Runtime Optimization of Widrow-Haff Classification Algorithm Using Proper Learning Samples

Mir-Hossein Dezfoulian*[1]
Assistant Professor
dezfoulian@basu.ac.ir

S.Younes Mirinezhad[2]
Student
y.mirinezhad93@basu.ac.ir

S. M. Hossein Mousavi[3]
Student
h.mosavi93@basu.ac.ir

Mehrdad Shafei Mosleh[4]
Student
m.shafaei93@basu.ac.ir

Abstract

This study works on the runtime optimization of Widrow-Hoff classification algorithm. The use of proper learning samples has a significant effect on the runtime and accuracy of supervised classification algorithms, in special Widrow-Hoff classification algorithm. In this study with synthesizing Multi Class Instance Selection (MCIS) algorithm and Widrow-Hoff classification algorithm, the runtime of algorithm has significantly reduced. Results of this, of vantage sample of accuracy and time, have been assessed, and simulations are indicating MCIS with the aid of proper measures is able to select the data having most effectiveness on classification. In the case, if Widrow-Hoff classifier has less and important samples (achieved by MCIS), it would be able to save significant amount of time in classification process.

Keywords: Widrow-Hoff, Classification, Learning samples, Runtime Optimization, MCIS

Introduction

With advancement of technology in recent years, in particular, in the field of computer science, the way of calculation process has been smoothed. Taxonomy of data for research and science studies in different fields such: medicine, agriculture, marine industry, chemistry etc. is of great importance. With taking aid of different classification algorithms, it is possible to classify the row data. There are different databases for this purpose, and daily the efforts are made in order to optimize classification algorithms or to propose new classification algorithms. Hitherto, many algorithms for the purpose of classification have been proposed, each having its pros and cons. One disadvantage for classifiers is the effect of out-of-center (outlier) or noisy data on the classifier. A variety of different methods are presented for eliminating the data with destructive effect on the speed of the classifier. MCIS [1] is a method which this paper get the aid, to optimize the runtime of Widrow-Hoff algorithm [2].

An Overview on Prior Works

Up to now, for the classification of data many methods have been proposed, which having its pros and cons. In 1963, for the first time the SVM was invented by Vapnik, and in 1995 was developed as non-linear state [3]. The basis of the work it does, is the linear classification of data, and makes effort to select the line with the most safe-margin. Learning this algorithm is almost easy, and for the data with high dimensions is works pretty well. In 1965, Steinbuch, K, B. Widrow innovated an algorithm for two compatible classification channels, being able to be used as classification algorithm. They named it Widrow-Hoff [2]. This algorithm has a high speed, but is sensible to noisy and outlier data. In 1967, Cover and Hart proposed an algorithm (Knn) which was able to classify and approximate considering K to the nearest data to the intended data [4]. In 1984, Poole.L,G.Warnaka,R.Cutter made a digital filter using Widrow-Hoff algorithm for audio noise reduction [5]. Also, in 2014, Rodriguez_Quinones raised the measurement accuracy of the 3D laser scanner using NNs and Widrow-Hoff [6]. In 2016, Viswanatha, Vishwanath could classify the face accurately using the linear discriminative regression learning [7]. In 2013, Jingnian Chen and his colleagues proposed a method for optimizing SVM by eliminating outlier data [1], [8]. The basis of works this algorithm does, is analysis for the data distances from each other, and succeeded to recognize non-outlier data.

Proposed Method

As mentioned before, one of the problems in finding the classifier is the existence of outlier or noisy data. In this study for optimizing the functionality of Widrow-Hoff algorithm, MCIS has been engaged for eliminating the outlier data. It works as first, the existing data using MCIS will be analyzed and the more suitable data for classification using Widrow-Hoff algorithm, will be chosen. Then the classifier structures from the reduced data. The less the number of existing data is, the faster the speed of classifier would be. This acceleration would be more tangible while working on large data.

