor. The outdoor images are categorized into several classes such as city, landscape, etc[29].

It is stated that conventional learning algorithms have two problems: 1) A large amount of labelled training samples are needed, and it is very tedious sand error-prone to provide such data; 2) The training set is fixed during the learning and application stages. If the application domain changes new labelled samples have to be provided to ensure the effectiveness of the classifier[29]. These problems were overcome by using a bootstrapping approach is presented in[14]. It starts from a small set of labeled training samples and by using a co-training approach, where two classifiers that are statistically independent are employed for co-training and co-annotating the unlabelled samples, the algorithm successively annotates a larger set of unlabelled samples[10].

The distance metric method[15] was trained with the objective to the K-Nearest Neighbors (kNN) always belonging to the same class, whereas examples that are from different classes are isolated by a big margin. The margin criterion leading to a convex optimization on the basis of the hinge loss works by SVMs[36]. On the different data sets of changing size and hardship, this work finds that metrics which are trained in this manner result in considerable enhancements in kNN classification.

Relevant Component Analysis (RCA)[16] in the context of distance metric learning which forms an intermediate between PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis) in its use of labelled data. Specifically, RCA makes use "chunklet" message it is typically a subset of a class which increases classification results.

Boosting distance metric learning framework[17] aims at retrieving similar images from a database, the semantic annotations of whose could yield the medical expert with larger insight into all the interpretations that are possible of the query image. Likely on the other side, retrieval of images which appear artificial is same as the query could lead the users to an inaccurate diagnosis because the queries are unrelated semantically and undesirable. Therefore, learning distance metric needs to maintain both visual resemblance and their semantic similarity at the same time. The focus of the study is on the retrieval of medical image. The problem indicated in this work is significant in the case of several image retrieval systems. The Positive Semi Definite (PSD) is computationally intense particularly when the data is of high dimensionality and is also cast as a Semi Definite Programming (SDP), when the Distance Metric Learning (DML) task imposes restraint on the solution.

By proposing an Online Algorithm for Scalable Image Similarity (OASIS)[18] learning, it learns a bilinear similarity measure over sparse representations helps in solving most of the problems encountered in CBIR systems. An experiments result concludes that the proposed OASIS is accurate for different scales with hundreds and thousands of images. It is typically a linear metric learning technique but the complexity is more in this approach. These problems are solved by using kernel based distance learning methods.

In[19] kernel methods proposing an automatic image annotation the result of which is a unified framework which is inclusive of: 1) Multiple features used for image representation, 2) A feature integration and selection technique. 3) And an automatic strategy for semantic image annotation[33]. An elaborate experimental assessment that is demonstrated yields the efficiency to the new framework to construct image representations that are meaningful for the purpose of learning and helpful semantic annotations to be used for image retrieval.

In[20] a novel neural network based on the nonlinear Kernel Least Mean Square (KLMS). The proposed approach allows the users by selecting an initial query image and search incrementally target images via relevance feedback. If users aren't satisfied with the retrieved results, relevance feedback method enhances the performance of the proposed system by updating a boundary for separating relevant images from irrelevant ones[20].

In[21] proposed a new KPCA framework uses different kernelizing Mahalanobis distance learners to increase the speed of algorithm. In recent work[22] the relationship between metric and kernel learning for a bigger class. These methods are different from other techniques in two key facets. First and foremost, their design is such as to learn with a multiple kernels are used by the algorithm proposed for learning; and second, they generally run in an approach of batch learning that there is no scaling to large-scale applications. In contrast, the online learning algorithms that are used for learning a similarity function along with multiple kernels[21].

In[23] an efficient Simple NPKL algorithm was proposed. 1. This Simple NPKL algorithm is with linear loss. It helps in coming up with a closed-form solution which can be resourcefully computed by the Lanczos sparse Eigen decomposition technique. 2. Another point is Simple NPKL algorithm having other loss functions (that includes square hinge loss, hinge loss, square loss) which can be re-formulated in the form of a saddle-point optimization issue, and can be resolved further by a fast iterative algorithm. The proposed work is mainly aims learning a kernel function/matrix that is consistent with the constraints given[28].

