Review Article | Impact Factor: 0.621 | ISSN: 2319-507X
Ruchida Sonar, IJPRET, 2014; Volume 2 (9): 319-329 | IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

INCREMENTAL LEARNING: A REVIEW

RUCHIDA SONAR, DR. P. R. DESHMUKH

Sipna, COET, Amravati

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

Abstract: Machine learning sometimes called as Data Mining means extracting useful information from data. Machine learning has to categorize data to learn new concept. And hence data classification techniques are used, from which SVM is one technique. Support Vector Machines (SVMs) have been successfully applied to solve a large number of classification and regression problems. However, SVMs suffer from the catastrophic forgetting phenomenon, which results in loss of previously learned information. Learn++ have recently been introduced as an online or incremental learning algorithm. As the areas of applications in data mining are growing substantially, it has become extremely necessary for incremental learning methods to move a step ahead. The strength of Learn++ lies in its ability to learn new data without forgetting previously acquired knowledge and with- out requiring access to any of the previously seen data, even when the new data introduce new classes. In this paper we discuss the areas, methods and types of incremental learning currently taking place and highlight its potentials in aspect of decision making. The paper essentially gives an overview of the current research.

Keywords: Data Mining, SVM, Catastrophic Forgetting, Online Or Incremental Learning.

Corresponding Author: MS. RUCHIDA SONAR

Access Online On:

www.ijpret.com

How to Cite This Article:

Ruchida Sonar, IJPRET, 2014; Volume 2 (9): 319-329

PAPER-QR CODE

INTRODUCTION

Data mining tasks, where specific knowledge has to be extracted from a huge amount of data, for inducing data several approaches have been studied in machine learning and formally in the field of learning theory. Machine learning offers one of the most cost effective and practical approaches to the design of pattern classifiers for a broad range of pattern recognition applications. Support Vector Machines (SVMs) have enjoyed a remarkable success as effective and practical tools for a broad range of classification and regression applications [1-2]. As with any type of classifier, the performance and accuracy of SVMs rely on the availability of a representative set of training dataset. The performance of the resulting classifier relies heavily on the availability of a representative set of training examples. In many practical applications, acquisition of a representative training data is expensive and time consuming. Such scenarios require a classifier to be trained and incrementally updated, where the classifier needs to learn the novel information provided by the new data without forgetting the knowledge previously acquired from the data seen earlier. Learning new information with- out forgetting previously acquired knowledge raises the stability. A completely stable classifier can retain knowledge, but cannot learn new information, whereas a completely plastic classifier can instantly learn new information, but cannot retain previous knowledge. For learning from new data involves discarding the existing classifier, merging the old and the new data and training a new classifier from scratch using the aggregate data. However, the result is in catastrophic forgetting [3], it can be defined as the system is not able to learn new patterns without forgetting previously learned ones. This approach is unfeasible if previous data are no longer available.

Incremental learning is the solution to such scenarios, which can be defined as the process of extracting new information without losing prior knowledge from an additional dataset that later becomes available. Incremental learning methods modify concept descriptions to accommodate new learning events (Winston, 1975; Michalski and Larson, 1978).

When we observe human learning we clearly see that it is incremental. People learn concept descriptions from facts and incrementally refine those descriptions when new fats or observations become available. Newly gained information is used to refine knowledge structure and models, and rarely causes a reformulation of all the knowledge a person has about the subject at hand.