3-1- MCIS

First, we select the one class data (accidentally) as positive class, and other data (the data of other classes) as for negative data. Afterwards, the data of positive class is clustered by K-Means, and the center of the cluster is considered as the representative of the cluster. Then the distance of all data of negative class will be calculated from the centers of positive class cluster, and label the data possessing a distance less than r as non-outlier(close to other classes) data. It should be observed the smaller r is, the lower the number of samples which label as non-outlier samples would be. These actions keep on going for all the existing classes till non-outlier data of all classes are labeled.

3-2- Classification

After proper data using MCIS are chosen, they will classify using Widrow-Hoff algorithm. Due to some data not being involved in this process of classification, it is expected the runtime of classification process be less than when Widrow-Hoff classifies using all the data, and considering Widrow-Hoff being sensitive to the outlier data, and these data get eliminated by MCIS, it is expected there would be changes in the accuracy of classification.

Experiments and Validations

For denoting the influence of choosing proper learned data on Widrow-Hoff classifier, first the non-outlier data using MCIS is identified, and then this data using Widrow-Hoff is classified and the results will be compared with other classification algorithms such as: KNN, SVM, MLP [9], FSM [10] and Widrow-Hoff algorithm.

4-1- Experiment Databases

Experiments took place on Pima-diabetes [11], Ionophere [12], Liver-Disorders [13], Breast-cancer [14], and in Table 1, name of each database and features, classes and chosen samples for learning and experiment is given. Each database uses 70% of samples for training and 30 for test.

Table 1-Databases and their features

	Pima diabetes	Ionosphere	Liver Disorders	Breast cancer Wisconsin
No. of samples	768	351	345	699
No. of features	8	34	7	10
No. of selected features	8	34	7	10
No. of classes	2	2	2	2
No. of learning samples	538	246	242	489
No. of test samples	230	105	103	210

4-2- Hardware

Due to comparison of Widrow-Hoff algorithm runtime with all data and chosen ones being dependent to a proper hardware, all the experiments were experimented using a system with Intel® Core™ i7-2670QM and 4GB of RAM. Considering the aim of study being the comparison of the accuracy of algorithms with the accuracy of Widrow-Hoff algorithm, the related results to classification algorithms FSM, MLP, SVM and KNN were extracted of related articles and web**.**

4-3- Results and Validations

In this paper, first Widrow-Hoff classifier was analyzed for observing the effect of outlier data. As Table 2 indicates, after eliminating outlier data (data with a distance more than r), runtime reduces significantly. This is better tangible while databases with higher dimensions are engaged, and with specification of proper k and r, it is possible to change the runtime.

Table 2-Comparison of Widrow-Hoff with all data and Widrow-Hoff with selected data

	Pima diabetes k = 2 r = 1.3	Ionosphere K = 4 و r = 6.7	Liver Disorders K = 1 و r = 0.2	Breast Cancer Wisconsin K = 3 و r = 1.8
accuracy Widrow-Hoff	65.8%	84.7%	40.2%	65.7%
runtime Widrow-Hoff	0.76 s	1.09 s	0.002 s	0.004 s
accuracy MCIS + Widrow-Hoff	65.8 s	85.6 s	60.7 s	65.07 s
runtime MCIS + Widrow-Hoff	0.0007 s	0.7 s	0.0001 s	0.001 s

For the purpose of comparing the accuracy of Widrow-Hoff after eliminating outlier data with other classification algorithms, the result of performance of each algorithm is performed in Table 3, which is extracted from following link: http://www.is.umk.pl/projects/datasets.html. Table 4 also presents the comparison results of the speed of Widrow-Hoff with the proposed method.

