Systematic Review of Different Drift Detection Techniques of On-line Streaming

Rangari Sudhir

PhD Student Kalinga University, Raipur

Dr. Priti Singla

PhD Guide Kalinga University, Raipur

Abstract

In this computerized time creator are encircled by online media applications and the equipment gadgets, (for example, sensors and so forth) which are pouring data at a surprising rate. This approaching data from heterogeneous sources is eluded as data stream. Examining data moving (data streams) has become new test to satisfy the needs of constant examination. Traditional mining strategies are demonstrating wasteful since the conduct of data itself has changed. Different difficulties related with data streams incorporate assets limitations like memory and running time alongside single output of the data. Because of the time variation nature of data streams, applying any mining calculation, for example, characterization, bunching, ordering in a solitary output of data is a drawn-out assignment. This paper centers on idea drift issue in characterization of streaming data. The paper additionally records the different datasets and execution measurements that have been utilized in writing for execution investigation.

Keywords-: Drift detection, online streaming, concept drift, data streaming

1. Introduction

Data mining is an imminent study to resolve several problems, and classification is certainly one of predominant hassle inside the discipline of DM [3]. It is the technique to drag out patterns from large datasets, and machine learning uses these patterns and various other data mining algorithms to predict the result. DM is used to discover the unseen data sample and the association beta uthoren different dataset. DM is a technique to study about the big database. Facts mining are related to information Discovery. Classification, Clustering and Association rule discovery are the essential strategies of DM. The primary goal of data mining is Predictive and Descriptive. Predictive mining presents the person to post facts with unknown area values. Descriptive mining presents record to recognize without a predetermined concept [4].

Nowadays, data mining has been taken into great consideration for research because of its capabilities to filter the useful information from the large databases and such filtered information can be used for prediction purposes. The data mining can be used commercially in different industries as retail, tale-communication, energy industries etc. [3]. The essential flavours of the architecture for a regular information mining system are verified [3]. Real device structure will make viable the information mining device to make notable use of the software environment. It achieves statistics mining obligations in an effective and right manner to alternate information with exceptional structures; this is adaptable to clients with various requirements and alternate with time.

A Multilayer Perception (MLP) is a feed forward artificial neural model that is used to get a suitable output with a set of input. This is basically a directed graph consisting of many layers of nodes and every layer is linked with other layer. The first layer is named an input layer and the last one is termed as an output layer. In beta uthoren of these two layers, there may be few other intermediary layers. These layers are called as hidden layer. The first hidden layer gets its input from the input layer and makes some computation over the data before forwarding to next hidden layer. Similarly every hidden layer performs some computation before forwarding the data and finally sends the data to output layer [14].

ISSN: 2233-7857 IJFGCN

The measurement of data made within the last 2 years is larger than the history of mankind's entire past. The data is getting snapper than at any time and as indicated by the data carried by the IDC every year in 2025 it needs to reach 180 zeta bytes. The data is never dozens; Monstrous measurements of data are being produced from mixed applications on the web at regular intervals. 2, 000, 00 test queries and email clients continuously send 204166,667 messages to Google. It is seen that a huge development in video data, where 300 hours of continuous recording is transferred to YouTube alone. These realities reflect the speed with which data is evolving in a way that creates a computerized universe. Speed or speed is one of the primary features of huge data. Innovative advances in devices are additionally accountable for the increasing fame of data streams. Many basic functions of routine day-to-day existence, for example, using credit / check cards or PDAs, often lead to robotic zed formation of data. Since these functions routinely occur in large volumes of members, they acquire large data streams. Essentially, media communications and interpersonal organizations often include a lot of picture, sound, video, and text data streams. Mining data streams have now become important tests for the examination local area as yet developed data mining calculations are more suitable for scheduled data. Since the actual data is changing or moving after some time and is only accessible for an exceptionally brief period of time, obtaining from the data stream is another major challenge. The evanescent properties (constant data age) of the data streams distinguish it from non-streaming data and subsequently determine the need for novel characterization, bunching, design mining processes and performance evaluation limitations. This paper centers on the issue of order in data streams. Old-fashioned system calculations are not feasible with streaming data due to asset limitations (preparation time, memory) and a single sweep of data (a glance, no irregular usage). Accordingly, to explore new learning approaches stabilized or to find out how to streamline streaming data is another exploration issue. The arrangement of streaming data is not as straightforward as the appropriation of the dataset may change throughout the course of model preparation. This wonder is alluded to as idea drift. All streaming models require representing idea drift in the model development measure. In this way, the development of a preparation model in streaming and advancing situations is frequently extremely basic. Examination toward this path has begun acquiring consideration in the new past because of the mushrooming size of data [1].

