2015 IEEE 27th International Conference on Tools with Artificial Intelligence

Active Learning for Multivariate Time Series Classification with Positive Unlabeled Data

Guoliang He [1], Yong Duan [1], Yifei Li [1], Tieyun Qian [1], Jinrong He [2,*], Xiangyang Jia [1]

[1] State Key Laboratory of Software Engineering, College of Computer Science, Wuhan University, Wuhan, China
[2] College of Information Engineering, Northwest A & F University, Yangling, China
glhe@whu.edu.cn, {dydm_13128, yfli814}@163.com, qty@whu.edu.cn, hejinrong@nwsuaf.edu.cn, jiaxiangyang@163.com

***Abstract*—Traditional time series classification problem with supervised learning algorithm needs a large set of labeled training data. In reality, the number of labeled data is often smaller and there is huge number of unlabeled data. However, manually labeling these unlabeled examples is time-consuming and expensive, and sometimes it is even impossible. Although some semi-supervised and active learning methods were proposed to handle univariate time series data, few work have touched positive and unlabeled data for multivariate time series (MTS) classification due to the data being more complex. In this paper we focus on active learning for multivariate time series classification with positive unlabeled data. First, we propose a sample selection strategy to find the most informative unlabeled examples for manual labeling. Second, we introduce two active learning approaches to obtain a high-confident training dataset for classification. Experiments on real datasets demonstrate the validity of our proposed approaches.**

Keywords-multivariate time series; active learning; positive and unlabeled data

I. INTRODUCTION

Traditional time series classification methods mostly apply supervised learning techniques on the labeled training dataset. In reality, the number of labeled data is often smaller and there is a huge number of unlabeled data as the quick development of network and information technology. However, manually labeling examples is time-consuming and expensive, and sometimes it is even impossible. Therefore, how to make better use of these unlabeled data to improve the performance of classification is critical. More specially, labeled examples of some classes are missing. For example, for a binary classification, only few labeled positive examples and lots of unlabeled examples in the original dataset.

To handle this issue, some semi-supervised methods have been proposed. For instance, Wei *et al.* [1] proposed a semi-supervised way to automatically identify unlabeled examples. To find an accurate boundary between positive and negative samples in unlabeled samples, Nguyen *et al.* [2] proposed an effective technique called LCLC (Learning from Common Local Clusters). However, the assumption that all samples within a cluster are from same class is unpractical.

However, semi-supervised methods are difficult to find confident examples with the limited labeled multivariate time series (MTS) data because comparing with univariate time series, MTS data has multiple variables and data distribution is complex. Moreover, it could not be efficient to deal with multiple sub-concepts in a class when the size of labeled examples is extremely smaller. On the other hand, since active learning methods cannot be directly applied without initially making an assumption on label assignment when the original data is extremely limited, few active learning works have touched MTS data. Therefore, it is a challenging work to obtain sufficient and reliable labeled MTS data from positive and unlabeled data for classification.

In this paper, we focus on active learning for multivariate time series classification with positive unlabeled data. Due to human resources being very expensive, we try to manually label the least number of examples and obtain a confident and informative dataset, which is sufficient to learn an efficient classification model. We make several contributions. First, we propose a sample selection strategy to find the most informative unlabeled examples for manual labeling. Second, we introduce two active learning approaches to obtain a high-confident training dataset for classification. Our experimental results on benchmark datasets show that when the percentage of labeled positive examples is low, the performances of our proposed methods are more competitive than existing approaches.

The remainder of this paper is organized as follows. In Section II, we review the problem definition and related work. Section III introduces sample selection strategy and active learning framework. In Section IV, we perform a comprehensive set of experiments on several datasets. Finally, we conclude our work and suggest directions for future work in Section V.

II. RELATED WORK

As mentioned, traditional classification methods need lots of labeled training examples for learning. In practice, labeled data are often insufficient and labeling examples is time-consuming and expensive. To make better use of unlabeled examples, semi-supervised approaches have been proposed to improve the performance of time series

* Corresponding author

1082-3409/15 $31.00 © 2015 IEEE
DOI 10.1109/ICTAI.2015.38

IEEE computer society

classification. For instance, Wei and Keogh [1] proposed a semi-supervised way to automatically label the nearest example closest to labeled examples iteratively. However, it's hard to find a good stopping criterion to decide when to stop the iteration, which leads to automatically labeled examples being not confident. To find an accurate boundary between positive and negative samples in unlabeled samples, Nguyen *et al.* [2] proposed an effective technique called LCLC (Learning from Common Local Clusters). The unlabeled samples are clustered and all samples within a cluster are assumed to be from same class. Then, a chaining approach was applied to find the boundary between positive and negative clusters in terms of their similarity distances. However, the assumption that all samples within a cluster are from same class is unpractical. To overcome the drawbacks of LCLC, Nguyen *et al.* [3] proposed a novel ensemble based approach to obtain multiple diverse classifiers.

