Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 3, March 2016, pg.76 – 80

ONLINE SENSOR BASED C4.5 ALGORITHM FOR STRUCTURAL HEALTH MONITORING

Author Name 1: Rajashree.N

Department of CSE

IFET College of Engineering,

Villupuram

Author Name 2: Vishnu Priya.R

Ass.Prof, CSE

IFET College of Engineering

Villupuram

Abstract-The process of implementing a damage detection and characterization strategy for engineering structures is referred as Structural Health Monitoring. Structural health monitoring (SHM) uses building vibration responses from online sensors to diagnose and evaluate possible structural damages. This is a data classification problem. The builder states come from many on-line sensors. Normal classification methods, such as support vector machine (SVM), cannot classify this large data stream. In this paper, the convex–concave hull and SVM are modified into online versions, such that the novel data classifiers for the SHM can be trained online. We design a laboratory-scale prototype for the experimental evaluation. The proposed method using C4.5 algorithm is widely used because of its quick classification and high precision. Experiment results show that the various entropy based approach is effective in achieving a high classification rate.

Key words: Structural health monitoring, support vector machine, building vibration responses, damage detection.

I. INTRODUCTION

Data mining is an interdisciplinary subfield of computer science It is the computational process of discovering patterns in large datasets("big data") involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating. Data mining is the analysis step of the "knowledge discovery in databases" process, or KDD. Large quantities of data to extract previously unknown, interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection), and dependencies (association rule mining). This usually involves using database techniques such as spatial indices.

Historically, the monitoring of structures has involved many ingredients of the modern SHM paradigm, such as data collection and processing followed by diagnosis. At the simplest level, recurrent visual observation and assessment of structural condition (cracking, spalling and deformations) could be viewed as SHM, yet the aim of present-day research is to develop effective and reliable means of acquiring, managing, integrating and interpreting structural performance data for maximum useful information at a minimum cost while either removing or supplementing the qualitative, subjective and unreliable human element. Historical developments in SHM have generally focused on subsets of the SHM paradigm, but in recent years, a few research teams have begun to focus on, or at least recognize (Fanelli 1992), the need for a holistic approach to optimization of SHM.

II. RELATED WORK

V. Alves [1] Data driven SHM methodologies take raw signals obtained from sensor networks, and process them to obtain features representative of the condition of the structure. New measurements are then compared with baselines to detect damage.

G. Cauwenberghs [2] An on-line recursive algorithm for training support vector machines, one vector at a time, is presented. Adiabatic increments retain the Kuhn Tucker conditions on all previously seen training data, in a number of steps each computed analytically. The incremental procedure is reversible, and decremental “unlearning” offers an efficient method to exactly evaluate leave-one-out generalization performance.

J. Cervantes [3] The reduced support vector machine (RSVM) is extension method of smooth support vector machine (SSVM) for handling computational difficulties as well as reduces the model complexity by generating a nonlinear separating surface for a large dataset.

Y. Engel [4] Sparsity of the solution is achieved by a sequential sparsification process that admits into the kernel representation a new input sample only if its feature space image cannot be sufficiently well approximated by combining the images of previously admitted samples. This sparsification procedure is crucial to the operation of KRLS, as it allows it to operate on-line, and by effectively regularizing its solutions.

III. PROPOSED WORK

Compared with the other SVM classifiers, our approach has a classification accuracy as good. The approach is significantly faster than all the other classifiers. A c4.5 technique is used to improve the damage detection. Find a minimum execution time of the model that has maximum accuracy. Classify by that model which has minimum execution time.

Modules used in proposed system are given,

1. Data Preprocessing
2. Clustering With Convex–Concave Hull
3. Online SVM And C 4.5
4. SHM Via Online SVM

Data Preprocessing

This step is to preprocess extracted words. There are four aspects: splitting, eliminating stop words, stemming and removing specific tags. Splitting: combined words are split into verbs and nouns. Because of naming convention, verb and noun are combined in operation names, service names and messages of web services. Eliminating stop words: words are filtered by stop list. The stop list contains a few common words (including preposition, article, etc.), which are recognized that they are not useful in information retrieval.

Clustering With Convex–Concave Hull

We have used the convex–concave hull method to find extreme points of a large data set and trained the SVM classifier with only the extreme points (which are much less than original data). However, the convex–concave hull method is a batch process, it cannot be applied for online dynamic data set. Here, we modify the convex–concave hull method using online clustering strategy. The obtained online convex–concave hull can find extreme points from online data stream. The objective of the convex–concave hull method is to detect the objects that are close to exterior boundaries.

Online SVM And C4.5

SVM classification can be grouped into two types: linearly separable and linearly inseparable cases. In linearly inseparable case, the convex–concave hulls B(X) are intersected. Because support vectors are generally located on the exterior boundaries of data distribution, they are not vertices of $B(X+)$ and $B(X-)$. On the other hand, the vertices of the convex–concave hull are the border points, and they are possible to be the support vectors. In our online SVM classification, the penal factor can be selected very small, because all misleading points almost disappear by the concave algorithm.

