ISSN 2278-3091
Volume 9 No.2, March -April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse111922020.pdf
https://doi.org/10.30534/ijatcse/2020/111922020

Aspect Oriented Concept Drift Detection in High Dimensional Data Streams

M. Sankara Prasanna Kumar[1], A. P. Siva Kumar[2], K. Prasanna[3]
[1]Research Scholar, JNTUH-Hyderabad, Assistant Professor, AITS, Rajampet, Andhra Pradesh, India, sankaraprasannakumar@gmail.com
[2]Assistant Professor, Department of Computer Science and Engineering, JNTUCE, Anantapuramu, Andhra Pradesh, India, sivakumar.ap@gmail.com
[3]Associate Professor, Department of Information Technology, AITS, Rajampet, Andhra Pradesh, India, prasanna.k642@gmail.com

ABSTRACT

The drift of the concept is the critical goal of data mining over data transmission, which often deotes the diversity between the pair of sequentially transmitted data tuples. The drift of the concept can be incremental, which increases gradually in the face of data transmission. The other dimension of the drift concept is sudden drift that manifests itself considerably between the pair of tuples of transactions transmitted in sequence. Contemporary Concept drift identification approaches are primarily intended to deal with incremental or sudden drift. In this document, aspect-oriented concept drift detection (AOCDD) is projected into high-dimension data streams. To report concept drift, the AOCDD represents the diversity of data projection for the aspects that are used to frame the record structure in the target data streams. The experiments carried out the reference data sets as flows, which show the importance and scalability of the AOCDD for the detection of drift. The performance advantage of the proposal is scaled by comparing the experimental results with another contemporary model in recent literature.

Key words: Drift detection methods, Concept-drift, Aspect Pattern Weights, online imbalance datasets.

1. INTRODUCTION

Operating with complete data sets in a static environment allows store data sets to be durable and the process can be continuously accessed. Furthermore, the theory behind the data remains unchanged. However, with the advancement of technology in the recent past, a large possibility of applications made, researchers will focus on dynamic environments, such as banking, telecommunications, genomics, e-commerce and the social media platforms they generate. data streams in real time and on a continuous basis. The data generated from these types of uses continuously form the flows at a higher frequency. Maintaining such a large amount of data through storage and post-processing will be impossible due to volume involvement. In specific defined applications such as "social networks" and "quick responses to the banking mandate".

Consequently, data streams require novel algorithms from such applications through limited storage usage, one-time data scanning, and real-time response. In addition to these additional demands, the application also challenges the change of the whole theory on which the algorithm is based. The term derived, the theory behind it, is the collection of data changes after defined stability stages during the period.

The inability of the algorithms to include the alterations in the results of the data transmission to the method with less precision and total effectiveness [1]. And since machine learning is based on cases that are constant, functioning in the dynamic atmosphere requires a timely identification of the concept of drift [2]. Some of the examples of the concept of drift can be seen in the detection of scams and garbage and in the estimation of the climate.

The concept of drift is divided into 2 subtypes, according to the rate of change: the concept of fast or fast drift and the concept of gradual drift. The word sudden drift reflects a sudden change in data collection theory, while gradual drift occurs over a period of time. To ensure an accurate classification, it is important to classify the established drift in 1 of the 2 classifications, as well as to reduce false negatives.

Through the importance of categorizing the concept of drift, multiple methods are suggested for the detection of concept drift. The 3 important approaches in this classification are drift detection methods (DDM), inline algorithms, and ensembles. In between these three, the ensemble method has become more popular in expressions of precise categorization. And the set method joins the votes for each intended class tag.

This document focuses on improving the efficiency of the data flow drift detection concept. In contrast to previous approaches that focused on sudden or incremental detection of deviations, the suggestion in this article reached the importance of detecting sudden and incremental deviations in the model.

Identifying the concept of sudden and incremental drift, this article suggested a new concept of drift detection using "Aspect Pattern Weights (APW)" that is targeted as an unreliable mining procedure and capable of determining sudden and incremental shunts in the model.

Section 2 describe the literature work related to the proposed method. In Section 3, procedure and its data structure presented. Section 4 describe the experimental study and result analysis, followed by conclusion in Section 5.

2. RELATED WORK

The work [3] shows that real-time-based applications are spam identification, preference prediction, cooperative filtering, etc. they need a fundamental alteration in the data flows and their intrinsic associations. The current solutions to the problems to be addressed in the concept of drift will be broadly classified into 2 categories [4], [5]. In the primary phase, the classifier is programmed by the researchers to automatically increase the parameters of the adaptation method according to the new data. And the second phase of the method incorporates an additional drift detector in the statistical method. The statistical method and the detector ensure that the drift time is recorded. When compared to the primary phase, the methods in the second phase performed 2 operations: mitigate and detect the concept of drift and record the time of occurrence of the drift.

