

EDIFICATION TRAINING FOR PARTICIPATING IN VARIOUS ACTIVITIES THROUGH ONLINE

K.M. Hemavathy^{1*}, M. Beemamehraj²

¹*Computer Science Engineering Bharath University Chennai, India,

²Assistant Professor Computer Science Engineering Bharath University Chennai, India

¹*hemram2011@gmail.com, ²beemamehraj@gmail.com

***Corresponding Author: -**

Email ID - hemram2011@gmail.com

Abstract: -

This paper analyze the specification of online multitask learning for recovering various classification process that is in parallel related, focusing at every part of data received by each accurately and efficiently. Statistical machine translation systems are usually trained on large amounts of micro-blog sentiment detection on a faction of users, which classifies micro-blog posts which are generated by each user into expressive or non-expressive categories. This particular online learning task is challenging for a number of reasons. To achieve the major requirement of online applications, a highly efficient and scalable problem that can give sudden assumption with low learning cost. This requirement leaves conventional batch learning algorithms out of consideration. Then, novel organization methods, be it group or online, often encounter a dilemma when applied to a group of process, i.e., on one hand, a single classification model trained on the entire collection of data from all tasks may fail to capture characteristics of individual task; on the other hand, a model trained separately on individual tasks may suffer from insufficient training data. To rectify this problem in this paper, we propose an Edification training for participating in various activities through online, from this we can geographically model over the entire data of all process. Another part individual model for various related process is combined inferred by to make cost effective in the global model through a Edification training via online approach. We defined the effectiveness of the proposed system on a synthetic dataset. Here the evaluation had done three real-life problems bioinformatics data classification, spam email filtering, and micro-blog sentiment detection.

Keywords: - on-line learning, Artificial intelligence, multitask learning, classification



Distributed under Creative Commons CC BY-NC 4.0 OPEN ACCESS

I. INTRODUCTION

Classical machine learning methods are often formulated as a single task learning problem, which by definition learns one task at a time. On the contrary, multitask learning aims to solve multiple related learning tasks in parallel. Many real-world problems are essentially multitasked learning, although they are often broken into smaller single learning tasks, which are then solved individually by classical learning methods.

Multitask learning has been extensively studied in machine learning and data mining over the past decade. Empirical findings have demonstrated the advantages of multitask learning over single task learning across a variety of application domains. The classical multitask learning methodology often makes two assumptions. First, it assumes there is one primary task and other related tasks are simply secondary ones whose training data are exploited by multitask learning to improve the primary task. Thus, the classical multitask learning approach focuses on learning the primary task without caring how the other tasks are learned. Second, the classical multitask learning problem is often studied in a batch education setting, which assumes that the preparation data of all tasks are available. On one hand, this assumption is not realistic for many real-world problems where data arrives sequentially. On the other hand, the batch multitask learning algorithms usually have fairly intensive training cost and poor scalability performance, as far as large real applications is concerned.

II RELATED WORKS

[2] In this paper two named-entity recognition models which use characters and character n -grams either exclusively or as an important part of their data demonstration. The first model is a character-level HMM with minimal perspective information, and the second model is a maximum-entropy conditional markov model with considerably richer context features. Our best model achieves an overall F1 of 86.07% on the English test data (92.31% on the development data). This number represents a 25% error reduction over the same model without word-internal (substring) features.

[14] This paper, have shown how to extract a hypothesis with small risk from the ensemble of hypotheses generated by an arbitrary on-line learning algorithm run on an independent and identically distributed example of data. Using a simple large difference argument, we confirm tight data-dependent bound for the risk of this hypothesis in terms of an easily computable statistic associated with the on-line performance of the ensemble. Via sharp point wise bounds on, we then obtain risk tail bounds for kernel Perceptron algorithms in terms of the spectrum of the experimental kernel matrix. These boundaries reveal that the linear hypotheses found via our approach achieve optimal tradeoffs between hinge loss and margin size over the class of all linear function, an issue that was left open by prior results. A unique feature of our approach is that the key tools for our analysis come from the model of prediction of individual sequence; i.e., a copy making no probabilistic assumptions on the source generates the data. In fact, these tools twist out to be so powerful that we only need very elementary statistical facts to obtain our final risk bounds.

[3] LIBSVM is a library for Support Vector Machines (SVMs). We have been actively developing this package ever since the year of 2000. The goal is to help users to easily apply SVM to their application. LIBSVM gain wide popularity in machine learning and a lot of other areas. In this editorial, we present all implementation detail of LIBSVM. Issues such as solve SVM optimization problems, theoretical convergence, probability estimates, multi-class classification, and parameter selection are discussed in detail.