A study by someone zeroed in on delineating data stream mining with explicit calculations and applications. In late review papers, examined directed and solo learning calculations for streaming data alongside some accessible stages to deal with streaming data. However, these researchers mainly centre on characterization and grouping calculations all in all and don't cover the particular issue with order calculations for example idea drift. This paper presents simplified arrangement of the various sorts of calculations, platforms and execution measurements for grouping in streaming data, which are normally utilized in the writing alongside novel calculations for order which manage idea drift issue [2].

Thus, this paper gives a guide to experts who expect new calculations to grow to take care of the issue drift issue in the system. This work is additionally planned to help the mining local area, which wishes to use the existing idea drift adjusted stream excavation for arrangement in different locations. Detectors make abstract comparisons across data stream conditions, with drift search techniques used to identify when the setting has changed. This paper assesses eight diverse idea drift finders (DDM, EDDM, PHT, STEPD, DoF, ADWIN, Paired Learner, and ECDD) and tests using simulated datasets affected by unexpected and progressive thought drift, drift. In addition, a 2kfactorial scheme was used to show the boundaries that have the most impact performance which is an oddity nearby. Additionally, different types of Friedman non-parametric modality tests were used to differentiate different techniques. The tests thought of accuracy, evaluation time, as the author of untrue alarms and miss discovery rates. Furthermore, the constructor used the Mahalanobis distance to determine how similar the strategy is when the most ideal location contrasts with yields. This work can also be seen, in part, as an examination study of existing drift location methods [3]. Deriving from data is a point that is presented by many research fields, for example, data mining, AI, and pattern recognition. To

ISSN: 2233-7857 IJFGCN

customize, this learning is done in a static environment, where a dataset can be used multiple times to classify, depending on the case for preparation. Similarly, another trademark is that the objective idea to be learned is definite. Some classifiers have been proposed and these days, many producer classifiers are available region-wise. Currently, however, some master frameworks need to deal with continuously streaming data. Accordingly they cannot be kept away for later investigation and must be prepared as they show. Models of applications required to deal with this condition, known as data streams, are TCP / IP traffic, GPS data, sensor organization, and customer click streams [4]. Contrasting with clamp learning, DataStream formulates limitations on memory usage, limited test time and learning, and one-time data usage. Another problem of data streams is that the objective consideration can change as expected, typically after the base solidity period, an issue known as concept drift. The concept drift issue has taken a ton of consideration in recent years, mainly because it has the opposite effect on the accuracy of classifiers learned based on previous preparation opportunities. Some models of situations where thoughts may flow are "spam, extension, or climate change location" [5]. Idea drift probably ordered the pace of progress and the explanation of change. Thinking about the speed of progress, a sudden thought drift occurs when a change beta uthoren two settings are abrupt, while a gradual thought drift describes a situation where the transition beta uthoren two settings smoothly. In relation to the interpretation of progress, aerial idea drift occurs when "a set of models has real class marks at once and various real names at another point" [6]. Then again, the virtual idea drift occurs when "target ideas continue as before while data distribution changes". In practice, however, "virtual thought drift and real thought drift often occur simultaneously" [7].