Emphasizing the identification of the data flow leads, several research techniques are suggested. These assists are grouped into 3 types: concept drift identifiers, window segment method, and assembly item classifiers. To maintain accurate identification rates, window methods retrieve existing data and advanced classifiers. However, the main restriction of these techniques is that the size of the window limits the discovery process.

Set classifiers, on the other hand, adjust grouping models and classifier organizations according to decreased production due to the drift principle. Drift detectors constantly track the classifier output and use trigger points to increase warnings of possible shunts [6] and classifiers.

The sliding window model is a significant result in the classification that limits the volume of cases from recent data. This method together with the traditional learning program produces novel classifiers for the data flow. However, detecting the optimal sixe window is a difficult task. Windows that are small can identify sudden alterations, but huge windows fail to identify these types of rapid alterations. Small windows may also have less precision at constant times, but large windows had comparatively higher precision rates at constant times [7].

The work [3] suggested detection models for DDM. The model measures the total error of the classification that is generated from the method, as well as its SD-standard deviation. Since this model measures the aggregate of false negatives and positives, it is ineffective in detecting individual deviations until the aggregate differs. And this problem is highlighted even more in the data that is unbalanced.

Proposal [8] suggested the DDM variant. The old DDM model maintains slow and constant variations when measuring the distance between 2 errors. However, the model waits until it gets at least 30 errors, to measure and incorrect data for the implementation of unbalanced activities.

The work [9] investigated another model called a model to calculate the loss. The STEPD model relates the precision of the existing window to the added precision minus the precision of the existing window. Take a similar ratio test to compare levels of Accuracy.

The The work [10] shows that the DDM are lengthened for use in the OCI (online imbalance data sets). Apprehensions regarding DDMs are addressed with data that is unbalanced; it depends on the platform of the alterations in the true positives (TP). On the other hand, in cases where a drift occurs without presenting abnormalities in the TP, the method cannot detect the shunts. And, for example, without copying the alterations in the "PPV, F-score and TPR", the derivations are likely from the data that is unbalanced to balanced data. The DDM-OCI model will not be successful in detecting this concept of drift.

Another limit of this model will be that it will lead to "FP (false positives)" on a common basis. And the statistical task will spread inappropriately in a stable atmosphere. Consequently, the confidence levels of TPR in [3] have become inappropriate for null dissemination.

To overwhelm the confines of the DDM-OCI model, the "Four linear velocities" model was suggested [11]. And look at 4 confusing matrix contrast rates with the above method. However, the assessment of the method represented that FP scare was still being triggered.

Set classifiers for data streams quickly emerged as a well-known method of enabling machine learning. In contemporary work, there are 2 subtypes of set classifiers: block and online bases. Set in real time or online after the instance of each data and advance their classifiers. The functional set block detects the pieces of data before advancing the classifiers. Therefore, the line-based set will be more suitable for abrupt deviations, and the block-based model is appropriate for incremental deviations.

The work [7] proposes the "Aggregate set methods (AUE2)", based on the weight of drift classification. The model predicts the class mark and tracks the precision of small blocks to detect slow fluctuating drift. And because the approach involves large class tagging, the grouping approach is optimized for fewer instances.

The work [12] [29] demonstrates a community of groups and classifiers used to classify the data and thus reduce the burden. Effort [13] shows that core density is used to detect patterns and clusters not seen. Spatial path offsets are also derived from 2D data.

No matter what specific attempts are made to detect deviations, simplification in the midst of evolution and existing perceptions are not achieved to satisfactory levels of precision. The paper [14] addresses the problem by suggesting "level wise sets". For each level of diversity, the set is assigned, and whenever a new concept needs to be simplified, sets at various levels are linked.

For programmers, re-detection remains an important task. Novelty is identified when the novel feature is nurtured by the current sequence of data. The papers [15-20] present that many frameworks are proposed to address novelty precision for detection.