[4] Statistical machine translation systems are usually trained on large amounts of bilingual text and monolingual text in the target language. In this paper we explore the use of transductive semi-supervised methods for the effective use of monolingual data from the source language in order to improve conversion quality. We intend several algorithms with the intend, and at hand the strengths and weakness of everyone. We present detailed experimental evaluations on the French–English Europol data set and on data from the NIST Chinese–English large data way. We confirm a significant progress in translation quality on both tasks.

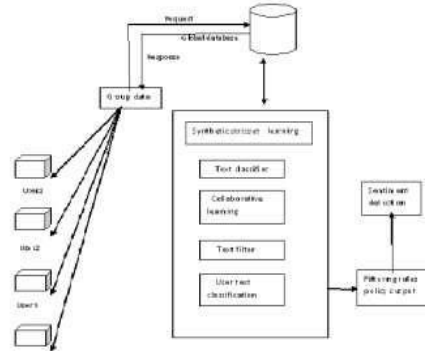
[6] This paper presents a new technique called “t-SNE” that visualizes high-dimensional data by giving each data point a location in a two or three-dimension map. The technique is a variation of Stochastic Neighbor Embedding that is much easy to optimize, and produce significantly better visualizations by reducing the tendency to crowd points together in the midpoint of the map. t-SNE is better than presented techniques at creating a single map that reveals structure at many special scales. This is particularly important for high-dimension data that lie on several dissimilar, but linked, low-dimensional manifolds, such as images of objects from several classes seen from several viewpoints. For visualizing the structure of very outsized data sets, we show how t-SNE can use accidental walks on neighborhood graphs to allow the implicit structure of all of the data to influence the way in which a subset of the data be displayed. We demonstrate the performance of t-SNE on a wide variety of data sets and compare it with many other non-parametric visualization techniques, including Isomap, Samsmon mapping, and Locally Linear Embedding. The visualization produced by t-SNE are significantly better than those produced by the other techniques on almost all of the data sets.

III IMPLEMENTATION OF COML BY ONLINE LEARNING METHODS

We investigate the problem of online multitask learn, which differs from the classical multitask learn in two aspects. For most, our objective is to improve the learning performance of all tasks instead of focusing on a single main task. Second, we structure the multitask learning problem in an online learning setting by assuming that the data for each task arrives in sequence, which is a more reasonable scenario for real-world application. Unlike group learning technique, online learn methods learn over a progression of data by processing each sample ahead arrival. At each encircling, the learner first receives one instance, makes a calculation, and receives the true label.

The error information is then used to update the learning model. Propose a novel collaborative online multitask learning (COML) technique to attack the aforementioned challenges. The basic idea is to first build a generic global model from large amount of data gathered from all user, and then consequently leverage the global model to build the personalized classification models for individual users through a collaborative learn process. We plan this idea into an optimization problem under an online learn setting, and propose two dissimilar COML algorithms by exploring different kinds of online learning methodologies.

SENTIMENT DETECTION USING SYNTHETIC DATASET:



IV. METHOD

Synthetic dataset:

It was designed by Sheldon to show that solving multiple related tasks jointly outperforms the solution that treats each task in isolation.

The goal is to discriminate two classes (positive class and negative class) in a two-dimensional plane with non – linear decision boundaries.

ALGORITHM

4.1 COML Algorithm: COML operates in a sequential manner.

```

Algorithm 1: Collaborative online multitask learning


---


Input: training data of  $K$  users  $(\mathbf{x}^k, y^k)$ ,
 $k = 1, \dots, K$ , parameters  $\eta_1 \geq 0, \eta_2 \geq 0$ ,
 $C \geq 0$ 
Initialize  $\mathbf{w}_0^k \leftarrow \mathbf{0}, \mathbf{u}_0 \leftarrow \mathbf{0}$ ;
for  $t \leftarrow 1$  to  $T$  do
    // Update the collaborative models
    for  $k \leftarrow 1$  to  $K$  do
        receive global model weight vector  $\mathbf{u}_t$ ;
        receive training instance  $(\mathbf{x}_t^k, y_t^k)$ ;
        set loss
         $\ell(\mathbf{w}^k; (\mathbf{x}_t^k, y_t^k)) = \max(0, 1 - y_t^k \mathbf{w}^k \cdot \mathbf{x}_t^k)$ ;
        if  $\ell(\mathbf{w}^k; (\mathbf{x}_t^k, y_t^k)) = 0$  then
             $\tau = 0$ 
        else
             $\tau = \min \left\{ C, \frac{\eta_1 + \eta_2 - y_t (\eta_1 \mathbf{w}_t + \eta_2 \mathbf{u}_t) \cdot \mathbf{x}_t^k}{\|\mathbf{x}_t^k\|^2} \right\}$ 
        end
        update  $\mathbf{w}_{t+1}^k = \frac{\eta_1 \mathbf{w}_t + \eta_2 \mathbf{u}_t + \tau y_t \mathbf{x}_t^k}{\eta_1 + \eta_2}$ ;
    end
    // Update the global model
    for  $k \leftarrow 1$  to  $K$  do
        receive training instance  $(\mathbf{x}_t^k, y_t^k)$ ;
        set loss
         $\ell(\mathbf{u}_t; (\mathbf{x}_t^k, y_t^k)) = \max(0, 1 - y_t^k \mathbf{u}_t \cdot \mathbf{x}_t^k)$ ;
        if  $\ell(\mathbf{u}_t; (\mathbf{x}_t^k, y_t^k)) > 0$  then
             $\sigma = \min \left\{ C, \frac{1 - y_t^k \mathbf{u}_t \cdot \mathbf{x}_t^k}{\|\mathbf{x}_t^k\|^2} \right\}$ ;
            update  $\mathbf{u}_{t+1} = \mathbf{u}_t + \sigma y_t^k \mathbf{x}_t^k$ ;
        end
    end
end