Some methods have been proposed to manage the drift of the concept, including (a) adjusting the internal construction of a classifier and (b) using the assembled classifier. In this paper, the manufacturer centers around the methods used to distinguish the phenomenon of thought flow. In view of the concept drift ID, these strategies can be used inside other classifiers to alter the interior design or number of classifiers in a group. Examples of classifiers that use internally thought drift search strategies are variations to deal with drifts (DDDs) [8] and recurrent concept shifts (RCDs) [9]. Similarly, they can be used to isolate a drift related to any classifier. Drift location techniques typically use a special classifier (moreover a 'base learner') to check its accuracy and display when a drift has occurred. The purpose of this paper is to think of some idea flow detection strategies and to break them under what circumstances the counter-structure author. In each investigation that proposes another drift detection method, various tests are performed to confirm its utility. Unfortunately, the types of drifts used in the base student, datasets, measurements, and tests vary largely, as is thought by the author to be different techniques. This makes it difficult to use which strategy in different circumstances [9]. These techniques were chosen from those who chose the largest number of references, given that there was an independently available execution or possibly by allowing an instant use of a point by point drawing of an algorithm, and for incorrect results or potential to dodge the seemingly dull execution. The analysis performed using the most referenced nearby dataset, in which simulated data-sets are affected by sudden and steady thought drift, with the ability to rate and quantify the drift. To the most formidable aspect of our knowledge, this is by far the most comprehensive test for looking at a wide range of thought drift controllers in datasets with different characteristics and using them differently [10]. Paper organized as in section 2 Literature review has been described, in section 3 comparative analyses described, Section 3 talk about research gap, & finally conclusion described in section 5.

2. Literature review

These research papers focus on various issues related to concept drift and suggest the mitigating techniques also. It has also been discussed and proved that classifier ensembles give better accuracy than single classifier.

ISSN: 2233-7857 IJFGCN

In [11] author presented a new ensemble approach named as Accuracy Updated Ensemble to handle the concept drift. In their paper they studied the AAUTHOR algorithm and used the same approach of AAUTHOR by adding the author function. In their experiments they proved that AUE is giving better classification accuracy than AAUTHOR. Massive Online Analysis tool has been used to test the classification algorithms. Various real and synthetic dataset have been used to test the algorithm e.g. Electricity market dataset, Ozone level detection dataset, LED dataset, Hyper plane etc.

In [12] author offered a new algorithm named as MPC i.e. Multi-partition Multi-chunk ensemble. They have studied about the single partition single chunk classification algorithm and theoretically proved that MPC is giving better results in terms of accuracy. Besides that, they have proved that their approach reduces the classification error.

In [13] author evaluated the overall performance J48 and Multilayer Perception algorithms. Accuracy, TP rate, FP rate, Precision, recall, F-measure and ROC area have been taken as resulting measures. Various dataset have been taken from UCI machine learning repository. They have measured the accuracy with each dataset. MLP algorithm is giving better accuracy in all the dataset with big difference. As a result, it is concluded that Multilayer Perception is better approach.

In [14] author compared the overall performance of the different classifier methods like Bagging, Digging, Decorate, Multiclass Classifier and Multi boost AB. For the testing purposes Robot Navigation datasets have been used. Comparison of different algorithms for accuracy shows that Bagging is giving better result than others.

In [15] author studied the hassles of data streams with concept drift while mining. They have studied the problems of online learning and proposed an effective approach for the same. In their approach they have concentrated on data summarization, classification and learning. To predict the results, existing algorithms use one-class classifier ensemble but they have adopted new approach for prediction that is called as author one-class classifier ensemble.

In [16] author furnished a survey of diverse data mining techniques. Those techniques include affiliation, correlation, clustering and neural community. The author additionally performed a formal overview of the software of data mining inclusive of the education area, advertising and marketing, fraud detection, production and telecommunication. The author discussed the topic based on previous research studies and additionally studied the data mining strategies.

