A New Fuzzy Based Ensemble Classifier for Analysis of ECG Signal

Girish B Umaratkar¹, Sachin A Murab²

¹Student, ²Assistant Professor

Abstract: ECG(electrocardiogram) reflects the state of cardiac heart and hence is like a pointer to the health conditions of a human being. However, ecg being anon-stationary, continuous in nature and abruptly changing signal, the irregularities may not be periodic and may show up at different intervals. For taking intelligent health care decisions, ecg signal needs to be analyzed accurately. Clinical observation of ecg can take long hours and can be very tedious. Moreover, visual analysis cannot be relied upon. Thus, our basic objective is to come up with an ensemble based classification technique that will classify ecg signal with the more accuracy. This objective has motivated us to search and experiment with various ecg signals by categorizing it in correct class and simultaneously achieving maximum accuracy of the ensemble classifier. This paper deals with the implementation of a fuzzy based ensemble classifier that performs the computations by using fuzzy inference system (FIS) to classify the ecg and to achieve the maximum accuracy. Overall, we have tried to minimize the concept drift evolved in the ecg signal and maximize the accuracy because the error rate introduced due to concept drift is inversely proportional to the accuracy of ensemble based classifier. The result shows that the ensemble classifier with the fuzzy based technique is more accurate up to 99% in classification of ecg signal.

Keywords: ensemble, ecg, fuzzy system.

I. INTRODUCTION

Data Stream Mining can be considered a subfield of data mining, machine learning, and knowledge discovery. Data Stream Mining is the process of extracting meaningful information from continuous, rapid data streams that have many attributes and terms. It refers to informational structure extraction as models and patterns from continuous data streams. Data Streams have different challenges in many aspects, such as computational, storage, querying and mining as it can be conceived as a continuous and changing sequence of data that continuously arrive at a system to store or process. While processing the data, noise, errors, unwanted data, missing values have to be removed. There are many proposed classification algorithms for concept drifting data streams. These algorithms support multidimensional analysis and decision making. Additional data analysis techniques are required for in-depth analysis, characterization of data changes over time. In addition, huge volumes of data can be accumulated beyond databases and data warehouses. In applications like video surveillance, weather forecasting, telecommunication, sensor networks, satellites, call records, vital signals monitoring data stream mining plays a key role to analyze the continuous data. The meaningful, effective and efficient analysis of this data in such different forms becomes tedious task and also the issue of memory constraints has to handle as enormous data is generated continuously.

Fig 1: General Process of Data Stream Mining [9]
In many data stream mining applications, the goal is to predict the class or value of new instances in the data stream given some knowledge about the class membership or values of previous instances in the data stream. Machine learning techniques can be used to learn this prediction task from labeled examples in an automated fashion. In many applications, the distribution underlying the instances or the rules underlying their labeling may change over time, i.e. the goal of the prediction, the class to be predicted or the target value to be predicted, may change over time. This problem is referred to as concept drift.

To address the problem of concept drifting in real time data stream mining, contributions have been made in literature regarding various ensemble based classification algorithms techniques for health care application [4]. The accuracy of the ensemble based classification algorithms is still challenging for taking intelligent health care decision. Hence, the work is focused on the ensemble based classification technique which has to be applied on the vital signal ECG of human body.

II. DATABASE DETAILS

An ECG signal is composed of successive repetition of “PQRST” in monotony. In the beginning, a crust is generated from the linear signal to form the “P” wave. The declining linear wave soon gets a downward deflection labeled as “Q” wave. A sudden upright deflection can be observed just beyond the Q wave to form a highcone that is “R” wave. On its decline a slight downward deflection is the “S” wave. A noticeable hinge after the S wave is known as “T” wave that marks the end of a segment of the ECG signal.

III. FLOWCHART OF THE DESIGN

![Flowchart of proposed design](image-url)
IV. ENSEMBLE BASED CLASSIFICATION

An ensemble classifier is a classification algorithm that learns a set of classifiers instead of learning a single classifier, and then combines the predictions of these classifiers to produce the final prediction. The key step of an ensemble classifier is forming an ensemble of diverse classifiers from a single training set.