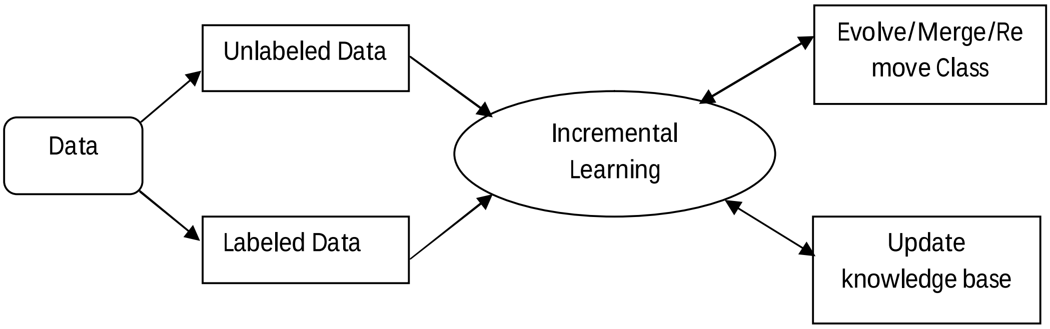

Figure 1 shows representative diagram of incremental learning.

There are two major reasons why humans must learn incrementally:

1. Sequential flow of information. A human typically receives information in step and must learn to deal with a situation long before all the information about it is available. When new information does become available, there is rarely time to reformulate everything known about all the concepts involved.
2. Limited memory and processing power. People cannot store and have easy access to all the information they have been exposed to. They modify the generalizations when new facts are available.

2. EXISTING LEARNING METHODS AND INCREMENTAL LEARNING

First and foremost we will discuss about the methods of unsupervised learning to supervised techniques, with the need for incremental learning. The paper proceeds in stages putting light on how incremental learning evolved with these learning methods. The focus of our paper is on the methods, the approaches and their novelty that are used for incremental learning referencing the type of application as well.

In the course of understanding what precisely an incremental approach is, it is necessary to understand that incremental learning can be in terms of the newly gained knowledge as well as evolving new class or a cluster. It can even merge or reform the classes. Precisely we can frame the learning to be one that is –

1. Capable to learn and update with every new data labeled or unlabeled.

2. It will use and exploit the knowledge in further learning.

3. Will generate a new class/cluster as required and take decisions to merge or divide them as well.

4. Will enable the classifier itself to evolve and be dynamic in nature with the changing environment.

With these factors, it can be rightly said that the incremental learning forms a full package that is up all the time. Considering the traditional methods and where incremental learning stands with respect to each of them along with the applications, is discussed ahead in this section.

2.1 Mathematical Representation and algorithm

When we talk about incremental learning, it is all about the learning approach or the classifier who is capable to perform the activities with respect to the environment. Mathematically it is represented as-

Let U = {ud1 , ud2 , ud3 ,...udn } be the new unlabeled data and

L = {ldi : C j | i = 1 to n , j = 1 to m} .

Let Ic be the classifier that is used for incremental learning. So, we can have-

K = f (Ic(U x), K prev) where K = {Cx , KB}

Here the value of Cx can be of existing class or new generated one. K governs the entire process. This is modeled and learned at every stage of new data availability. The learning process is summarized in the algorithm as follows:

1. For every Dx| Dx U or L

2. Do

Use KBprev

If (Dx U)

Classify Dx, with f (IC) GenerateK

Update KBnew←K+KBprev

Assign KBprev ←KBnew

3. TYPES OF INCREMENTAL LEARNING

Incremental learning by developing the two concepts of

1. Informational incremental learning and

2. Operationally incremental learning.

These concepts are applied to the problem of learning containment decision lists for demonstrating its relevance. Informational incremental algorithms are required to work incrementally as usual, i.e. they have no permission to look back at the whole history of information presented during the learning process. Operationally incremental learning algorithms may have permission to look back, but they are not allowed to use information of the past in some effective way. Obviously, the latter concept depends on some more detailed specification of how to process information presented.

Both types of learning [4] have a restricted power compared to unconstrained learning. The algorithmic learning theory is providing firm results in this regard. Within certain formal settings, operationally incremental learning is particularly restrictive. Especially, operationally incremental learning seems to be the crux of the limitations of strong case-based learning approaches. This may be understood as a hint how to relax requirements for enlargement of the scope of ease-based learning devices.