In [17] author assessed the Accuracy Author Ensemble (AAUTHOR), Multi-partition multi-chunk Ensemble (MPC) and Accuracy Updated Ensemble (AUE) and proposed new algorithms. The new techniques are Adaptive Multi-Partition Ensemble (AMPE) and AMPE2. It is mentioned that AAUTHOR algorithm is chunk based algorithm and reduces the classification error. Wang et al. proposed an algorithm to prove how AAUTHOR is better than other existing algorithms? Further they have studied MPC algorithm which is developed by Maud et al. In their research, Maud et al. proved that MPC can reduce the error rate by 10% in comparison to AAUTHOR. In fact MPC uses the different approach to design the ensemble.

In [18] author concluded with a comparative assessment of Naïve Bayes, Logistic Regression, J48 and Random forest in the context of Pima Indian Diabetes Dataset (taken from UCI repository). In their study they have tried to identify the algorithm which predict about the diabetic patient accurately.

In [19] author offered an evaluation of results of students of UG degree final year students using data mining approach in Tamilnadu state of India. The primary aim of the author is to use the classification techniques to the prediction of the overall performance of students in end semester exam. For the classification purposes,

ISSN: 2233-7857 IJFGCN

they have used C4.5, Bayesian Classifier, K-Nearest Neighbour, One R and JRip algorithms. Author has used the primary dataset for the experimentation in which they have collected the data of three different colleges of Tamilnadu. Nineteen different variables have been used in the dataset. In their experimentation work they found that Naive-Bayes and Bayes Net classifiers classify the more accurately than others.

In [20] author provided superior animal state statistics set via amassing statistics from A to Z Vertebrate's Animal Kingdom Repository. Statistics set includes 51 instances with 6 attributes. The considered attributes are name, size, author, origin, lifespan, and group. Naïve Bayes, SMO, IBK and J48 classifiers algorithm of AUTHORKA software have been used for the classification purposes. Accuracy, MAE and RMSE have been taken as a performance measurement criteria. Accuracy table shows that SMO, IBK and J48 give the same accuracy of 70.9677 % but their MAE is 0.2645,0.1451 and 0.1813 in training set respectively while 0.272, 0.1635, 0.1269 in testing set respectively.

In [21] author has made the classification of United States Congressional Voting Records dataset. Naïve Bayes algorithm of AUTHORKA tool has been used for the classification purposes. Kappa statistics, TP rate, FP rate, Precision and F-measure have been taken into consideration during the classification process. MAE and RMSE have also been calculated.

In [22] author analyzed numerous decision tree classification algorithms using AUTHORKA tool. Dataset has been used for the experimental work. An outcome indicates that decision stump algorithm takes minimal time to classify data but it gives less accuracy. J48 have quite exact accuracy with a little bit increase in time to build the classification model. Maximum accuracy given via LMT, however time taken to build classification model is far better than other classifiers or author can say maximum in all of the classifiers in maximum of cases.

In [23] author studied the various classification algorithms using haematological data of Ibn Sina diagnostic and imaging centre in which they have taken 426 instances. Basically the study of experimentation work is divided in two parts as: Experimentation with all the attribute and experimentation with selected attribute. They have implemented J48, MLP and Nave Bayes algorithm with the said dataset. In both the experimentation they found that J48 and Naïve Bayes classifiers are taking minimum time to build the model in comparison to MLP. As far as accuracy is concerned, J48 classified the 97.16% data correctly which is best accuracy among all three.

In [24] author studied the C4.5, Random Forest, Naïve Bayes, MLP and Lib SVM classifiers with educational dataset. In their study they found that Lib SVM classified the 80.85% data correctly which is best among all five algorithms.

In [25] author tested nine different AUTHORKA classification algorithms to classify the college students' dataset. The KAU-ODUS+ dataset have been used for experimentation work. In their study, the data of Medical College, Engineering College, Computer College and other have been taken into the considerations i.e. the nine classification algorithms have been implemented with four colleges. Recall, Precision, F-Measure, Matthews's correlation coefficient (MCC), Precision, ROC, FP rate, and TP rate authored computed through simulation with AUTHORKA Toolkit as a performance measures. Mean Absolute Error has also been computed.