In[24] an extension of the level technique that was actually developed for the optimization of non-smooth objective functions, to render convex-concave optimization, and then employ it to multiple kernel learning was proposed. The extended level technique gets over the difficulties of the already available techniques. Also it should be noted that this technical work was also motivated by the earlier work on OMKL[25]. The OMKL technique was introduced to learn the classifiers with optimal combination of multiple kernels[28]. The major aim of this OMKL learning classifier is to measuring the image similarity function from triplets for CBIR retrieval[26]. In specific, caution is required in the design with respect to several classifiers[27] with the triple constraints.

3. Experimentation Results

This section studies about an extensive collection of experiments used for evaluating the efficiency of the algorithms proposed for visual similarity search in CBIR have been conducted. The data sets and source code implementation of the experiments can be seen in the project website (http://OMKS.stevenhoi.org/). Adopt five generally available image data sets from http://omks.stevenhoi.org/. The benchmark of tasks related to image retrieval, classification and recognition widely uses these set of data values[28]. The performance of various algorithms in an individual manner on a stochastically picked subset were then evaluated public spaces (name it "Public," it is also utilized for evaluating the impact of parameters) along with the entire indoor collection[26]. For every data set (except for "Oxford" which has no categorical info), randomly choose a subset from every class to assure that all the classes are given with the same number of images as the one which has got the least number images in the actual data set. This can prevent the performance to be dominated by some unknown single class of huge number of images. Based on the data set, thereafter select randomly 50 percent examples from every class to create a training set, 10 percent examples to generate a validation set, 10 percent examples to make a query set, and the remaining 30 percent examples to make up the test set for the purpose of retrieval evaluation[28]. The Caltech10 database with different samples with six category names for the given queries is as follows: 1 (roulette wheel), 2 (billiard), 3 (skyscraper), 4 (bear), 5 (Minotaur), 6 (laptop) are shown in Figure 1[28].

Figure 1. "Caltech10" database by various algorithms with category names for the queries are as shown below: 1 (roulette wheel), 2 (billiard), 3 (skyscraper), 4 (bear), 5 (Minotaur), 6 (laptop).

The standard mean Average Precision (mAP) has been adopted to assess the retrieval outcome. The mAP value is computed on the basis of the average AP value of all of the queries[28]. The precision value is defined as the ratio of examples that are relevant over the total number of retrieved examples, whereas recall is defined as the ratio of the relevant examples that are retrieved over the total number of relevant examples in the database. The performance evaluation results given in Table 1 are measured using the following methods Euclidean Distance Based Learning (EDBL), Support Vector Machines (SVM)[28], Relevant Component Analysis (RCA), OASIS (Online Algorithm for Scalable Image Similarity), KLMS (Kernel Least Mean Square) and OMKS (On-line Multiple Kernel Similarity) the performance is measured in terms of top-n (n = 1, 2, ..., 5) precision and the mAP values. Figure 2 and Figure 3 illustrates the information of the top-n precision outcomes on two sampled data sets. In this graph we compare the results of the methods EDBL, SVMs, RCA, OASIS, KLMS and OMKS for Public dataset[28]. It shows that the proposed OMKS results have higher precision values for public and Caltech dataset is illustrated in Figure 2 and Figure 3.

Table 1. Experimental Results of mAP Performance

Algo-rithm	Met-ric	Public	Indoor	Caltech 5000	Image CLEF+	Oxford
EDBL	mAP	0.1629	0.0441	0.1836	0.4128	0.4365
	SD	0.00175	0.00075	0.0048	0.0108	0.0000
SVM	mAP	0.1632	0.0443	0.1839	0.4129	0.4368
	SD	0.0018	0.0008	0.00485	0.0110	0.0015
RCA	mAP	0.1635	0.04435	0.1841	0.4132	0
	SD	0.0019	0.00085	0.00492	0.0112	0
OASIS	mAP	0.1683	0.0462	0.1843	0.4532	0.6078
	SD	0.0049	0.0009	0.0059	0.0121	0.0422
KLMS	mAP	0.1758	0.0462	0.1912	0.1965	0
	SD	0.0049	0.0028	0.0078	0.0110	0
OMKS	mAP	0.2085	0.0628	0.3278	0.4848	0
	SD	0.0185	0.0039	0.0064	0.0285	0