However, these semi-supervised methods are unable to accurately extract the reliable positive and negative examples from unlabeled data, resulting in inaccurate classifiers. To enhance its performance, active learning is an efficient way to find the most valuable instances from lots of unlabeled data for manual labeling. It makes use of user's feedback to enlarge the size of the training data and enhance the confidence of examples, which could improve effectiveness of learning algorithm. So far many sample selection approaches have been proposed in literature [4-6]. Reference [7] was a comprehensive survey of recent works in this field.

Among all sample selection methods, uncertainty sampling is a popular way. It tries to find the least uncertain example, which belongs to several classes with the same probability. For instance, the margin sampling strategy selects the unlabeled examples which lies within the margin of the current SVM since these examples are the most likely to become new support vectors [8,9]. To solve the outlier problem by uncertainty sampling, Zhu *et al.* [10] presented two density based techniques, which prefer not only the most informative example in terms of uncertainty criterion, but also the most representative example in terms of density criterion. Sheng-Jun Huang *et al.* [11] proposed a systematic way for measuring and combining the informativeness and representativeness of unlabeled instance by its prediction uncertainty for active learning. Moreover, Lindenbaum *et al.* [12] proposed a sample selection method by finding an example with the highest utility, and taking its effect on the resulting classifier into account. To estimate the probability of its possible labels in the process of computing the expected utility of an example, they adopted the random field model.

However, existing active learning methods cannot be directly applied without initially making an assumption on label assignment when the original data is extremely limited. To handle this issue, Sun *et al.* [13] proposed using canonical correlation analysis to investigate the correlation between different views of the available data. Experiments demonstrated that this measure can be used as a selection criterion for active learning with only a single labeled example in each class. To deal with few labeled positive examples and a huge number of unlabeled examples, Garg *et al.* [14] introduced an active learning based method using SVM classifier. Alireza Ghasemi *et al.* [15] proposed an uncertainty-based active learning algorithm to select most uncertain samples. They firstly estimated probability density of positive and unlabeled points separately, and then computed expected value of informativeness to get rid of a hyper-parameter. However, it is not suitable to be used in applications because the prior probability of positive class is usually not available in reality.

Due to no sample selection method being always better than other techniques on all kinds of data sets, Ferdowsi *et al.* [16] proposed an online algorithm for active learning that switches between different candidate instance selection strategies for classification in imbalanced data sets. To combine with semi-supervised method, Shusen Zhou *et al.* [17] proposed an active deep network (ADN) algorithm to address sentiment classification on insufficient labeled data, and incorporated pool-based active learning with the ADN method to achieve higher classification performance.

III. THE PROPOSED TECHNIQUE

In this section, we present two active learning techniques with the smallest cost to obtain most reliable and informative positive and negative examples from positive unlabeled data. First, we discuss how to select valuable samples for manual labeling. Then, two active learning techniques are proposed to obtain sufficient labeled examples as training data.

A. Sample selection straegy

As mentioned above, to select important examples for manual labeling, many techniques have been proposed, such as marginal sampling, density-weighted method, etc. However, these methods did not consider its affection for the learning model in the next iteration.

To select the most important sample for manual labeling, here we consider its uncertainty and utility simultaneously. The more uncertain an unlabeled sample, the more difficult the current learning model to identify its class. The utility of an unlabeled sample denotes it could supply how much information to update the learning model, that is to say, a more accurate classifier could be learned if we identify an unlabeled sample with a high utility in the process of training a learning model.

Now we define the uncertainty and utility of an unlabeled sample.

Definition 1. Uncertainty. For an unlabeled sample u, suppose its nearest positive sample is P and nearest negative sample is N. we can calculate the uncertainty of this unlabeled sample u by

$$\mathrm{UCT}(u) = \frac{\min\{Sim(u,P), Sim(u,N)\}}{\max\{Sim(u,P), Sim(u,N)\}} \quad (1)$$

where Sim(u, P) and Sim(*u*,N) is the similarity distance *u* and its nearest positive sample P and its nearest negative sample N, respectively. For the similarity distance, there are several methods [18-21]. Here we adopt Eros distance method to measure the similarity between two MTS examples.

