International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Edification Training for Participating in Various Activities through Online

S. Thomas Fleming[1], P. Selvaraj[2]

[1]M. Tech-Information Technology, SRM University, Chennai, Tamilnadu, India

[2]Assistant Professor, Department of Information Technology, SRM University, Chennai, Tamilnadu, India

Abstract: ***The specification of online multitask learning for participating in various activities for recovering various classification process that is in parallel related , focusing at every part of data received by each accurately and efficiently. Statistical computational linguistics systems area unit sometimes trained on giant amounts of micro-blog sentiment detection on a bunch of users that classifies micro-blog posts generated by every user into emotional or non-emotional classes. 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 classification methods, be it batch 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 independently on individual tasks may suffer from insufficient training data. To rectify this problem in this paper, we propose Edification training for participating in various activities through online; from this we can geographical model over the entire data of all process. Another part individual model for various related process are 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 spam email filtering, bioinformatics data classification, and micro-blog sentiment detection.***

Keywords: data mining, artificial intelligence, multi task learning

1. 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 multitask 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 learning setting, which assumes that the training 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 multitasks learning algorithms usually have fairly intensive training cost and poor scalability performance, as far as large real applications is concerned.

In this paper, we investigate the problem of online multitask learning, which differs from the classical multitask learning in two aspects. First, our goal is to improve the learning performance of all tasks instead of focusing on a single primary task. Second, we frame the multitask learning problem in an online learning setting by assuming that the data for each task arrives sequentially, which is a more realistic scenario for real-world applications. Unlike batch learning techniques, online learning methods learn over a sequence of data by processing each sample upon arrival. At each round, the learner first receives one instance, makes a prediction, and receives the true label.

The error information is then used to update the learning model. Our early study on this work was first motivated by the need to classify online user-generated content (UGC), e.g., micro-blog posts or spam email tags. For UGC, each individual exhibits uniqueness, but also shares certain characteristics with others in the group. It is thus desirable to develop an efficient and scalable classifier that can solve individual task by adapting to the global knowledge shared by all users. Consider the real problem of micro blog sentiment analysis on a group of users where the goal is to classify micro-blog posts generated by each user into several emotional or non-emotional categories in a near real-time manner.

When finding this downside by classical machine learning techniques, we are going to face a quandary, i.e., one international classification model trained on the whole assortment of information from all users might fail to capture the peculiarity of individual users and so typically works poorly. On the opposite hand, a totally personalized model for every user is also inaccurate as a result of insufficient coaching knowledge, particularly at the first stage of the training task. This so motivates U.S.A.to check on-line multitask learning techniques. We propose a novel collaborative online multitask learning (COML) technique to attack the aforementioned challenges.

The basic plan is to 1st build a generic world model from great deal of information gathered from all users, then after leverage the world model to make the customized classification models for individual users through a cooperative learning method. We formulate this idea into an optimization problem under an online learning setting, and propose two different COML algorithms by exploring different kinds of online learning methodologies.

2. Existing System

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 learning setting, which assumes that the training 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 multitasks learning algorithms usually have fairly intensive training cost and poor scalability performance, as far as large real applications is concerned.

It is thus desirable to develop an efficient and scalable classifier that can solve individual task by adapting to the global knowledge shared by all users. Consider the real problem of micro blog sentiment analysis on a group of users where the goal is to classify micro-blog posts generated by each user into several emotional or non-emotional categories in a near real-time manner. When solving this problem by classical machine learning techniques, we will face a dilemma, i.e., a single global classification model trained on the entire collection of data from all users may fail to capture the peculiarity of individual users and thus often works poorly.