```

At each learning around, it collects the current global set of data, one from each of the engaged user/tasks, which are employed to update the global classification model. It builds a global classification model to exploit the commonality among tasks.

Passive Aggressive framework: This framework is used to build a global model using data collected from all users. Use the shared information

4.2 CONFIDENCE WEIGHTED LEARNING Algorithm:

Confidence weighted learnAlgorithm, planned by Crammer et al., models the linear classifier hypotheses with a multivariate Gaussian distribution over weight vector, which is then used to control the direction and scale of parameter updates. This algorithm performs well on quite a few tasks [1]

(4.2.1) Adaptive regularization of weights (AROW):

It solved the unconstrained objective function on each round [1].

(4.2.2) Kullback-Leibler (KL) divergence sense:

Adaptive regularization of weights ensures that the updated distribution is similar to the current distribution. The weight division is updated by minimizing the KL divergence between the new and old weight distributions, together with the KL divergence between the weight distributions of individual and global models [1].

(4.2.3) Gaussian distribution:

We maintain a global classification model parameterized by a Gaussian distribution over weight vectors with mean μ and covariance Σ , and use the global model information together with each user data to train the collaborative models [1].

(4.2.4) Comparison:

To examine the effectiveness of learning multiple related tasks together, we compare cw method with Passive Aggressive and Adaptive regularization of weights algorithms as below.

Global model:

At each learning round, the algorithm receives a training sample from each task, and uses the sample to update its weight vector.

Personal model:

Every task is associated with a personalized classification model.

Simple model:

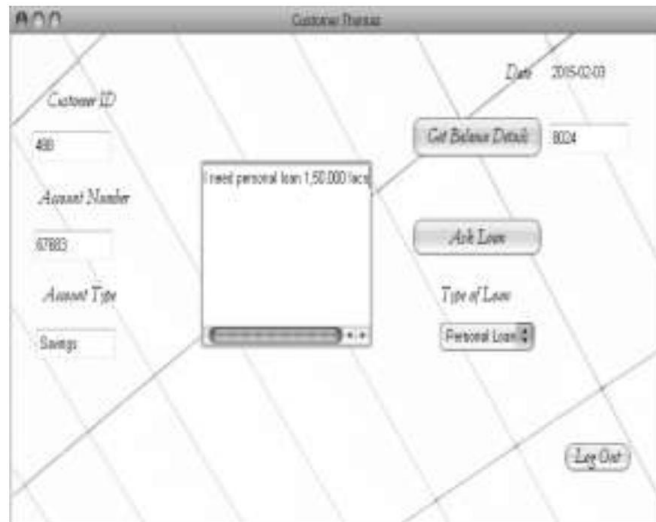
It simply switches between the global model and personal models according to their cumulative error counts in previous online learning rounds. At each round it sets its weight vector to that of the best model i.e.; one with the least cumulative errors to-date.



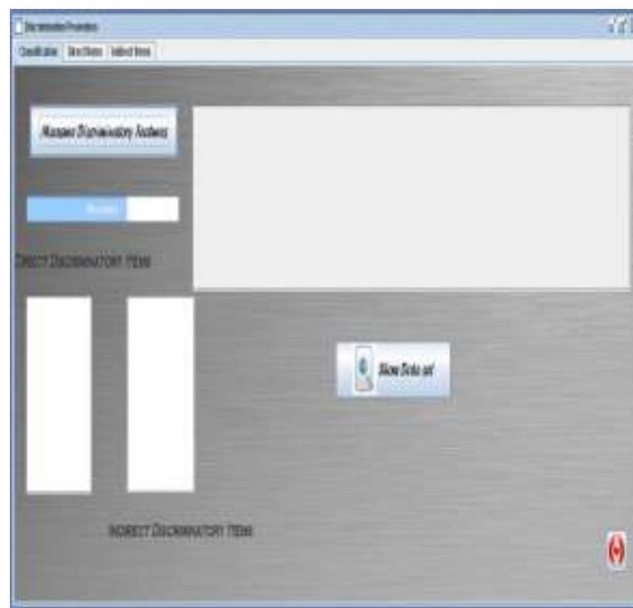
This figure shows about the data collection of the users like name, sex, age etc



The above figure shows the account login of the particular customer in the bank.