From formula (1) we can see that the uncertainty of an unlabeled sample u is higher when Sim(u,P) and Sim(u,N) are close to each other. According to the definition of uncertainty, we plan to select the a few unlabeled samples with the highest uncertainty for active learning. On the other hand, the uncertainty of an unlabeled sample denotes the correct degree of the class of this sample, which does not represent its effectiveness for the next learning iteration if it has been selected for active learning.

For instance, Figure 1 plots some labeled and unlabeled samples, where solid and circle symbols present labeled and unlabeled samples, respectively, and a dot or diamond sample denote its class is positive or negative in nature. In this figure A and B have the largest uncertainty. However, there are no unlabeled samples around A while a few unlabeled samples around B. Comparing A and B, labeling B is more effective because it could lead to its neighbors being easy to be labeled while A being similar to a outlier.

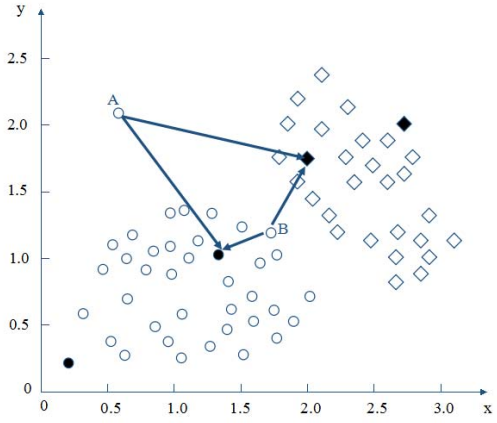


Figure 1. The utility of two unlabeled examples A and B

To deal with this issue, we assess the utility of each chosen unlabeled sample. Considering Figure 2, p_1, p_2, n_1 and n_2 are labeled samples, and A, B and C are unlabeled samples the highest uncertainty. To label optimal unlabeled samples for next learning, we set p_1, p_2, n_1, n_2, A, B and C as centers, and each unlabeled sample u is assigned to one of the these centers with the nearest neighbor method. By counting the number of neighbors of A, B and C, the effectiveness of an unlabeled sample with the highest uncertainty is determined. Specially, for an unlabeled sample A with the highest uncertainty, its utility is

$$\text{Utility}(A) = \frac{N_i}{|U| - K} \quad (2)$$

Where N_i is the number of neighbors of A, |U| is the number of the unlabeled samples and K is the number of unlabeled samples the highest uncertainty.

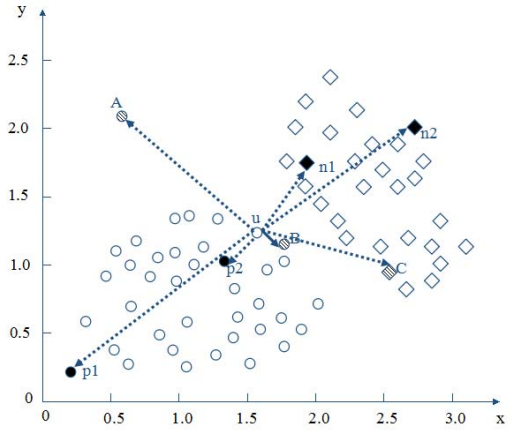


Figure 2. The profile to evaluate the utility of an unlabeled example B

B. Active learning framework

Based on the sample selection strategy presented in the above, the next work is to manually label the minimal number of unlabeled samples for training an efficient learning model. Now, we discuss two active learning methods to obtain a confident training dataset.

Brute-Force Active Learning (BFA)

The most straightforward way for active learning unlabeled examples is using the brute force method, and the framework is described in Table 1.