Disadvantages

1. The authors used feature hashing to solve multitask learning problem.
2. Unlike the existing batch multitask learning studies, our work is closer to the online multitask learning methodology.
3. Many real-world problems are essentially multitask learning, although they are often broken into smaller single learning tasks,

3. Proposed System

We investigate the problem of online multitask learning, which differs from the classical multitask learning in two aspects. First, our goal is to improve the learning performance of all tasks instead of focusing on a single primary task. Second, we frame the multitask learning problem in an online learning setting by assuming that the data for each task arrives sequentially, which is a more realistic scenario for real-world applications. Unlike batch learning techniques, online learning methods learn over a sequence of data by processing each sample upon arrival. At each round, the learner first receives one instance, makes a prediction, 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 users, and then subsequently leverage the global model to build the *personalized* classification models for individual users through a *collaborative* learning process. We formulate this idea into an optimization problem under an online learning setting, and propose two different COML algorithms by exploring different kinds of online learning methodologies

Advantage:

1. The confidence weighted learning algorithm maintains a probabilistic measure of confidence in each component of its weight vector using a Gaussian distribution.
2. It also uses an improved update strategy, leading to extra robustness in the case of non-separable data empirically.
3. We have a tendency to assess the performance of our rule on an artificial dataset and 3 real-life datasets.

4. Modules Details

1. Data Model Classification
2. Classical Learning Process
3. Multitask Learning Process
4. Global Model by Online Learning

1. Data Model Classification

It is thus desirable to develop an efficient and scalable classifier that can solve individual task by adapting to the global knowledge shared by all users. Consider the real problem of micro blog sentiment analysis on a group of users where the goal is to classify micro-blog posts generated

2. Classical Learning Procees

To evaluate the efficacy of the proposed technique, we conduct experiments by comparing our algorithms against a variety of state-of-the-art techniques on a synthetic dataset and three real-life applications, including online peptide binding prediction in bioinformatics, and micro-blog sentiment detection