The work [21] suggested the other existing method called "EXStream" in recent works. The EXStream method is the combination of 3 methods used in the concept of drift discovery. In the middle of three, one is, the detection of the concept of drift through the "classification imbalance"; and 2 others are identifying the drift in the attribute scenario and the values are shown directed to these attributes accordingly. The first method that represents drift through the imbalance class is limited to data flow categorization methods. Two other methods together look at the extent of drift despite the directed procedure that is applied to the data. And these 2 methods deliberate the data stored in the specified time and report the attributes that are in contradiction to the base attributes by deliberating the represented values for each resulting attribute as the vector. Identically, the final method in the middle of 3 detects the different values that are displayed in the recordset that is stored in the buffer and the recordset that is used to denote the concept. In the concern of identifying differences between values and attributes, EXStrean is using the model called "IME (Interaction Based Method for Explanation) [22]" that it planned to allow for single estimates. The method limits are a weak implementation to identify the incremental impact of the drift and the identification of the drift is limited to the categorization procedure on the data flows.

The work [23] projected a conceptual drift detection method that verifies the magnitude of the change between explanations of multiple models. The explanations of the model are calculated through the values projected to the attributes, which is a similar context of the model proposed for this manuscript. However, the method is not considering the format of the identified drift, in addition to computing explanations of multiple models that seek prior knowledge of the projected values, which is often a critical constraint in detecting concept drift.

The paper [24] presents another important paper in the more recent literature: "identification of density variation based on the nearest neighbor (NN-DVI)" for the principle of drift detection within data streams. This work used the "closest neighbor method" to detect the difference in data density in the specified area. The model that splits the specified data into the regions and tries to look at the drift at the regional level that further uses these leads at the region level to finalize the drift state of all the data. However, the method limited its execution to "high-dimensional data" that is delaying drift detection due to overloading the procedure, further resulting in drift detection inaccuracy.

In this document, "Identifying the Unit of Concept for Pattern Diversity Measurements (AOCDD)" was intended to overcome the limitations observed in current approaches. The recommended approach is independent of the mining method applied to the target data stream and shows the ability to accurately identify drift length. Contrary to the current SPC process ExtreamModel [23], which focuses on regional deviations, the suggested AOCDD process shows the status of the vertical and horizontal distribution similarities in the transmission buffers that were converted in the window, compared to the source records. in the SPC ExtreamModel [23].

3. METHODS AND MATTER

The proposed method consists of 2 stages, one of which is to anticipate the coincidence of the aspect and the second stage is to anticipate the coincidence of the drift of the idea to reflect the inclemency or the abrupt "current window" of the objective data flow in the log level. To verify the appearance level similarity performance, the suggested method evaluates the assisted values for the weights displayed for each aspect to represent the level of appearance distribution similarity, and to evaluate the "log level similarity", the Visualized method assesses "pattern weights" represented by projected aspect values that showed aspect level similarity. Later, the suggested method predicts the extent of "conceptual drift" as it is abrupt, not drift, or incremental. The summary of this method is represented in Table 1. The data structure used and the identification of the suggested drift method are discussed in the following segments.

Table 1: The Summary of the suggested detection of the drift Concept

- Identify the similarity of the aspect -level // evaluate for all aspects the sum of the resemblance for destination to source and source to destination, If the similarity reaches 0.7, there will be no drift associated with the resulting dimension at that point.
- When the aspect level similarity is greater or equal to the defined threshold similarity.
  - Identify the “similarity aspect -pattern” // for every distinctive pattern in decreasing sequence of length of pattern (minimum length of the pattern is 0.7*count of aspects identified with the similarity is higher than the 0.7), if at this point the aspect ratio is greater than enough 0.7, perform the "similarities pattern level test," applying the "similarity pattern" from destination-source and source-destination when the similarity is reaches 0.7, exit check at that stage of "pattern level similarity."
  - If the "similarity aspect model" is either higher or equal to the threshold defined.
    - In theory there is no drift.
    - Identify whether the drift is incremental.
  - Else, approve the drift is happening incrementally in concept // “aspect level similarity” is maximum, “aspect pattern level similarity” is minimum.
  - Elseif “aspect level similarity” is approximately equal to “similarity-threshold”, then Label the drift as Incremental.
- If not, label the drift as sudden

3.1 Structure of the Data

The set of notation buffered windows $BW = \{bw_1, bw_2, \ldots bw_{|BW|}\}$ so that the standardized records are stored in each buffered window. Let the note $v_j = \{sv_1, sv_2, \ldots sv_{|asp|}\}$ be the record of the buffered window bw_i will be outlined through values represented for aspects $asp = \{s_1, s_2, \ldots s_{|asp|}\}$ fixed arrangement of records. Every buffer window $\{bw_i \exists\ bw_i \in BW\}$ that the BW buffered windows set reflects the variation in design related to other buffered BW set windows.

The proposed drift detection technique is 2 phases. The initial step is to define "the similarity of aspect stage," and the next phase is to define "similarity of pattern stage."