TABLE 1. BRUTE FORCE ALGORITHM FOR OBTAINING A SUFFICIENT LABELED DATASET

Input: a training dataset with a few positive samples P and huge number unlabeled samples U
Output: the labeled dataset D
1: initialize D = P
2: labeling a confident negative sample u in unlabeled data U
4: D = P + {u}; U= U- {u}
3: Do
4: Selecting an optimal unlabeled sample x within U
5: Labeling x manually
6: D = P + {x}; U= U- {x}
7: While (stopping condition is not satisfied)
8: Return D

Given a positive unlabeled dataset, to find maximal uncertain samples within the margin area between positive and negative data, we first try to find a confident negative example from unlabeled data (line 2). Since the most dissimilar example far from original positive examples is uncertain to be a negative due to only a few number of positive examples and the diversity of class. Here we manually label the dissimilar examples until it is a negative example. Then, we select an optimal unlabeled sample using

sample selection strategy (line 4) and label it until the stopping condition is satisfied. As for the stopping criterion, we will discuss it later.

Semi-supervised Active Learning (SSA):

For the brute force method, a confident example can be manually labeled in each iteration. As it is going on, newly labeled samples can offer more information about the data distribution, and it is convenient to use semi-supervised methods to expand the size of labeled training data. There is a popular belief that the larger the size of the training data, the more efficiency of the learning model, especially for K-NN method.

To enlarge the scale of labeled data, we discuss how to combine active learning and semi-supervised learning methods to obtain large number of labeled training data. While it is a simple idea, for clarity we provide the framework of semi-supervised active learning in Table 2.

TABLE 2. SEMI-SUPERVISED ACTIVE LEARNING FOR OBTAINING A SUFFICIENT LABELED DATASET

Input: a training dataset with a few positive samples P and lots of unlabeled samples U
Output: the labeled dataset D
1: initialize D = P
2: labeling a confident negative sample u in unlabeled data U
4: D = P + {u}; U= U- {u}
3: Do
4 Learning a classification model with labeled data D
5: Selecting an optimal unlabeled samples x within U
6: Labeling x manually
7: Automatically labeling some confident unlabeled samples Y
8: D = P + {x} + Y; U= U- {x} - Y
9: While (stopping criterion not satisfied)
10: Return D

This procedure is similar to the brute force framework except line 7, which is to automatically label some confident unlabeled samples Y with a semi-supervised method. In the process of the semi-supervised learning, in terms of the uncertainty of unlabeled examples, which has been evaluated in the process of sample selection, an unlabeled example that is the most close to labeled examples can be labeled automatically.

Stopping Criterion:

In the above a sample selection strategy is proposed to label an important and unlabeled example by considering their uncertainty and utility at each iteration. If the utility of an unlabeled sample is small, it means that this sample can't offer useful information to improve the classifier model. According to this idea, we stop the process of the active learning when the utility of the chosen unlabeled sample is smaller than a user-specified threshold. It is obvious that the smaller the threshold, the larger the number of unlabeled examples needed to be labeled manually.

After obtaining the labeled training data, many algorithms can be used for time series classification. The simple technique 1-NN (classification based on the nearest neighbor) with Euclidean distance was shown to be the most effective method. Therefore, in this paper we use 1-NN method for classification after the labeled training data are obtained using active learning method.

IV. EXPERIMETNAL EVALUATION

In this section, we empirically study the proposed method for MTS classification on positive unlabeled data. The algorithms are implemented in matlab, using a PC computer with Intel Core 3.6GHz CPU and 8 GB of main memory.

A. Datasets

The experiments are carried out on several real-world datasets, that is, EEG, Wafer, and ArabicDigits [22,23]. For the EEG dataset, it contains 1200 samples, each class has 600 samples. The length of an example is 256. And 600 examples are used for training while other 600 examples for testing.

There are 1194 samples in Wafer dataset, among which 1067 are normal and 127 samples are abnormal. The length of an example is between 104 and 198. And 634 examples are used for training while other 560 samples for testing.

There are ten classes for the ArabicDigits dataset, here we just use four of it, where two classes '0' and '1' as dataset AD1 and two class '4' and '5' as dataset AD2. Each class consists of 660 examples in the training data, and 220 examples in the testing data. The length of an example is between 4 and 93.

Table 3 shows the summary of all of the datasets used in the experiments.

Similarly, we also use the F-measure to evaluate the performance of these approaches. The F-measure is defined as F=2*p*r/(p+r), where p and r means the precision and recall of classification, respectively. F-measure is larger when both of precision and recall are good.