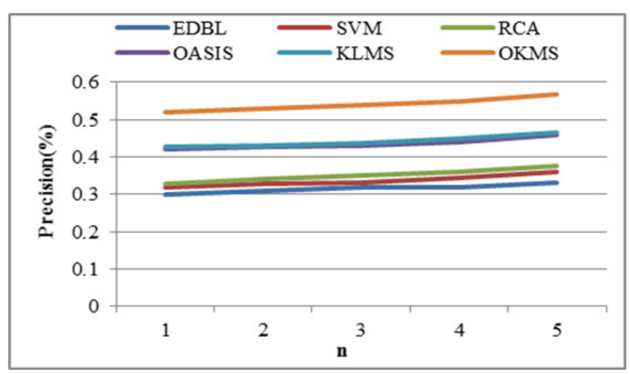

Figure 2. Top-n precision results over "Public" dataset vs. methods.

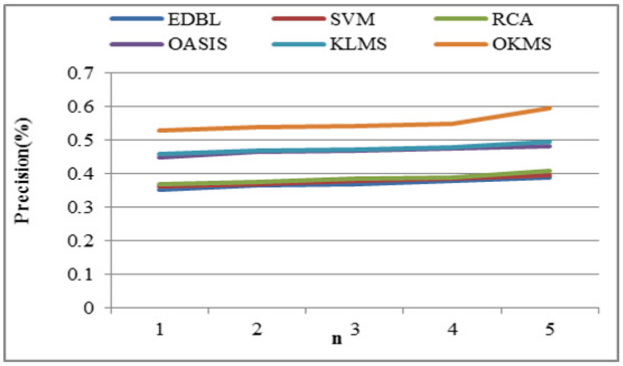

Figure 3. Top-n precision results on "Caltech" dataset vs. methods.

Figure 4 and Figure 5 reveals the results of the evaluation under different values of RatioTrain that are utilized for constructing the similarity functions on the two sampled data sets[29]. The results of all the algorithms that are compared share the similar kind of performance trend when the number of training triplets sees an increase. Particularly, the more the value of RatioTrain, the better is the retrieval performance that can be accomplished by the learning algorithms. Moreover, when RatioTrain is huge enough, for example, over 40 percent, many of the learning algorithms attempt to go smaller as an improvement that is chiefly attributed to adequately big amount of training data.[28] Eventually, the earlier experiments were performed, for every case that is under different values of RatioTrain, the OMKS algorithm proposed can outperform significantly compared to the other contending algorithms[28].

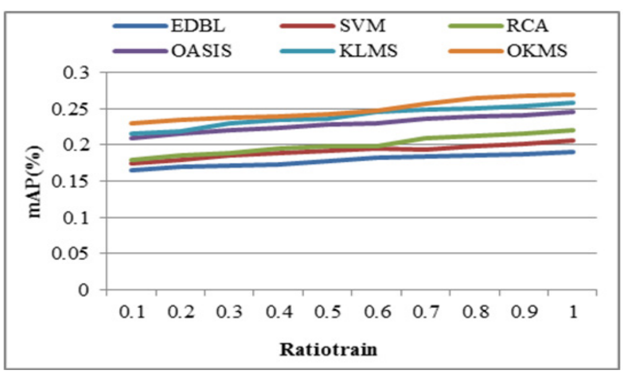

Figure 4. Evaluation corresponding to RatioTrain on "Public" data sets vs. methods.