SHM Via Online SVM

The building vibration data are measured from the accelerometers on each floor of the building. After numerical integrations, the displacement and velocity data are used for SHM. Novelty detector. In order to obtain labels of the data in SHM, Novelty detector is needed . SVM Classifier-This is a main block for SHM. After the SVM is trained in the block “SVM Training,” it is used as a classifier to detect damages. Outliers detection. The block compares the labels from SVM classifier and Novelty detector. If the labels of the aforementioned two blocks are the same, the block just outputs this label. Online Clustering.-In this block, the borders of the convex–concave hull are calculated online using the results SVM Training- SVM is trained with the data set obtained by Online Clustering; trained SVM is in the SVM Classifier block for identifying damage. The system architecture is given,

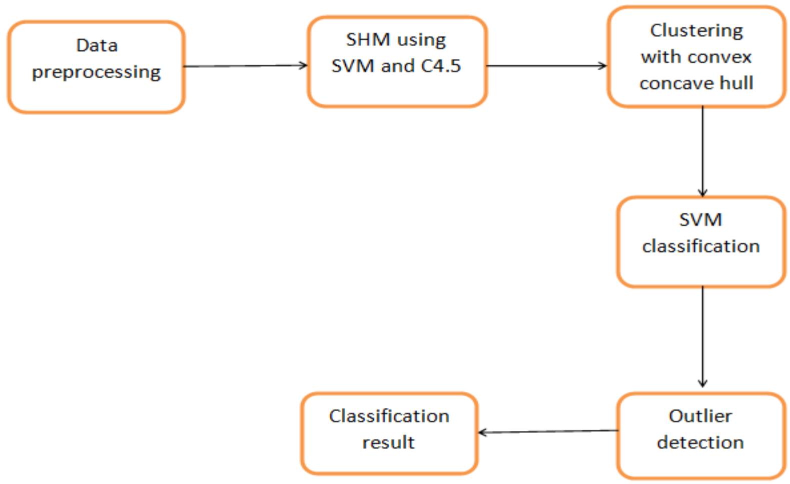

Fig 1.system architecture

IV. C4.5 ALGORITHM

C4.5 algorithm which highlights some efficiency improvements. C4.5 by adopting the best among three strategies for computing the information gain of continuous attributes. All the strategies adopt a binary search of the threshold in the whole training set starting from the local threshold computed at a node. The first strategy computes the local threshold using the algorithm of C4.5, which, in particular, sorts cases by means of the quicksort method. The second strategy also uses the algorithm of C4.5, but adopts a counting sort method. The third strategy calculates the local threshold using a main-memory version of the Rain Forest algorithm, which does not need sorting. Our implementation computes the same decision trees as C4.5 with a performance gain of up to five times.

V. SYSTEM IMPLEMENTATION

The system output is mainly based upon building structures damage detection.it will be evaluated using c4.5 algorithm. The evaluation data can be done by the convex–concave hull method using online clustering strategy SVM classification can be grouped into two types: linearly separable and linearly inseparable cases. In linearly inseparable case, the convex–concave hulls B(X) are intersected. the convex–concave hull are calculated online using the results SVM Training- SVM is trained with the data set obtained by Online Clustering; trained SVM is in the SVM Classifier block for identifying damage.

VI. CONCLUSION

The process of implementing a damage detection and characterization strategy for engineering structures is referred as Structural Health Monitoring. Compared with the other SVM classifiers, our approach has a classification accuracy as good. Experiment results show that the various entropy based approach is effective in achieving a high classification rate.

REFERENCES

- [1] V. Alves, "Novelty detection for SHM using raw acceleration measurements," Struct. Control Health Monit., vol. 22, no. 9, pp. 1193–1207, Sep. 2015.
- [2] K. P. Bennett and E. J. Bredensteiner, "Geometry in learning," in *Geometry atWork*, C. Gorini, Ed. Washington, DC, USA:Mathematical Association of America, 2000, pp. 132–145.
- [3] G. Cauwenberghs and T. Poggio, "Incremental and decremental support vector machine learning," in *Advances in Neural Information Processing Systems*, vol. 13. Cambridge MA, USA: MIT Press, 2001.
- [4] J. Cervantes, X. Li, W. Yu, and K. Li, "Support vector machine classification for large data sets via minimum enclosing ball clustering," *Neurocomputing*, vol. 71, no. 4–6, pp. 611–619, Jan. 2008.
- [5] C.-C. Chang and C-J. Lin, LIBSVM: A library for Support Vector Machines, 2001. [Online].
- [6] A. K. Chopra, *Dynamics of Structures: Theory and Application to Earthquake Engineering*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.
- [7] Y. Engel, S. Mannor, and R. Meir, "The Kernel recursive least-squares algorithm," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2275–2285, Aug. 2004.
- [8] W. Eddy, "A new convex hull algorithm for planar sets," *ACM Trans. Math. Softw.*, vol. 3, no. 4, pp. 398–403, Dec. 1977.
- [9] C. R. Farrar and Worden, *Structural Health Monitoring: A Machine Learning Perspective*. Hoboken, NJ, USA: Wiley, 2013.
- [10] E. Forstner and H. Wenzel, "The application of data mining in bridge monitoring projects: Exploiting time series data of structural health monitoring," in *Proc. 22nd Int. Workshop Database Expert Syst. Appl.*, 2011,pp. 297–301.