TABLE 3. DATASETS USED IN THE EVALUATION EXPERIMENTS

	EEG	Wafer	AD1	AD2
Num of variables	64	6	13	13
Max length	256	198	93	93
Min Length	256	104	4	4
Num of Samples	1200	1194	1760	1760

B. Analysis of sample selection strategy

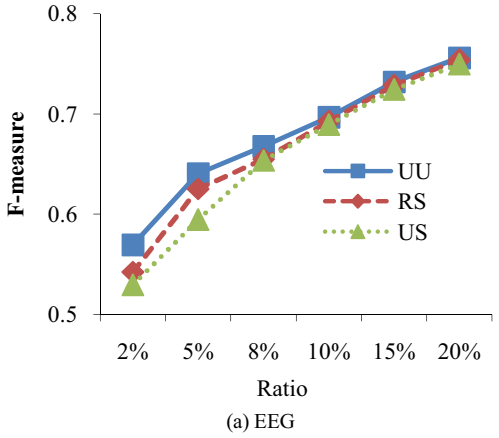
To illustrate the efficiency of our proposed sample selection method (UU) in terms of uncertainty and utility, in this section we first compare with two classical sample selection approaches as following.

Uncertainty sampling (US): selects examples with high uncertainty from the unlabeled data according to the labeled examples.

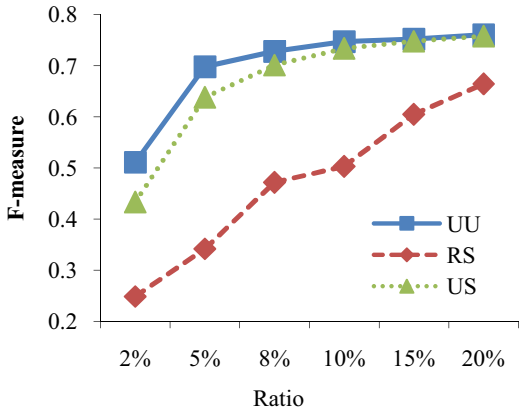
Random sampling (RS): selects examples from the unlabeled data randomly.

To analyze the function of these three sample selection methods, we compare their F-measure when they manually label the same number of unlabeled examples. To show the changes in the performance of classification as the manually labeled examples gradually augmented, we do experiments on different percentage of manually labeled examples.

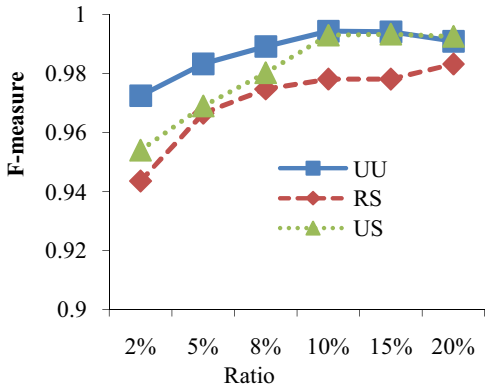
Figure 3 shows classification results of the three sample selection strategies. It is clear that UU produces the best results consistently across all labeled examples on each dataset, especially for the size of the manually labeled examples is smaller. When the number of manually labeled examples is larger, the difference of performances of UU, US and RS method is reduced.



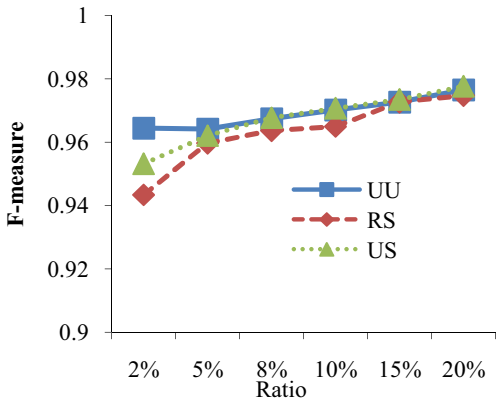
(a) EEG



(b) Wafer



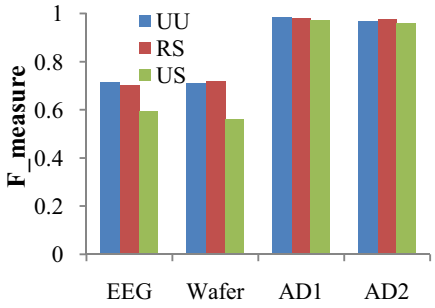
(c) AD1



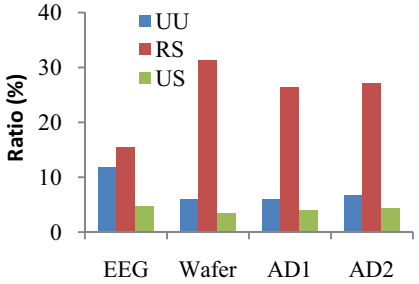
(d) AD2

Figure 3. The classification performances of three sample selection strategies with different ratio of manually labeled examples to original unlabeled examples

On the other hand, we also analyze the performance of these three methods using the active learning process when the stopping Criterion is satisfied, which is natural for applications. The comparison of the classification performances are shown in Figure 4. The performance of UU and RS methods are similar while much more examples need to be labeled by experts, which is more time-consuming and expensive. For US method, its performance is the worst although fewer examples need to be manually labeled. Considering the performance and the cost, UU method is much effective among these three selection methods.