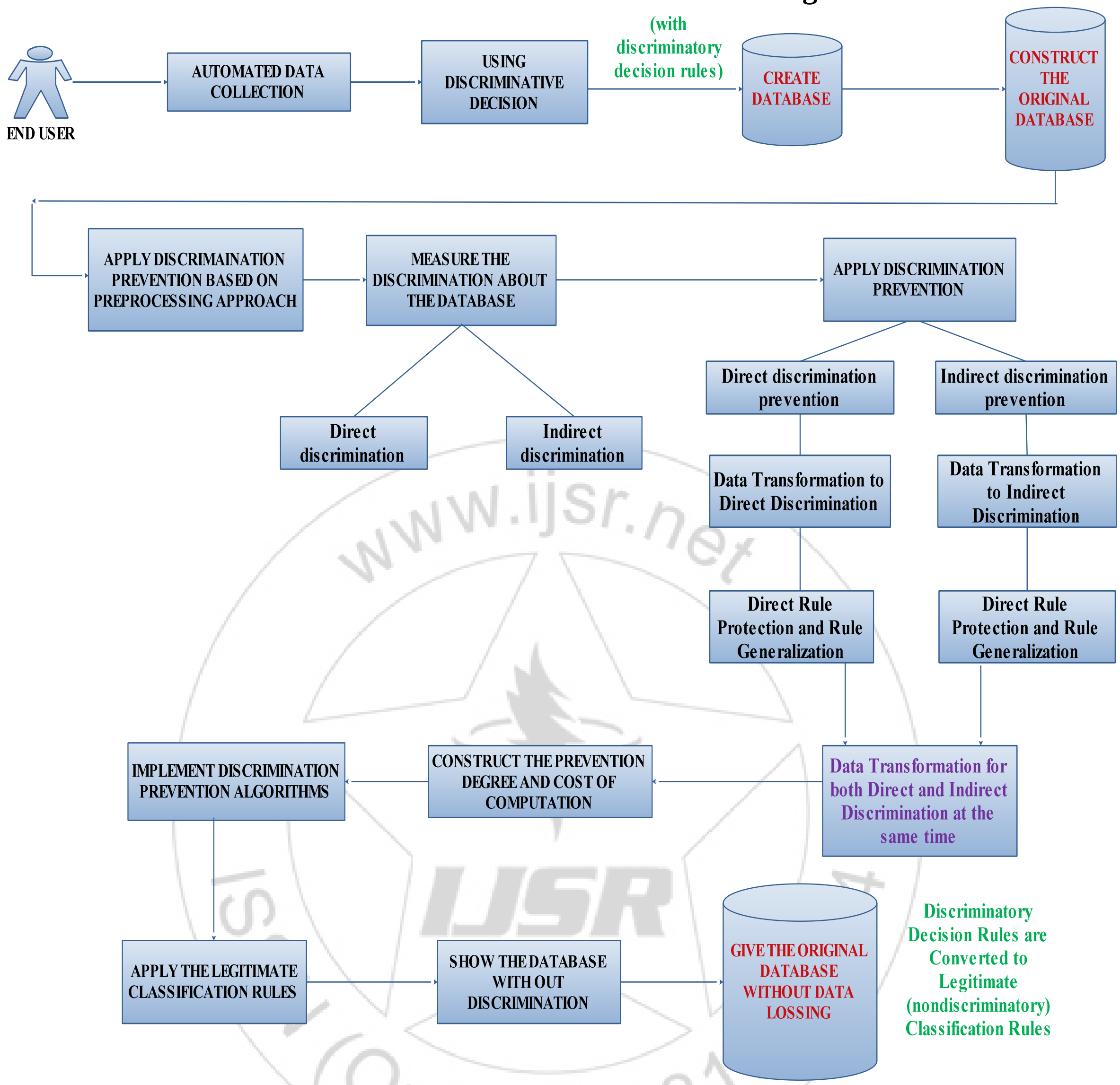
3. Multitask Learning Process

Specifically, the collaborative online multitask learning operates in a sequential manner. At each learning round, it collects the current global set of data; one from each of the engaged users/tasks, which are employed to update the global classification model. At the same time, a collaborative personalized model is maintained for each user/task. The individual collaborative classification model is subsequently updated using the latest individual data and the global model parameters. Therefore, our approach can leverage global knowledge for classification, while adapting to individual nuances via the collaborative learning way.

4. Global Model by Online Learning

The performance of the proposed collaborative online multitask learning methods is better than that of the two batch learning algorithms .It should be noted that compared to online learners who update models based only on the current sample, batch learning methods have the advantage of keeping a substantial amount of recent training samples, at the cost of storage space and higher complexity. In fact, the proposed COML algorithm is more efficient than batch incremental method.

5. Data Flow Diagram



6. Conclusion

In this paper, we have a tendency to planned a cooperative on-line multitask learning methodology that's ready to benefit of individual and international models to realize an overall improvement in classification performance for put together learning multiple correlate tasks. We have a tendency to show that its ready to exceed each the worldwide and private models by coherently group action them during a unified cooperative learning framework. The experimental results demonstrate that our algorithms are each effective and economical for 3 real-life applications, together with in-line 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. How to incorporate hierarchies and clusters of tasks is also worthy of further study. In conclusion, our collaborative online multitask learning method is a significant first step towards a more effective online multitask classification approach.

References

[1] R. Carina, "Multitask learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, Jul. 1997.
[2] T. Eugenio, C. A. Michelle, and M. Pintail, "Learning multiple tasks with kernel methods," *J. Mach. Learn. Res.*, vol. 6, pp. 615–637, Apr. 2005.
[3] R. Collier 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.
[4] C. Wider, Y. Alton, N. C. Toussaint, and G. Retch, "Inferring latent task structure for multi-task learning via

multiple kernel learning," *BMC Bioinformatics*, vol. 11, Suppl. 8, p. S5, Oct. 2010.

[5] G. Li, S. C. H. Hoi, K. Chang, and R. Jain, "Micro-blogging sentiment detection by collaborative online learning," in *IEEE 10th ICDM*, Sydney, NSW, Australia, 2010, pp. 893–898.

[6] 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.

[7] N. Cesar-Bianchi, A. Condone, 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.

[8] K. Crammer, O. Decal, J. Kismet, S. Shalev-Shwartz, and Y. Singer, "Online passive-aggressive algorithms," *J. Mach. Learn. Res.*, vol. 7, pp. 551–585, Mar. 2006.

[9] K. Crammer and Y. Singer, "Ultraconservative online algorithms for multiclass problems," *J. Mach. Learn. Res.*, vol. 3, pp. 951–991, Jan. 2003.