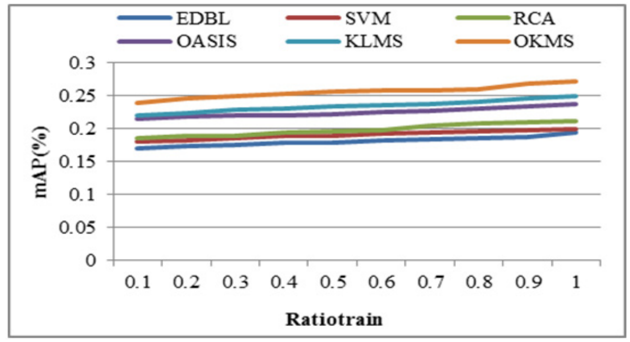

Figure 5. Evaluation of Ratio Train on "Caltech" data sets vs. methods.

5. Conclusion and Future Work

This paper discusses the problem of existing distance and kernel based learning algorithms for ranking images toward visual similarity search[28]. The limitations of traditional distance learning and kernel based learning algorithm for the purpose of ranking the images with respect to visual similarity search methods were overcome by a proposed novel OMKS learning method. By examining the ability of multiple kernels in merging the multimodal data, OMKS is much more efficient for improving retrieval results in CBIR. Designed a resourceful OMKS algorithm and have widely assessed the algorithms proposed for the cause of image similarity search on huge databases which is publicly available[28]. The major problem of the work it is easily applicable to public image databases, but applying this methods to medical image databases it becomes not practicable since noise in the images samples are not removed. Therefore it reduces the learning results for the purpose of visual similarity search. The future work can be extended to medical applications that address both theoretical and real-time challenges of the OMKS framework proposed for the use of large-scale applications[28]. Noise in the image samples can be removed in pre-processing stage, important features can be selected for ranking visual similarity search, and relevance feedback learning based content based image retrieval can also be extended additionally to enhance medical image applications in content based image retrieval in medical applications.