(a) F-measure of the classification

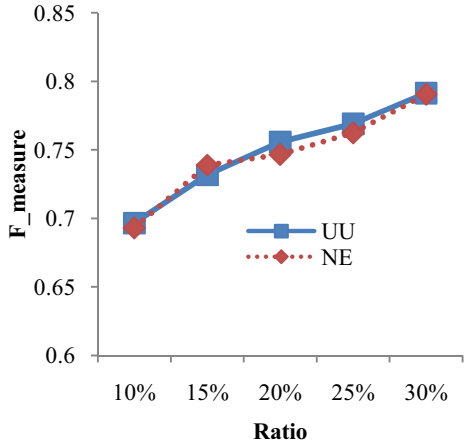


(b) Ratio of manually labeled examples to original unlabeled examples

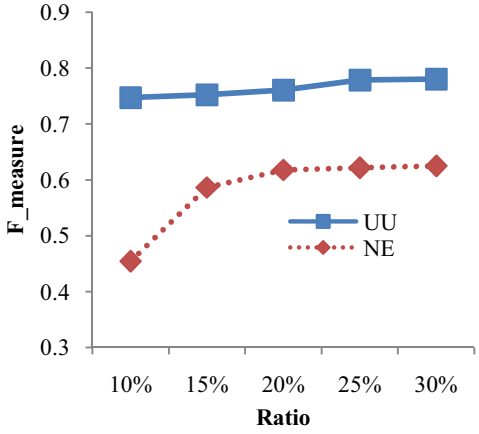
Figure 4. The results of three sample selection approaches with the same stopping condition

Moreover, we also compare UU method with a novel and effective sample selection approach, neighbor entropy active Learning method (NE) [16]. To be fair, similar to NE, we first select examples from the unlabeled data randomly and label them manually until the ratio of labeled examples to unlabeled ones is no less than 10%. Then, we adopt NE and UU to select samples for manually labeling. To analyze the function of two sample selection methods, we compare their F-measure when they manually label the same number of unlabeled examples. As mentioned before, to validate the effectiveness of our proposed sample selection method UU, we do experiments on different ratio of manually labeled examples to unlabeled examples.

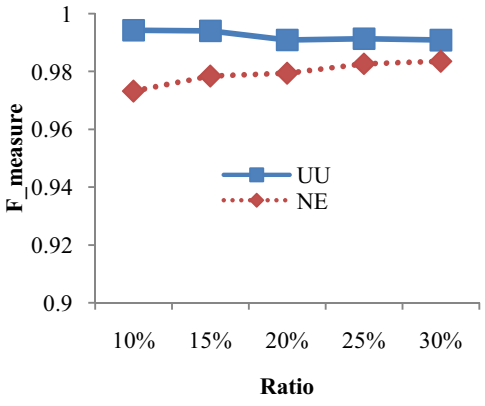
The classification results with UU and NE approaches are shown in Figure 5. From it we could see that our proposed method UU is always better than NE method across the entire range of different number of manually labeled examples. Note that, when the size of the manually labeled examples is larger, the difference between two methods is smaller. It is reasonable that when more labeled examples are added to the training data, their difference is doomed to decrease in the performance. Therefore, our proposed method could find more important example, which could improve the performance of classification.