3.2 Aspect Level Distribution Diversity

The purpose of this recommended test on the match aspect is to show that the distribution values of similarity are performed on the records of the current window for every aspect and tuples are registered for account. The suggested approach tests the weights of each entry displayed in the current window, which is “occurrence-ratio” in buffered window. And it is vice versa, evaluates the weight of every entry depicted to be value of resultant- aspect in buffered window which is “occurrence-ratio” the record set that is buffered. The aggregate weights displayed as values for the corresponding aspects for entire entries showed the weights of the respective buffered & buffered frames. And an average of 2 weights of the same dimension would later appear as "similarity distribution weight." The proportion of aspects with the distribution weight similarity that is greater than the indicated threshold will later be reflected as the "patterned window level aspect weight match weight." That is more than the threshold stated so that the model level distribution weight continues to be predicted. The entire process is discussed as follows:

For each cw Begin //for each current window, which is the most recent window that buffered set of records from data stream

$\forall_{k=1}^{|asp|} \{s_k \exists s_k \in asp\}$ //For each aspect s_k,

$\forall_{j=i+1}^{|BW|} \{bw_j \exists bw_j \in BW\}$ Begin //for each buffered window bw_j

Let the notation v_j^k refers the vector of the values observed for aspect s_k in all records of the buffered window bw_j.

Let the notation v_{cw}^k refers the vector of the values observed for aspects s_k in all records of the current window cw.

For each unique entry e_{cw}^k of the vector v_{cw}^k Begin

Find the occurrence ratio $ro\{e_{cw}^k \to v_j^k\}$ of the e_{cw}^k in vector v_j^k and add $ro\{e_{cw}^k \to v_j^k\}$ to the set ro_{cw}^k

End

For each e_j^k unique entry from vector v_j^k Begin

Find the occurrence ratio $ro\{e_j^k \to v_{cw}^k\}$ of the e_j^k in vector v_{cw}^k and add $ro\{e_j^k \to v_{cw}^k\}$ to the set ro_j^k ro_j^k

End

Find the mean $\langle ro_{cw}^k \rangle$ of the ro_{cw}^k and the mean $\langle ro_j^k \rangle$ of the set ro_j^k

If the mean $\langle ro_{cw}^k \rangle$ is greater than the given threshold τ, conclude the drift in concept between current window cw and buffered window bw_j under aspect s_k

End

End

End

- After aforesaid procedure which applied to entire aspects which are involved, identify aspects ratio which are not showing the distribution diversity.
- If ratio will be higher than specified threshold $a\tau$, confirm that there will be no drift displayed at the aspects level distribution and at that point enhancing the procedure to represent the drift through an “aspect pattern support weights” that displays the distribution diversity at record level.

3.3 Diversity of Distribution record level

In this article this prototype contains the 2nd stage of the whole drift detection process. That is important because, however, the drift concept is not shown in drift stages, and also fails to reminder the values of the pattern similarities shown for the ensuing aspects of the window that is buffered and the buffers that may result in an incremental "drift concept." Therefore, all "data level distribution similitudes" that detect the range of drift concept must be recorded, where they are seen, it is understood to be a drift concept suddenly or incrementally.

The correlation with the distribution at record level is measured using the suggested approach as follows: Deliberate the $asp = \{s_1, s_2, \dots s_{|asp|}\}$ set showing the chosen aspects from the first stage that reflect the "similarity of the distribution level dimension" method. The research [25] shows the set theory, the possible number of single subsets from the specified set is $2^{|asp|} - 1$. Consequently it is shown that the probable sum of the patterns is from the set of $asp$are $2^{|asp|} - (|asp| + 1)$ which consists of 2 or more components of each pattern and hence we are eradicating the size of sub-sets to be 1 that is of $|asp|$ number.

In order to compare patterns of the same size as set, these patterns are then divided into classes. And the maximum number of possible sets for this concern is $|asp| - 1$, the term $|asp|$ refers to the total number of aspects used every frame document. In comparison, the increasing number of patterns in each set is $\frac{|asp|!}{l!*(|asp|-l)!}$ the term $|asp|$ here indicates the number of aspects used to frame every record, and the word l indicates the equivalent pattern length. When the variable $|asp|$ matches a number, the corresponding variable is $(|asp| - 1)!$ It is the 1 which is the root of number convention[26].

