Submitted in partial fulfillment of the requirements of

BITS G513 Study in Advance Topic(SAT)

BY

RASHMI GULHANE

ID NO: 2015H112187P

Under the supervision of

Dr. POONAM GOYAL

Assistant Professor CSIS

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI, PILANI CAMPUS

(Dec 2016)

Acknowledgements

First of all I liked to express my Sincere gratitude to my supervisor **Dr. Mrs. POONAM GOYAL** for giving me this opportunity to work under her expert guidance and helping me explore this field.

I would also like to thank **Mrs.chandramani choudhary** for her valuable suggestion.

I am also thankful to CSIS Department, BITS Pilani for providing the opportunity to take this Course.

Table of Contents:

1. Acknowledgements……………………………………………………………………………1
2. Abstract……...………………………………………………………………………………2
3. Introduction...………………………………………………………………………………...4
4. Framework..…………………………………………………………………………………5
    4.1. Tag relevance formulation..……………………………………………………………6
        4.1.1. Visual neighbor search..………………………………………………………..6
        4.1.2. Tag relevance function..………………………………………………………...6
    4.2. Ranking-oriented learning..……………………………………………………………7
5. Experimental Configuration..……………………………………………………….......9
6. Conclusion..……………………………………………………………………………...10
7. Acknowledgement..……………………………………………………………………...11
8. References..……………………………………………………………………………...11

2.Abstract

Image tags are very important in online application such as browsing,sharing and social media for retrieving images.When user tags a image,the tag can be vague and imperfect as a result it does not matches the content of image which reduce the rendering of current image search.To provide a better result of image search query we link content of image to its tag and correct tag which are inappropriate for this purpose we do tag relevance learning.We give weights to neighbor depended on the tags it has and then determine tag of the image so a risk of making a wrong belief about image tag is avoided.We have done this task of ranking by using both supervised as well as unsupervised machine learning technique.

3.Introduction

The number of images on INTERNET has exploded in recent days so a good search technique is required to explore image collection in recent years,The tags given to image by user can be used to search index and search this images.As a result tag based image search is popularly used to retrieve image in social database.Many a times tag given by user is not accurate as user does not have that time and subject knowledge to name the image accurately.Because of this inaccuracy we cannot us the tag given by user as it is to retrieve images from search.So we need to establish relevance of the tag with respect to the visual content of image which will help us to improve the tagging process this is known as problem of tag relevance learning.

To solve this problem a classifier is applied which learns the relevance of each tag with the content of the image and give a classifier score.But this sort of classification has inefficiencies as the number of tags are huge and selection of training example is an issue.So we Apply unsupervised method for tagging. one of the most popular Approach used is KNN but the performance provided by this approach is also limited as it does not intend to quality of Tags given to an image.Motivated by this the author provides a new approach of tag relevance learning.Both supervised and unsupervised Approach is used.The Author Aims to improve the tagging process by using the visual content of images and the tags of neighboring kth images.The architecture of proposed framework is as follow:

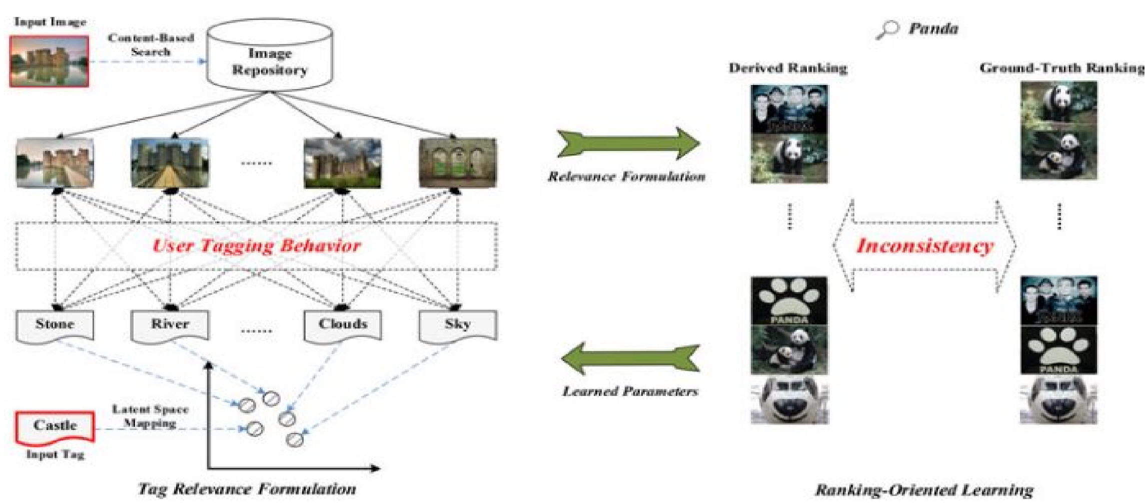

Fig.1.Architecture of proposed framework

4.Framework

The architecture of the system is shown in fig 1.It consists of two main part The first is tag relevance formulation and the second is ranking oriented learning.We Explain both the step in details in further part.For better understanding we list all the symbols and their definition in table 1.