This figure shows about the details of the customer whether he needs the personal loan or business loan



The above figure shows about the direct discrimination and indirect discriminations of dataset.

V.CONCLUSION

We proposed a collaborative online multitask learning method that is able to take advantage of individual and global models to achieve an overall improvement in classification performance for jointly learning multiple correlated tasks. We showed that it is able to outperform both the global and personal models by coherently integrating them in a unified collaborative learning framework. The experimental results demonstrate that our algorithms are both effective and efficient for three real-life applications, including online spam email filtering, MHC-I binding prediction, and micro-blog sentiment detection task. Although the collaborative online multitask learning algorithm was firstly designed to solve the UGC classification problem, it has potential applications outside of the domains studied here. We hope to be able to extend our experiments to a more substantial size dataset and also to more applications. Our methods assume uniform relations across tasks. However, it is more reasonable to take into account the degree of relatedness among tasks. In conclusion, our collaborative online multitask learning method is a significant first step towards a more effective online multitask classification approach. We will be implementing the cluster of tasks by using the collaborative online multitasking learning techniques. This work can be extended to get even better by using effective cluster on multitask learning techniques.

REFERENCES

- [1].Guangxia, Li,Steven, C.H.Hoi,Kuiyu, Chang,Wenting Liu,and Ramesh Jain vol.26,no.8,August 2014
- [2].Dan Klein. Joseph Smarr. Huy Nguyen. Christopher D. Manning "Named Entity Recognition with Character LevelModels". In Proceedings of CONLL-2003.
- [3].Chang CC, Lin CJ "LIBSVM: A library for support vector machines"ACM ... Communications and Control (ICECC), 2011International Conference.

- [4].Nicola Ueffing,” Transductive learning for statistical machine translation”In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pp. 25-32.
- [5].G. Li, K. Chang, S. C. H. Hoi, W. Liu, and R. Jain, “Collaborative online learning of user generated content,” in Proc. 20th ACM Int. CIKM, 2011, pp. 285–290.
- [6].L.J.P. van der Maaten and G.E. Hinton. “VisualizingHigh-Dimensional Data Usingt-SNE”. Journal of Machine Learning Research 9(Nov) PP: 2579-2605, 2008.
- [7].C. Widmer, Y. Altun, N. C. Toussaint, and G. Rätsch, “Inferring latent task structure for multi-task learning via multiple kernel learning,” BMC Bioinformatics, vol. 11, Suppl. 8, p. S5, Oct. 2010.
- [8].G. Li, S. C. H. Hoi, K. Chang, and R. Jain, “Micro-blogging sentiment detection by collaborative online learning,” in IEEE 10thICDM, Sydney, NSW, Australia, 2010, pp. 893–898.
- [9].L.Yang, R. Jin, and J.Ye, “Online learning by ellipsoid method,” in Proc. 26th Annu. ICML, Montreal, QC, Canada, 2009, p. 145.
- [10]. P. Zhao, S. C. H. Hoi, and R. Jin, “DUOL: A double updating approach for online learning,” in NIPS, 2009, pp. 2259–2267.
- [11]. R. Collobert and J. Weston, “A unified architecture for natural language processing: Deep neural networks with multitask learning,” in Proc. 25th ICML, Helsinki, Finland, 2008, pp. 160–167.
- [12]. K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, “Online passive-aggressive algorithms,” J. Mach. Learn. Res., vol. 7, pp. 551–585, Mar. 2006.
- [13]. N. Cesa-Bianchi, A. Conconi, and C. Gentile, “A second-order perceptron algorithm,” SIAM J. Comput., vol. 34, no. 3, pp. 640–668, Jul. 2005.
- [14]. N. Cesa-Bianchi, A. Conconi, and C. Gentile, “On the generalization ability of on-line learning algorithms,” IEEE Trans. Inform. Theory, vol. 50, no. 9, pp. 2050–2057, Sept. 2004.
- [15]. K. Crammer and Y. Singer, “Ultraconservative online algorithms for multiclass problems,” J. Mach



K.M.Hemavathy, Graduated B.E and her area of interest is Data mining and Networking.



M.BEEMAMEHRAJ Asst. Professor in Bharath Unviersity and her area of interest is Data