V. PROPOSED FUZZY BASED ENSEMBLE CLASSIFICATION ALGORITHM

**Input:** Data stream with class labels available from generated database.
- x = Permitted error
- y = Classifier Precision
- z = Number of Classifier
- w = Number of ECG records

**Output:** ECG Class, Ensemble Classifier error rate.

**Classification of Testing data**
1. Classifier \((1) \leftarrow \text{Ensemble set of } (Z1, Z2, Z3, ..., Zl)\) Classifier.
2. CE \(\leftarrow\) Classifier Error of Z Classifier in matrix of zeros.

An FIS Training:
3. New Classifier Required \(\leftarrow\) true;
4. Get data chunk T from input stream with class label
5. CE\(_{i}\) = classification error for data set W using Classifier \(Z_{i}\)
6. while (New Classifier Required)
7. New Classifier Required \(\leftarrow\) false;
8. for classifier \(Z_{i}\) in Classifier
9. GO TO An FIS Training
10. anfi_output = Evaluate FIS training
11. Class number = round (anfi_output)
12. if CE\(_{i}\) <= x
13. New Classifier Required \(\leftarrow\) true;
14. GO TO an FIS Training
15. endif
16. endif
17. endfor
18. for classifier \(Z_{i}\) in Classifier
19. if CE\(_{i}\) <= min error
20. min error = CE\(_{i}\)
21. endif
22. endif
23. endfor
VI. EXPERIMENTAL RESULTS

Fig. 5: Original ECG Signal 100.dat

Fig. 6: Normalized ECG Signal 100.dat

Fig. 7: Wavelet decomposition for 100.dat
VII. APPLICATION
The proposed work leads the building of Real-Time Data Stream Mining system that can analyze medical data stream and make real-time prediction of ECG signal with more accuracy. The RT-DSM system can help the medical professionals that can be used in medical field of prognosis and diagnosis the chronic diseases related with heart such as arrhythmia, premature ventricle contraction, branch bundle block, atrial flutter etc. Most of the existing software technologies are case-based data mining systems. They only can analyze finite and structured data set and can only work well in their early years and can hardly meet today's medical requirement.

An objective of a health process is one where patients can stay healthy with the support of expert medical advice when they need it, at any location and any time. An associated aim would be the development of a system which places increased emphasis on preventative measures as a first point of contact with the patient. As the vital signal ECG is a pointer for predicting health status of a human heart we motivated to apply RT-DSM ensemble based algorithm on ECG signal.

VIII. CONCLUSION
In this paper we have investigated the major issues in classifying large volume, high speed and dynamically changing streaming...
ECG data and proposed a fuzzy inference system (FIS) approach.

the experiments are carried out on vital signal ECG of human body and applied the fuzzy based ensemble classifier on it. The result shows that the ensemble classifier with the fuzzy based technique is more accurate up to 99% in classification of ECG signal.

REFERENCES

[5] Charu C. Aggarwal, Jiawei Han, Jianyong Wang and Philip S. Yu, Fellow, IEEE, “A Framework for On-Demand Classification of Evolving Data Streams”, IEEE Transactions on Knowledge and Data Engineering, Vol. 18, No. 5, may 2006, pp. 577-589.
[10] Mohammad M. Masud, Jing Gao, Latifur Khan, Jiawei Han, “A Practical Approach to Classify Evolving Data Streams: Training with Limited Amount of Labelle Data”, IEEE International Conference on Data Mining, 2008, pp. 929-934.
[12] JingGao,WeiFan,Jiawei Han, “On Appropriate Assumptions to Mine Data Streams: Analysis and Practice”,IEEE International Conference on Data Mining - ICDM, DOI: 10.1109/ICDM.2010.125, 2010, pp. 1175-1180.