Table 1 Summary of symbols and definitions

Symbols	Definitions
$\mathcal{X}$	Social image collection
$\mathcal{T}$	Tag vocabulary
x	Tagged image, $x \in \mathcal{X}$
t_i	i-th tag in the vocabulary, $t_i \in \mathcal{T}$
z_l^x	l-th visual neighbor of x, $z_l^x \in \mathcal{X}$
v_l	Weight of z_l^x
w_{ij}	Correlation between t_i and t_j
s_{t_i}	Number of images tagged with t_i
s	Number of total images, $s = \|\mathcal{X}\|$
m	Number of unique tags, $m = \|\mathcal{T}\|$
k	Number of visual neighbors
p	Dimension of the latent space
$\mathbf{u}_i$	Representation vector of t_i in the latent space
$\mathbf{e}_i$	Unit vector with 1 in the i-th position
$\mathcal{Y}$	Set of all possible rankings over images
Y	Ranking over images, $Y \in \mathcal{Y}$
Y_q^*	Ground-truth ranking of the images with respect to t_q, $Y_q^* \in \mathcal{Y}$
$\mathcal{I}_{t_q}^+$	Set of the relevant images with respect to t_q
$\mathcal{I}_{t_q}^-$	Set of the irrelevant images with respect to t_q
n	Number of training instances
d	Dimension of image feature vector

Table 1.Consist of symbols Table

4.1. Tag relevance formulation

4.1.1. Visual neighbor search

To find the visual neighbour we have used machine learning technique which uses the multi-modal information associated with an image.Finding visual neighbor is a two step process 10 feature extraction 2)similarity measure.In the first step we use 5 type of low-level features to represent images which includes histogram of dimension 64,color correlation of 144 dimension,edge direction histogram of 73 dimension,wavelet texture of 128 dimension,block wise color moment of 225 dimension.This are standard features provided in NUS-WIDE-LITE[4] and image is represented in different form such as color,shape and texture.This features are combined into different categories 1) Late fusion - it learns each feature one by one before giving tag relevance score and then integrates them,As we learn each feature individually the cost associated with it is too high 2) Early fusion - integrates all the feature and then learn the tag relevance score.As range of all the features are different we first normalize each feature in the range of [0,1] and then merge then up.We find the K nearest neighbor of each image by finding the euclidean distance.

4.1.2.Tag relevance function

The author has proposed a framework which construct a tag relevance function using the neighbor voting scheme where a tag of an image is given score by considering the tagging information of visual neighbor.The equation(eq 2) for tag relevance function is as follow:

$$r(t_i, x) = w_{ii} \sum_{l=1}^{k} v_l \varphi(z_l^x, t_i) + \sum_{j=1, j \neq i}^{m} w_{ij} \sum_{l=1}^{k} v_l \varphi(z_l^x, t_j) ,$$

The equation means for each image we calculate the relevance score by considering the weight assigned to k nearest neighbor with respect to that tag and weight assigned to all other k nearest neighbor with respect to all other tags.W is a symmetric matrix and values less than and greater than zero are allowed in this matrix.Equation 2 can be written in the following form to form eq.3.

$$r(t_i, x) = \mathbf{e}_i{}^T W \Phi_x{}^T \mathbf{v} .$$

As there are many parameters involved and too many parameter may decrease the model stability.So we convert the representation in latent space.By converting the above equation(eq 3) to latent space we get a new formulation as follow :

$$r(t_i, x) = \mathbf{e}_i{}^T U^T U \Phi_x{}^T \mathbf{v} ,$$

4.2 Ranking-oriented learning

Our Approach is designed to learn two parameter v and U in a supervised fashion.We learn this parameter by applying two machine learning algorithm cutting plane algorithm and pegasos algorithm.

4.2.1 Cutting plane algorithm:

This Algorithm[1] is a higher version of SVM and is used for multiclass problems.This algorithm finds a set of constraint so that this constraints can provide solution filling all the constraints at a error tolerance of e. The algorithm of cutting plane is given in Algorithm 1.It starts with an empty set of w.Then it iteratively finds the most violated constraint for each tag.The algorithms stops when it cannot find any constraints to be added into constraint set.