4. ENSEMBLE OF CLASSIFIERS

In this paper we describe an ensemble of classifiers approach: ensemble systems have attracted a great deal of attention over the last decade due to their empirical success over single classifier systems on a variety of applications. An ensemble of classifiers system is a set of classifiers whose individual decisions are combined in some way to obtain a Meta classifier. One of the most active areas of research in supervised learning has been to study methods for constructing good ensembles of classifiers. The main discovery is that ensembles are often more accurate than the individual classifiers that make them up. A rich collection of algorithms have been developed using multiple classifiers, such as AdaBoost [7] and its many variations, with the general goal of improving the generalization performance of the classification system. Using multiple classifiers for incremental learning, however, has been largely unexplored. Learn++, in part inspired by AdaBoost, was developed in response to recognizing the potential feasibility of ensemble of classifiers in solving the incremental learning problem. Learn++ was initially introduced in [5] as an incremental learning algorithm for MLP type networks. A more versatile form of the algorithm was presented in [6] for all supervised classifiers.

5. LEARN++

The Learn++ algorithm, exploits the synergistic power of an ensemble of classifiers to incrementally learn new information that may later become available. Learn++ generates multiple weak classifiers, each trained with different subsets of the data. For each database k, k=1,…,K that becomes available, the inputs to Learn++are (i)Sk □ □□ xi , yi □ | i □ 1, , mk □, a sequence of mk training data instances xi, along with their correct labels yi, (ii) a weak classification algorithm BaseClassifier to generate weak classifiers, and (iii) an integer Tk specifying the number of classifiers (hypothesis) to be generated for that database. We require that BaseClassifier obtain at least50% correct classification performance on its own training dataset, to ensure a meaningful classification performance for each classifier.

6. SUPERVISED AND SEMI-SUPERVISED INCREMENTAL LEARNING

In case of incremental learning in supervised or semi-supervised environment, what needs to be looked at is that the training data can arise at later stages. Rather that restricting the environment to the specific data, incremental learning unfolds the learning.

With pattern based techniques, [8] propose learning of new chunk of patterns keeping intact the previous ones on the basis of neural network, where the same can be applied in text domain [9].

In order to avoid the training phase, and update with each new training data that is evolved over the time, ensemble base methods are used [4]. [10] Suggest a Learn++, an approach inspired by Adaboost, working on neural network based ensemble classifiers working on digital optics database. Further [11] propose ADAIN, an adaptive framework which focus on use of nonlinear regressive models, but comparatively faster than Learn ++. [12] Adopt the Gaussian mixture model and Resource allocating NN for the learning with its application in a dormitory to study the habits of the students.

Performance driven data selection model is another approach [13] where selective incremental learning occurs for the unlabeled data, taking decision to learn from specific data sets that are classified to learn further.

A wide application to have a robot to learn manipulative tasks by use of Markov methods is [14]

Proposed, where initial stage teaching occurs and later the robot learns.

Extending the approach in medical image segmentation, [15] propose Ripple down rules for knowledge acquisition, where as Bayesian approach is proposed [16] to detect emergency in health surveillance.

Alike [17] propose its use in sports video view classification explicitly in baseball presenting a new distance measure along with a threshold criteria building positive and negative model pools. The discovery of interesting patterns further has given rise to [18] learning in detection of objects occurring in the images in hierarchy. Incremental detection and classifying the new images with existing objects is applied here. In case of face recognition too, incremental learning has taken a step ahead, [19] propose method for adaptive learning of new features and classifiers. Here the feature space is tuned with use of neural networks where Resource Allocating Network and Long Term Memory model is used.

SVM are found to be effective in large number of classification and regression problem. [20] Discuss of their use in the optical character recognition has found a new area, with ensembles of SVMs in operation. Ensemble based methods are further used with dynamic weighting scheme that is built for extra training models over the earlier ones giving incremental approach for new training data sets when available in batches as suggested by [21]. [22] Focus on ensemble methods to learn concept drift; that are characterized by non-stationary environment applying the approach on weather prediction system where the Learn++ approach is extended.