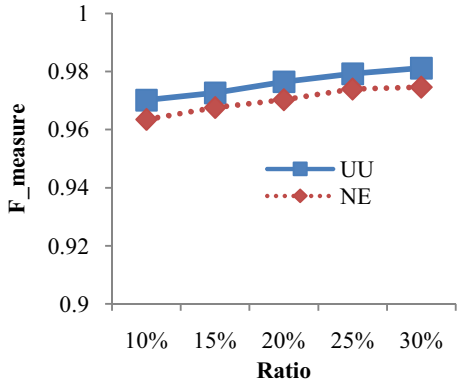
(a) EEG



(b) Wafer



(c) AD1



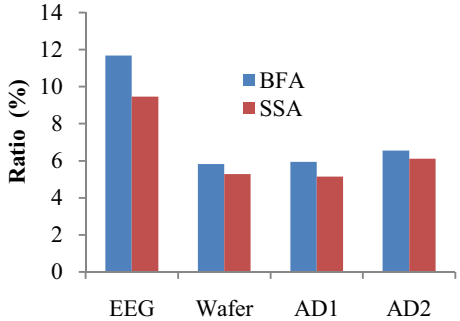
(d) AD2

Figure 5. The classification results of our proposed sample selection method (UU) and neighbor entropy based sample selection method (NE) with different ratio of manually labeled examples to original unlabeled examples

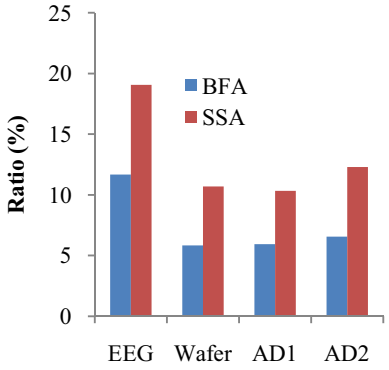
C. *Analysis of two active learning methods*

To get sufficient labeled examples, in section III two active learning methods are proposed. In general, we usually think that larger number of training data could improve the classification performance of the learning model. In this section we do some experiments to compare the performance of two proposed active learning methods. The experimental results of both methods are shown in Figure 6 and 7. It is obvious that in BFA much number of examples need to be labeled manually while much number of examples could be labeled in the process of SSA. It is reasonable because many examples could be labeled automatically using semi-supervised method in SSA. Meanwhile, F-measure of SSA is on EEG, Wafer AD1 and AD2, respectively, while that of BFA is on four datasets, respectively. We note that the classification performances of both active learning methods are similar.

Therefore, we can get two conclusions. Firstly, larger number of training data could not always improve the classification performance of the learning model. Smaller number of training data may learn a better learning model for classification when it consists of more informative examples. Secondly, SSA could reduce the amount of manual annotation, which is critical for many applications as the big data having coming.



(a) The ratio of manually labeled examples to original unlabeled examples



(b) The ratio of manually and automatically labeled examples to original unlabeled examples

Figure 6. The ratio of labeled examples to original unlabeled examples in the process of two active learning approaches on datasets

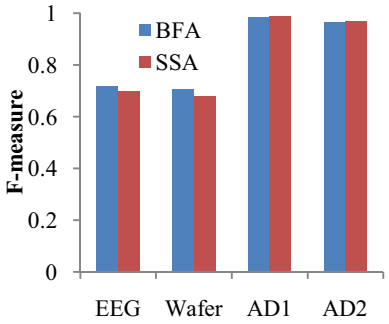


Figure 7. The Comparison of the classification performance with two active learning approaches on datasets

D. *Comparison of semi-supervised methods*

The last experiment analyzes the function for our proposed active learning methods against existing state-of-the–art semi-supervised methods for time series classification. That is Semi-06 [1], Semi-08 [24] and LCLC [2]. The experimental results on four datasets are listed in Table 4.

It is clear that our proposed methods are more efficient than existing semi-supervised methods. The reason is that BFA and SSA focus on selecting most informative examples while semi-supervised methods label examples which is closest to labeled ones from the unlabeled data. Therefore, after labeling, BFA and SSA could obtain more distinctive labeled example for classification.

TABLE 4. OVERALL PERFORMANCE OF VARIOUS TECHNIQUES

Data Set	BFA	SSA	Semi-06	Semi-08	LCLC
EEG	0.72	0.70	0.56	0.57	0.45
Wafer	0.71	0.67	0.06	0.08	0.13
AD1	0.98	0.99	0.45	0.76	0.37
AD2	0.97	0.97	0.80	0.81	0.78

V. CONCLUSIONS AND FUTURE WORK

In this paper we focus on addressing multivariate time series classification on few labeled positive examples and a huge number of unlabeled examples. First, we introduce a valid sample selection strategy for manual labeling. Then, two active learning frameworks are proposed to obtain sufficient and reliable labeled training data for classification. The experimental results show clearly that our proposed methods are competitive.