Algorithm 1 Cutting Plane Algorithm

Input: training instances $(t_1, Y_1^*), \ldots, (t_n, Y_n^*)$, regularization trade-off λ, error tolerance ε
Output: model parameters $\mathbf{v}$ and U, slack variable ξ
1: Initialize $\mathcal{W} \leftarrow \emptyset$
2: **repeat**
3: **for** $q = 1, 2, \ldots, n$ **do**
4: $\widehat{Y}_q \leftarrow \arg\max_{Y \in \mathcal{Y}} \Delta(Y_q^*, Y) + f(t_q, Y)$
5: **end for**
6: **if** $\frac{1}{n} \sum_{q=1}^{n} \Delta(Y_q^*, \widehat{Y}_q) - \frac{1}{n} \sum_{q=1}^{n} \left[f(t_q, Y_q^*) - f(t_q, \widehat{Y}_q) \right] > \xi + \varepsilon$ **then**
7: $\mathcal{W} \leftarrow \mathcal{W} \cup \left\{ (\widehat{Y}_1, \ldots, \widehat{Y}_n) \right\}$
8: Optimize Eq. (12) over $\mathcal{W}$
9: **end if**
10: **until** $\mathcal{W}$ has no change during iteration

Algorithm 1:Cutting plane Algorithm

The Optimizing function which is applied is as follow(eq.12):

$$
\begin{aligned}
&\min_{\mathbf{v},U,\xi} \quad \frac{\lambda}{2}\|\mathbf{v}\|_2^2 + \frac{\lambda}{2}\|U\|_F^2 + \xi \\
&\text{s.t.} \qquad \forall(Y_1,\ldots,Y_n) \in \mathcal{Y}^n : \\
&\frac{1}{n}\sum_{q=1}^{n}\Big[f(t_q, Y_q^*) - f(t_q, Y_q)\Big] \geqslant \frac{1}{n}\sum_{q=1}^{n}\Delta(Y_q^*, Y_q) - \xi \ .
\end{aligned}
$$

This function is different as compared to original structured SVM the standard regularization term eqn(12) and use the Frobenius norm.The slack variable is the only one which is shared across all the constraints.The optimization function constraint the requirement that true ranking average score should be greater than all possible ranking.The other issue is to reoptimize when a new constraint will be added for that we applied pegasos algorithm(Algorithm 2)[2].

Algorithm 2 Pegasos Algorithm

Input: working set $\mathcal{W}$, regularization trade-off λ, initialization values $\mathbf{v}_1$ and U_1, maximum iteration T

Output: model parameters $\mathbf{v}_{T+1}$ and U_{T+1}

1: **for** $t = 1, 2, \ldots, T$ **do**

2: Find $(\hat{Y}_1, \ldots, \hat{Y}_n) \in \mathcal{W}$ achieving the largest margin

3: Set $\eta_t = 1/(\lambda t)$

4: Compute $\nabla_{\mathbf{v}_t}$ and ∇_{U_t}

5: Update $\mathbf{v}_{t+1} = \mathbf{v}_t - \eta_t \nabla_{\mathbf{v}_t}$
$U_{t+1} = U_t - \eta_t \nabla_{U_t}$

6: Project $\mathbf{v}_{t+1} = \min\left\{1, \frac{1/\sqrt{\lambda}}{\|\mathbf{v}_{t+1}\|_2}\right\} \mathbf{v}_{t+1}$
$U_{t+1} = \min\left\{1, \frac{1/\sqrt{\lambda}}{\|U_{t+1}\|_F}\right\} U_{t+1}$

7: **end for**

Algorithm 2:Pegasos Algorithm

7 . Experimental configuration

We used the dataset NUS-WIDE-LITE[4] which has 55615 images with these tags.Preprocessing is applied to filter out the tags as the tag set is noisy.For training we consider half of the tags and the rest half for testing.Mean Average precision(MAP) is used as an evaluation measure.

8.Conclusion:

We have used both supervised and unsupervised approach for tag relevance learning thereby giving better performance than using either one.we have applied knn for finding the kth nearest neighbor then have applied a relevance function which considers pairwise correlation between kth nearest neighbor.For calculating pairwise correlation we need optimizing parameter which we get by applying Pegasos algorithm.

9.Acknowledgments:

This work is accomplished by using concept from paper[5] and algorithm use are applied from other research paper in the reference.

10.References:

[1] Joachims T, Finley T, Yu C (2009) Cutting-plane training of structural svms. Mach Learn 77(1):27–59

[2]. Shalev-Shwartz S, Singer Y, Srebro N, Cotter A (2011) Pegasos: Primal estimated sub-gradient solver for svm. Math Program 127(1):3–30

[3]Tang J, Hong R, Yan S, Chua TS, Qi GJ, Jain R (2011) Image annotation by knn-sparse graph-based label propagation over noisily tagged web images. ACM Transactions on Intelligent Systems and Technology 2(2):14:1–14:15

[4]Chua TS, Tang J, Hong R, Li H, Luo Z, Zheng Y (2009) Nus-wide: a real-world web image database from national university of singapore. In: Proceedings of ACM International Conference on Image and Video Retrieval, pp 48:1–48:9

[5]Chaoran Cui1 · Jialie Shen2 · Jun Ma3 ·Tao Lian3(2016) Social tag relevance learning via ranking-oriented neighbor voting.In : Processding of Multimed Tools Appl.