7. CURRENT SCENARIO

With the related work that is been taking place, the objective of this paper is to make the researcher familiar with the concepts. Currently work that is been carried out focuses on query formulation and data distributions.

To focus a few, [23] present a survey on the different techniques of incremental learning of HMM parameters. The work reviews batch learning techniques for estimation of HMM parameters, trying to remove the impact of use of HMM as priori methods with limited data.

Further, incremental and reinforcement learning and whole system learning could be the next big thing. [24] Propose use of incremental reinforcement learning designed for operation in multi- agent scenario. The work is based on modified version of Q-Learning, where the agent is faced with number of tasks that it learns.

Incremental learning in query formulation is explored by [25], where it assists the user to form query, from arbitrary to structured one trying to bridge the gap between the queries formed to the effectiveness that it can be retrieved.

Very recently incremental approach for managing the concept drift of data distributions [26] is proposed. The domain identified was remote sensing images for land cover classifications. With the kernel function in operation and on the basis of Markov chain, the learning process is active, making the addition and deletion of the training vectors.

8. DISCUSSION AND CONCLUSION

The paper tries to highlight the areas and methodologies that presented incremental approach, thought still lot more are to be explored. Considering the various areas and the work that is taking place, it is for the researcher to have a broad view and work on domain where the process of incremental learning will help in making strong decisions. Incremental learning that is selective in terms of the datasets used at the same time adaptive and dynamic with ability to take decision accurately is what currently looked at. Accuracy that considers impact of the decisions taken should equally be taken into consideration for the same. The following table tries to summarize the domains and the methods that are in use. (Since it is impossible to cover each and every aspect of the

Methods and domain, there could be more applications and domain that have been remained unexplored in this paper.)

Table 1: Incremental methods and domains: A summary of survey

Methods/ Algorithms	Application Domain
SVM, Ensemble methods, Dynamic weighing	Optical character/ Text document
Neural Network, Centroid based methods	Web pages / Document layout
Minimum bounding boxes	Mobile
Pattern matching – Neural Network	Text classification
Ensemble based methods, Learn++, ADAIN	Digital optics
Gaussian distributions, Neural Networks	Students behavior pattern
Bayesian, GRIN, BIRCH, DBSCAN	Relational databases/ Warehouse
Bayesian learning, Resource Allocation Network	Medical Image segmentation/ Sports video
Concept drift, Ensemble methods	Weather Prediction/ Data streams

The point is not where one intends to work that is in supervised or unsupervised environment but to identify the domain where the learning will be effective at the same time can it be used further for forecasting. When we refer to forecast, it can be with respect to the weather, the sales of an industry, the attrition rate and so on. At this point the methods and approaches proposed aim at giving accuracy at the same time better decisions. The point of discussion here is not just the application that you want, but the reason you want to have incremental learning.

With the current work that has been done in various fields, one more factor that has to be discussed is to evolve incremental learning with a feedback mechanism. The impact of the decisions and learning from the impacts and incorporating these decisions for further learning is required. The main aspect that can be part of further research is this decision support mechanism, that itself will evolve with every new decision taken, irrespective of the application it is been employed at.

Finally to have this, it equally essential to identify what statistical methods would be used and for what purposes. Each and every existing method has its own flavor and pre-requisites and that is what should be exploited further to come up with novel methods that can be used for incremental learning.

REFERENCES

1. V. Vapnik, Statistical Learning Theory. New York: Wiley, 1998.

2. N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.

3. R. French, "Catastrophic forgetting in connectionist networks: Causes, Consequences and Solutions," Trends in Cognitive Sciences, vol. 3, no.4, pp. 128-135, 1999.

4. P. Jantke, "Types of incremental learning," Training Issues in Incremental Learning, A.Cornuejos (ed.) pp.26-32, The AAAI Press: Menlo Park, CA 1993.

5. M.A. Maloof and R.S. Michalski," Selecting examples for partial memory learning," Machine Learning, vol. 41, pp. 27–52, 2000.