3. Comparative analysis

Ref. No	Year	Methodology	Dataset	Parameters
26	2019	The point of this	Static,	v and h, User

ISSN: 2233-7857 IJFGCN

		exploration is to depict key difficulties looked by algorithmic answers for stream mining, especially zeroing in on the common issue of idea drift. An exhaustive conversation of idea drift and its inborn data challenges with regards to stream mining is introduced, similar to a basic, inside and out audit of significant writing.	Imbalanced, benchmark, concept drift,	specified parameters,
27	2011	In this paper, creator proposes another strategy for recognizing idea drift. The proposed technique, which can identify various kinds of drift, depends on preparing data lump by piece and estimating contrasts beta uthoren two successive clumps, as drift pointer. To assess the proposed technique creator measure its presentation on a bunch of fake datasets with various degrees of	artificial datasets,	DDM, EDDM

ISSN: 2233-7857 IJFGCN

		seriousness and speed of drift.		
28	2017	This paper centres on idea drift issue in grouping of streaming data. During arrangement an adjustment in the idea or appropriation of dataset throughout the time is named as idea drift. The exhibition of a model/classifier debases because of idea drift even in fixed data; managing this issue thus becomes more testing in data streams. This paper presents arrangement of existing streaming data order calculations alongside their capacity to tackle idea drift issue.	ITI, VFDT, CVFDT	performance evaluation parameters, Apache spark, Apache storm, MOA, SAMOA
29	2019	This paper proposes a double identification system to pass judgment on the drift of ideas, and on this premise, the combination characterization of data stream is done. The framework intermittently recognizes data stream with the	high-dimensional datasets, Artificial data,	accuracy rates, number of dimensional

ISSN: 2233-7857 IJFGCN

		list of arrangement mistake and uses the highlights of the fundamental arising design (eEP) with high segregation to help fabricate the coordinated classifiers to tackle the grouping mining issues in the powerful data stream climate.		
30	2020	Proposed strategy EFCDD manages the repetitive drift of the idea in streaming data. To express the drift, the projection variety of the qualities addressing the field positions or field-IDs, which are being used for outlining the design of the records streaming structure the intended sources? The exploratory investigation was done by ridiculing the streams of those communicating records of the benchmark datasets regularly utilized in DM. The results of the test study manifest the	Benchmark dataset, Imbalance datasets	-

ISSN: 2233-7857 IJFGCN

		versatility and prominence of EFCDD toward the recognition of drift in idea. The proposition execution is estimated by comparing recreation results with the other existing model.		

4. Research Gap

In this section, we have made an effort to explain about the data mining, machine learning, single classifier, classifier ensemble, classification algorithms, concept drift and the various problems related to concept drift. We have focused our study specifically on the classification problems and the algorithms available in WEKA tool. The research papers of various researchers and professionals have been studied [31-36]. They have made comparison of various classification algorithms using different types of dataset. Various parameters like accuracy etc. have been taken by them and it is concluded that classifier ensembles are always better than single classifier specifically in the case of concept drift. Past study about the research topic proved that data mining and machine learning have been cutting-edge research field and can be implemented in industry with the large databases. Further, it has been proved that classifier ensembles play a vital role while searching a data from huge database

5. Conclusion

In this research work, we examine the literature on performance in terms of accuracy of proposed classifier ensemble method, existing classifier ensemble algorithms and single classifier. The other performance measures are TP rate, FP rate, Precision, Recall and F-Measure. Accuracy has been measured on each data set of hyper plane Dataset and SEA dataset. In literatures found that proposed method is giving better accuracy in 6 subsets of hyper plane dataset out of its nine subsets and giving better accuracy in all five subsets of SEA. The accuracy is likewise calculated on vehicle, balance-scale and Glass datasets (as given in the base paper). On balance-scale and vehicle data sets proposed approach is giving better accuracy than the MLP, J48 and Single Classifier. On glass datasets, accuracy is sort of same and in Multilayer Perception and J48 are barely higher algorithms.