Table 3-: Comparison of the accuracy of typical algorithms with the proposed method

	MCIS + Widrow-Hoff	Widrow-Hoff	SVM	KNN	MLP	FSM
Pima diabetes	65.8%	65.8%	77.5%	76.7%	76.4%	75.4%
Ionosphere	85.6%	84.7%	93.0%	98.7%	96.0%	92.8%
Breast cancer Wisconsin	65.07%	65.07%	96.9%	97.1	96.7%	98.3%

Table 4: Comparison of the speed of Widrow-Hoff with the proposed method

	MCIS + Widrow-Hoff	Widrow-Hoff
Pima diabetes	0.0007 s	0.004 s
Ionosphere	0.7 s	1.9 s
Breast cancer Wisconsin	0.001 s	0.004 s
Liver Disorders	0.0001 s	0.002 s

Conclusion and Suggestion

Considering the table results are indicating, it can be said the more there are outlier data in the database, meaning the more data volume and the higher their dimension is, runtime of classification algorithms increases. Therefore, when some data are put away (outlier data), classification takes place using less data, and the speed of classification increases. In particular, rise of classification pace while on a big database poses itself more tangibly. it is suggested that combine MCIS algorithm with other classification algorithms such Fisher and Etc.

References

[1] Chen, Jingnian, et al. "Fast instance selection for speeding up support vector machines." Knowledge-Based Systems 45 (2013): 1-7.

[2] Steinbuch, K., and B. Widrow. "A critical comparison of two kinds of adaptive classification networks." IEEE Transactions on Electronic Computers 5 (1965): 737-740.

[3] Cortes, Corinna, and Vladimir Vapnik. "Support vector machine." Machine learning 20.3 (1995): 273-297.

[4] Cover, Thomas M., and Peter E. Hart. "Nearest neighbor pattern classification." Information Theory, IEEE Transactions on 13.1 (1967): 21-27.

[5] Poole, L., G. Warnaka, and R. Cutter. "The implementation of digital filters using a modified Widrow-Hoff algorithm for the adaptive cancellation of acoustic noise." Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.. Vol. 9. IEEE, 1984.

[6] Rodríguez-Quiñonez, J., et al. "Improve 3D laser scanner measurements accuracy using a FFBP neural network with Widrow-Hoff weight/bias learning function." Opto-Electronics Review 22.4 (2014): 224-235.

[7] Vishwanath, P., and V. M. Viswanatha. "FACE CLASSIFICATION USING WIDROW-HOFF LEARNING PARALLEL LINEAR COLLABORATIVE DISCRIMINANT REGRESSION (WH-PLCDRC)." Journal of Theoretical and Applied Information Technology 89.2 (2016): 362.

[8] Burges, Christopher JC. "A tutorial on support vector machines for pattern recognition." Data mining and knowledge discovery 2.2 (1998): 121-167.

[9] Pal, Sankar K., and Sushmita Mitra. "Multilayer perceptron, fuzzy sets, and classification." Neural Networks, IEEE Transactions on 3.5 (1992): 683-697.

[10] Zeng, Zheng, Rodney M. Goodman, and Padhraic Smyth. "Learning finite state machines with self-clustering

recurrent networks." Neural Computation5.6 (1993): 976-990.
[11] Smith, J.W., Everhart, J.E., Dickson, W.C., Knowler, W.C., & Johannes, R.S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In Proceedings of the Symposium on Computer Applications and Medical Care} (pp. 261--265). IEEE Computer Society Press.
[12] Sigillito, V. G., Wing, S. P., Hutton, L. V., \& Baker, K. B. (1989). Classification of radar returns from the ionosphere using neural networks. Johns Hopkins APL Technical Digest, 10, 262-266.
[13] PC/BEAGLE User's Guide (written by Richard S.Forsyth) tekken from : https://archive.ics.uci.edu/ml/datasets/Liver+Disorders
[14] Wolberg, W.H., & Mangasarian, O.L. (1990). Multisurface method of pattern separation for medical diagnosis applied to breast cytology. In Proceedings of the National Academy of Sciences, 87, 9193—9196.