We note that the balance of the training data among different classes is an important element in the classification, which affect the performance of the learning model. In future, we plan to perform research on sample selection method to obtain a balanced labeled training data for multivariate time series classification.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China (61572376), the National Key Scientific Instrument and Equipment Development Project (2011YQ170065), and the Program of Introducing Talents of Discipline to Universities (B07037).

REFERENCES

[1] Li Wei, Eamonn Keogh. Semi-Supervised Time Series Classification. KDD, 2006.

[2] Minh Nhut Nguyen, Xiao-Li Li, See-Kiong Ng. Positive Unlabeled Learning for Time Series Classification. IJCAI, 2011.

[3] Minh Nhut Nguyen, Xiao-Li Li, See-Kiong Ng. Ensemble Based Positive Unlabeled Learning for Time Series Classification. DASFAA, 2012.

[4] Miloš B Stojanović, Miloš M Božić, Milena M Stanković, Zoran P Stajić. A methodology for training set instance selection using mutual information in time series prediction, Neurocomputing, 2014, 141:236-245

[5] Wang Zhenyu, Wang Xizhao. Active learning algorithm based on neighborhood entropy, Pattern recognition and artificial intelligence, 2011, 24(1): 97-102.

[6] Mohammad Sadegh Hajmohammadi, Roliana Ibrahim, Ali Selamat, Hamido Fujita. Combination of active learning and self-training for cross-lingual sentiment classification with density analysis of unlabelled samples, Information sciences, 2015,317:67-77.

[7] Burr Settles. Active Learning Literature Survey. Computer Sciences Technical Report 1648, 2010.

[8] Husheng Guo, Wenjian Wang. An active learning-based SVM multi-class classification model, Pattern recognition, 2015, 48(5):1577-1597.

[9] Simon Tong, Daphne Koller. Support vector machine active learning with applications to text classification, Journal of machine learning research, 2001, 45-66.

[10] Jingbo Zhu, Huizhen Wang, Benjamin K Tsou, Matthew Ma. Active learning with sampling by uncertainty and desnity for data annotations, IEEE Transactions on audio, speech, and language processing, 2010, 18(6):1323-1331.

[11] Sheng-Jun Huang, Rong Jin. Active Learning by Querying Informative and Representative Examples. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(10):1936-1949.

[12] Michael Lindenbaum, Shaul Markovitch, Dmitry Rusakov. Selective sampling for nearest neighbor classifiers, Machine Learning, 2004, 54: 125-152.

[13] Shiliang Sun, David R hardoon. Active learning with extremely sparse labeled examples, Neurocomputing, 2010, 73: 2980-2988.

[14] Priyanka Garg, S Sundararajan. Active learning in partially supervised classification, CIKM 2009.

[15] Alireza Ghasemi, Hamid R. Rabiee, Mohsen Fadaee, Mohammad T. Manzuri and Mohammad H. Rohban. Active Learning from Positive and Unlabeled Data. Data Mining Workshops, 2011.

[16] Zahra Ferdowsi, Rayid Ghani, Raffaella Settimi. Online active learning with imbalanced classes, IEEE 13th international conference on data mining, 2013. 1043-1048.

[17] Shusen Zhou, Qingcai Chen, Xiaolong Wang. Active deep learning method for semi-supervised sentiment classification, Neurocomputing, 2013(120):536-546.

[18] Mohammad Shokoohi-Yekta, Jun Wang and Eamonn Keogh. On the Non-Trivial Generalization of Dynamic Time Warping to the Multi-Dimensional Case. SDM 2015

[19] Tomasz Górecki, Maciej Łuxzak. Multivariate time series classification with parametric derivative dynamic time warping, Expert Systems with Applications, 2015, 42:2305-2312.

[20] Kiyoung Yang, Cyrus Shahabi. A PCA-based similarity measure for multivariate time series. MMDB, 2004:65-74.

[21] L. Karamitopoulos, G. Evangelidis, D. A. Dervos. Multivariate Time Series Data Mining: PCA-based Measures for Similarity Search. ICDM, 2008.

[22] Available from: <http://www.cs.cmu.edu/~bobski/>.

[23] Available from: <http://archive.ics.uci.edu/ml/datasets.html>

[24] Ratanamahatana C A, Wanichsan D. Stopping Criterion Selection For Efficient Semi-Supervised Time Series Classification, Software Engineering Artificial Intelligence Networking & Parallel/distributed Computing, 2008:1-14.