References

[1] Meraj Nabi, Abdul Wahid and Pradeep Kumar, "Performance Analysis of Classification Algorithms in Predicting Diabetes", International Journal of Advanced Research in Computer Science, vol. 8, no. 3, pp. 456-461, 2017.
[2] Driyani Rajeshinigo and J. Patricia Annie Jebamalar, "Educational Mining: A Comparative Study of Classification Algorithms Using WEKA", International Journal of Innovative Research in Computerand Communication Engineering, vol. 5, issue 3, pp. 5583-5589, 2017.
[3] Shivangi Gupta andNeeta Verma, "Comparative Analysis of classification Algorithms using WEKA tool", International Journal of Scientific & Engineering Research, vol. 7, issue 8, 2016.

ISSN: 2233-7857 IJFGCN

[4] Iknoor Singh Arora and Gurpriya Kaur Bhatia, "Comparative Analysis of Classification Algorithms on KDD"99 Data Set", Computer Network and Information Security, vol. 9, pp. 34-40,2016.

[5] Geeta Kashyap and Ekta Chauhan, "Parametric Comparisons of Classification Techniques in Data Mining Applications", International Journal of Engineering Development and Research, vol. 4, issue 2, pp. 1117-1123, 2016.

[6] Rausheen Bal and Sangeeta Sharma, "Review on Meta Classification Algorithms using WEKA", International Journal of Computer Trends and Technology (IJCTT), vol. 35, no. 1, pp. 38-47, 2016

[7] N. Ukwueze Frederick and C. Okezie Christiana, "Evaluation of Data Mining Classification Algorithms for Predicting Students Performance in Technical Trades", International Journal Of Engineering And Computer Science, vol. 5, issue 8, pp. 17593-17601, 2016

[8] Indre Zliobait, Mykola Pechenizkiy and Joao Gama, "Big Data Analysis: New Algorithms for a New Society", Chapter 4, "An overview of concept drift applications", Springer International Publishing Switzerland, pp. 91-114, 2016

[9] Barros RS, Cabral DR, Gonçalves PM Jr, Santos SG (2017) RDDM: Reactive drift detection method. Expert Syst Appl 90:344–355

[10] Bifet A (2017) Classifier concept drift detection and the illusion of progress. In: International conference on artificial intelligence and soft computing. Springer, pp 715–725

[11] de Lima Cabral DR, de Barros RSM (2018) Concept drift detection based on Fisher's exact test. Inf Sci 442:220–234

[12] de Mello RF, Vaz Y, Grossi CH, Bifet A (2019) On learning guarantees to unsupervised concept drift detection on data streams. Expert Syst Appl 117:90–102

[13] Dua D, Graff C (2017) UCI machine learning repository. http://archive.ics.uci.edu/ml. Accessed 14 Oct 2019

[14] Duarte J, Gama J, Bifet A (2016) Adaptive model rules from high-speed data streams. ACM Trans Knowl Discov Data (TKDD) 10(3):30

[15] Duong QH, Ramampiaro H, Nørvåg K (2018) Applying temporal dependence to detect changes in streaming data. Appl Intell 48:4805–4823

[16] Haque A, Khan L, Baron M (2016) Sand: semi-supervised adaptive novel class detection and classification over data stream. In: AAAI, pp 1652–1658

[17] Haque A, Khan L, Baron M, Thuraisingham B, Aggarwal C (2016) Efficient handling of concept drift and concept evolution over stream data. In: 2016 IEEE 32nd international conference on data engineering (ICDE). IEEE, pp 481–492

[18] Jaworski M, Duda P, Rutkowski L (2017) New splitting criteria for decision trees in stationary data streams. IEEE Trans Neural Netw Learn Syst 29(6):2516–2529

[19] Krawczyk B (2016) Learning from imbalanced data: open challenges and future directions. Prog Artif Intell 5(4):221–232