6. F.S. Osario and B. Amy," INSS: A hybrid system for constructive machine learning," Neurocomputing, vol. 28, pp. 191–205, 1999.

7. J. Ghosh and A,c, Nag," Knowledge enhancement and reuse with radial basis function networks," Proceeding of International Joint Conference on Neural Networks, vol. 9, no. 2, pp. 1322–1327, 2002.

8. S. Ozawa, S. Pang, N. Kasabov, Incremental Learning of chunk data for online pattern classification systems, IEEE Transactions on Neural Networks, Vo. 19 (6), 2008, pp 1061-1074.

9. Z. Chen, L. Huang, Y. Murphey, Incremental learning for text document classification, Proc. of IEEE Conference on Neural Networks, 2007, pp 2592-2597.

10. R. Polikar, L. Upda, S. Upda, V. Honavar, Learn ++: An incremental learning algorithm for supervised neural networks, IEEE Transactions on Systems, Man and Cybernatics, Vol.31 (4), 2001, pp 497-508.

11. H. He, S. Chen, K. Li, X. Xu, Incremental learning from stream data, IEEE Transactions on Neural Networks, Vol.22(12), 2011, pp 1901-1914.

12. A. Bouchachia, M. Prosseger, H. Duman, Semi supervised incremental learning, Proc. of IEEE International Conference on Fuzzy Systems, 2010 pp 1-7.

1. 13. R. Zhang, A. Rudnicky, A new data section principle for semi-supervised incremental learning, Computer Science department, paper 1374, 2006, http://repository.cmu.edu/compsci/1373.

13. Z. Li, S. Watchsmuch, J. Fritsch, G. Sagerer, Semi-supervised incremental learning of manipulative tasks, Proc. of International Conference on Machine Vision Applications, 2007, pp 73-77.

14. A. Misra, A. Sowmya, P. Compton, Incremental learning for segmentation in medical images, Proc. of IEEE Conference on Biomedical Imaging, 2006.

15. P. Kranen, E. Muller, I. Assent, R. Krieder, T. Seidl, Incremental Learning of Medical Data for Multi- Step Patient Health Classification, Database technology for life sciences and medicine, 2010.

16. J. Wu, B. Zhang, X. Hua, J, Zhang, A semi-supervised incremental learning framework for sports video view classification, Proc. of IEEE Conference on Multi-Media Modelling, 2006.

17. S. Wenzel, W. Forstner, Semi supervised incremental learning of hierarchical appearance models, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol.37, 2008.

18. S. Ozawa, S. Toh, S. Abe, S. Pang, N. Kasabov, Incremental Learning for online face recognition, Proc. of IEEE Conference on Neural Networks, Vol. 5, 2005 pp 3174-3179.

19. Z. Erdem, R. Polikar, F. Gurgen, N. Yumusak, Ensemble of SVMs for Incremental Learning, Multiple Classifier Systems, Springer Verlang,, 2005, pp 246-256.

20. X. Yang, B. Yuan, W. Liu, Dynamic Weighting ensembles for incremental learning, Proc. of IEEE conference in pattern recognition. 2009, pp 1-5.

21. R. Elwell, R. Polikar, Incremental Learning of Concept drift in nonstationary environments, IEEE Transactions on Neural Networks, Vol.22 (10), 2011 pp 1517- 1531.

22. W. Khreich, E. Granger, A. Miri, R. Sabourin, A survey of techniques for incremental learning of HMM parameters, Journal of Information Science, Elsevier, 2012.

23. O. Buffet, A. Duetch, F. Charpillet, Incremental Reinforcement Learning for designing multi-agent systems, Proc. of ACM International Conference on Autonomous Agents, 2001.

24. E. Demidova, X. Zhou, W. Nejdl, A probabilistic scheme for keyword-based incremental query construction, IEEE Transactions on Knowledge and Data Engineering, 2012, pp 426-439.

25. R. Roscher, W. Forestner, B. Waske, I2VM: Incremental import vector machines, Journal of Image and Vision Computing, Elsevier, 2012.