6. References

1. Wu X, Ngo CW, Hauptmann AG, Tan HK. Realtime near-duplicate elimination for web video search with content and context. IEEE Transactions on Multimedia. 2009; 11(2):196–207.
2. Yang X, Zhu Q, Cheng KT. Near-duplicate detection for images and videos. Proceedings of ACM Workshop on Large-scale Multimedia Retrieval and Mining; 2009. p. 73–80.
3. Law-To J, Chen L, Joly A, Laptev I, Buisson O, Gouet-Brunet V, Boujemaa N, Stentiford F. Video copy detection: a comparative study. Proceedings of ACM International Conference on Image and Video Retrieval; France. 2007. p. 371–8.
4. Datta R, Joshi D, Li J, Wang JZ. Image retrieval: Ideas, Influences, and Trends of the New Age. ACM Computing Surveys, 2008; 40(2):1–60.
5. Geetha P, Narayanan V. A survey of content-based video retrieval. Journal of Computer Science. 2008; 4(6):474–86.
6. Hu R, Ruger SM, Song D, Liu H, Huang Z. Dissimilarity measures for content-based image retrieval. Proceedings of IEEE International Conference on Multimedia and Expo; Hannover. 2008. p. 1365–8.
7. Park BG, Lee KM, Lee SU. Color-based image retrieval using perceptually modified Hausdorff distance. Journal on Image and Video Processing. 2008; 4:1– 10.
8. Rubner Y, Tomasi C, Guibas LJ. The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision. 2000; 40(2):99–121.
9. Hoi SCH, Liu W, Chang SF. Semi-supervised distance metric learning for collaborative image retrieval and clustering. ACM Transaction on Multimedia Computer Communication Applications. 2010; 6(3):18–26.
10. Lee JE, Jin R, Jain AK. Rank-based distance metric learning: An application to image retrieval. Proceedings of IEEE Conference Computer Vision and Pattern Recognition; USA. 2008. p. 1–8.
11. Lew MS, Sebe N, Djeraba C, Jain R. Content-based multimedia information retrieval: State of the art and challenges. ACM Transaction on Multimedia Computer Communication Application. 2006; 2(1):1–19.
12. Shi R, Feng H, Chua TS, Lee CH. An adaptive image content representation and segmentation approach to automatic image annotation. International Conference on Image and Video Retrieval (CIVR); Berlin, Heidelberg: Springer; 2004. p. 545–54.
13. Jin W, Shi R, Chua T-S. A semi-naive Bayesian method incorporating clustering with pair-wise constraints for auto image annotation. Proceedings of the ACM Multimedia; China. 2004. p. 336–9.
14. Feng H, Chua TS. A bootstrapping approach to annotating large image collection. Workshop on Multimedia Information Retrieval in ACM Multimedia; Singapore. 2003. p. 55–62.
15. Weinbergerand KQ, Saul LK. Distance Metric Learning for Large Margin Nearest Neighbor Classification. Journal of Machine Learning Research. 2009; 10(2009):207–44.
16. Bar-Hillel T, Hertz N, Shental, Weinshall D. Learning a Mahalanobis metric from equivalence constraints. Journal of Machine Learning Research. 2006; 6(1):937–65.
17. Yang L, Jin R, Mummert LB, Sukthankar R, Goode A, Zheng B, Hoi SCH, Satyanarayanan M. A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. IEEE Transaction Pattern Analysis and Machine Intelligence. 2010; 32(1):30–44.
18. Chechik G, Sharma V, Shalit U, Bengio S. Large scale online learning of image similarity through ranking. Journal of Machine Learning Research. 2010; 11(2010):1109–35.
19. Caicedo JC, Gonzalez FA, Romero E. Content-based histopathology image retrieval using a kernel-based semantic annotation framework. Journal of Biomedical Informatics. 2011; 44(4):519–28.
20. Modaghegh H, Javidi M, Yazdi HS, Pourreza HR. Learning of relevance feedback using a novel kernel based neural network. Australian Journal of Basic and Applied Sciences. 2010; 4(2):171–86.
21. Chatpatanasiri R, Korsrilabutr T, Tangchanachaianan P, Kijsirikul B. A new kernelization framework for mahalanobis distance learning algorithms. Neurocomputing. 2010; 73(10–12):1570–9.
22. Jain P, Kulis B, Davis JV, Dhillon IS. Metric and kernel learning using a linear transformation. Journal of Machine Learning Research. 2012; 13(1):519–47.
23. Zhuang J, Tsang I W, Hoi SCH. A family of simple non-parametric kernel learning algorithms. Journal of Machine Learning Research. 2011; 12(2011):1313–47.
24. Xu Z, Jin R, King I, Lyu MR. An extended level method for efficient multiple kernel learning. Proceedings of Advances in Neural Information Processing Systems; Vancouver, BC, Canada. 2008. p.1825–32.
25. Jin R, Hoi SCH, Yang T. Online Multiple Kernel Learning: Algorithms and Mistake Bounds. Proceedings of 21st International Conference Algorithmic Learning Theory (ALT'10); Singapore. 2010. p. 390–404.
26. Giri A, Meena Y. Review paper on content based image retrieval. The International Journal of Science and Technology. 2014; 2(4):125.
27. Sankari JM, Shoba LK. Mining student's social data to understand their learning experience. The International Journal of Science and Technology. 2014. 2(12):143.
28. Xia H, Hoi SCH, Jin R, Zhao P. Online multiple kernel similarity learning for visual search. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2014; 36(3):536–49.
29. Available from: www.citeseerx.ist.psu.edu
30. Kumar A, Kim J, Cai W, Fulham M, Feng D. Content-based medical image retrieval: A survey of applications to multidimensional and multimodality data. Journal of Digital Imaging. 2013.
31. Available from: dme.rwth-aachen.de
32. Raveaux R, Burie J-C, Ogier J-M. Structured representations in a content based image retrieval context. Journal of Visual Communication and Image Representation. 2013.
33. Available from: www.informed.unal.edu.co

34. Available from: www.cse.msu.edu
35. Available from: www.cfsites1.uts.edu.au
36. Available from: www.machinelearning.wustl.edu
37. Jayamangala H, Sheshasaayee A. A review on models and algorithms to achieve influence maximization in social networks. Indian Journal of Science and Technology. 2015.