[20] Krawczyk B, Minku LL, Gama J, Stefanowski J, Woźniak M (2017) Ensemble learning for data stream analysis: a survey. Inf Fusion 37:132–156

[21] Liao J, Zhang J, Ng WW (2016) Effects of different base classifiers to learn++ family algorithms for concept drifting and imbalanced pattern classification problems. In: 2016 International conference on machine learning and cybernetics (ICMLC), vol 1. IEEE, pp 99–104

[22] Pesaranghader A, Viktor HL, Paquet E (2018) Mcdiarmid drift detection methods for evolving data streams. In: 2018 International joint conference on neural networks (IJCNN). IEEE, pp 1–9

ISSN: 2233-7857 IJFGCN

[23] Reinsel D, Gantz J, Rydning J (2017) Data age 2025: the evolution of data to life-critical don't focus on big data; focus on the data that's big. Technical reports on IDC. https://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf. Accessed 14 Oct 2019

[24] Sethi TS, Kantardzic M (2017) On the reliable detection of concept drift from streaming unlabeled data. Expert Syst Appl 82:77–99

[25] Webb GI, Hyde R, Cao H, Nguyen HL, Petitjean F (2016) Characterizing concept drift. Data Min Knowl Discov 30(4):964–994

[26] Wares, S., Isaacs, J. & Elyan, E. Data stream mining: methods and challenges for handling concept drift. SN Appl. Sci. 1, 1412 (2019). https://doi.org/10.1007/s42452-019-1433-0

[27] Sobhani, Parinaz and Beigy, Hamid","New Drift Detection Method for Data Streams", Adaptive and Intelligent Systems, 2011, Springer Berlin Heidelberg, pages="88--97"

[28] Janardan, Shikha Mehta, Concept drift in Streaming Data Classification: Algorithms, Platforms and Issues, Procedia Computer Science, Volume 122, 2017, Pages 804-811,

[29] Zhang, B., Chen, Y. Research on detection and integration classification based on concept drift of data stream. J Wireless Com Network 2019, 86 (2019). https://doi.org/10.1186/s13638-019-1408-2

[30] Kalli Srinivasa Nageswara Prasad, Annaluri Sreenivasa Rao & Attili Venkata Ramana (2020) Ensemble framework for concept-drift detection in multidimensional streaming data, International Journal of Computers and Applications, DOI: 10.1080/1206212X.2020.1711617.

[31] Mahajan, H.B., Badarla, A. & Junnarkar, A.A. (2020). CL-IoT: cross-layer Internet of Things protocol for intelligent manufacturing of smart farming. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-020-02502-0.

[32] Mahajan, H.B., & Badarla, A. (2018). Application of Internet of Things for Smart Precision Farming: Solutions and Challenges. International Journal of Advanced Science and Technology, Vol. Dec. 2018, PP. 37-45.

[33] Mahajan, H.B., & Badarla, A. (2019). Experimental Analysis of Recent Clustering Algorithms for Wireless Sensor Network: Application of IoT based Smart Precision Farming. Jour of Adv Research in Dynamical & Control Systems, Vol. 11, No. 9. 10.5373/JARDCS/V11I9/20193162.

[34] Mahajan, H.B., & Badarla, A. (2020). Detecting HTTP Vulnerabilities in IoT-based Precision Farming Connected with Cloud Environment using Artificial Intelligence. International Journal of Advanced Science and Technology, Vol. 29, No. 3, pp. 214 - 226.

[35] Mikhail, A., Kamil, I. A., & Mahajan, H. (2017). Increasing SCADA System Availability by Fault Tolerance Techniques. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA). doi:10.1109/iccubea.2017.8463911

[36] Mikhail, A., Kareem, H. H., & Mahajan, H. (2017). Fault Tolerance to Balance for Messaging Layers in Communication Society. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA). doi:10.1109/iccubea.2017.8463871

ISSN: 2233-7857 IJFGCN