This helps to shape probable patterns and groups of the same duration when the procedure is completed.

step 1. Let every pattern in series with maximum length $ps_{|asp|-i}$ be at least long 1,

step 2. $\forall^{i=1}_{(|asp|-1)} \forall^{j=1}_{|ps_{|asp|-i}|} \{p_j \exists p_j \in ps_{|asp|-i}\}$ Begin //p_j is the pattern of length $|asp| - i$

a. Project v^j_{bw} set, // that includes entire likely values of p_j in the ensuing buffered window bw record collection

b. Similarly, v^j_{cw} projects the range, // which includes entire likely values of p_j the respective patterns in the current window cw .

step 3. Assess the distribution diversity between v^j_{cw} & v^j_{bw} sets shown. The procedure for determining the similarity of distribution between these two sets would be very similar to the procedure changed to determine the "similarity of distribution of the element"

step 4. Find "occurrence ratio (support weights)" of every entry e which is unique of set v^j_{bw} in reverence to the cw and total the equivalent "support-weights" to be asw^j_{bw}

1. Let the term tr denotes empty-set
2. $\forall^{m=1}_{|v^j_{bw}|} \{e \exists e \in v^j_{bw} \wedge e \notin tr\}$ Begin
3. $sw(e) = \frac{\sum_{a=1}^{|cw|}\{1 \exists e \subseteq \{r_w \exists r_w \in cw\}\}}{|cw|}$
4. $asw^j_{bw} = asw^j_{bw} + sw(e)$
5. End // of step 3
6. End //of step 2

step 5. Identically find "support (occurrence ratio) weights" of every entry e which is unique of set v^j_{cw} in respect to the bw , and total amount of the corresponding "support-weights" to be asw^j_{cw}

step 6. Finding ds "mean of the total support weights", asw^j_{cw} , asw^j_{bw} is more than specified $s\tau$ similarity threshold (0.7 /above),

$$ds = \frac{asw^j_{bw} + asw^j_{cw}}{2}$$

step 7. $if\ (ds \geq s\tau)$, at that point the current window cw will not be displayed at a drift in the case of buffered window bw, then finish the procedure then, if not, start the pattern set level procedure that matches the pattern set by size sequences.

step 8. At the conclusion of the test

- If the detected drift is more or less equal to the size of the patterns only, then the drift has been accepted to be gradual drift and, if not, a sudden drift.
- When the gradual drift is detected, merge the current buffered window with the next set of buffered rates and apply the same procedure.
- If the process represents coefficients of drifts beyond defined thresholds, no drift is reported in the records that buffered set as opposed to the' buffered window.'
- If in the former stage of drift detection, the coefficients of drift are smaller than those of the drift note, then authorize the drift.
- If drift coefficients exceed the coefficients noted in the former drift detection process and are below the defined thresholds, the drift is accepted incremental.

This process continues until the current window drift or similarity is accepted.

4. EXPERIMENTAL STUDY AND PERFORMANCE

Throughout this section, the experimental information produced for the model are clarified, such as target dataset, approach evaluation, use of success indicators and explanations made from the gained results. However, the AOCDD findings were compared with the SPC_ExtreamModel findings [23], which are obtained from the experiment performed on the same target dataset, in order to analyze the proposed model value for the detection of the concept-drift.

4.1 The Dataset

The effort [27] contributes to the KDD-Cup'99 as one of the most significant datasets adapted by the models listed in recent literature for the mining processing. The effort [28] introduces for the first time the data set launch and the use of the data set by Stolfo et al. Research [30] contributes to the DARPA'98 network analysis initiative, which has accumulated almost 4 GB traffic of the network. And this formatted binary data is used to determine the model efficiency of Intrusion Detection Systems. The new Dataset for Network Traffic holds the records about five million records with 41 aspects in each database. Therefore, partial analysis of a data collection obtained are used to perform experiments as this array of data sets holds numerous and manifold records that are vulnerable to numerous malevolent crimes and demonstrate concept-driving.

4.2 The Experimental Setup

For the experiment, the dataset is initially divided into different tuples, based on the values covered by any aspect used for the production of every record. Based on the contribution [31] the data-set-partitioning method is carried out by clustering x-means, where cluster numbers (input) are essential for the identification of drift with different spanned values between the aspects of the records. Furthermore, these clusters are used to determine the drift detection accuracy of both AOCDD and SPC Extream Model.

After the data is divided into tuples, the process is carried out to stream the target date when the records are buffered within the specified length. Then take a twice as many records as a new window cw. If it is not tuple first, the drift detection mechanism is implemented in the current window, and all buffered windows. $\{bw_i \exists\ 1 \le i \le S\}$ where the notation Srefers max tuples present as buffered window in case of buffered window existence. Therefore, when the current cw does not show any concept-drift in any buffered window, the current cw is collective with a buffered window bw_i that has similarities. Once the data streaming process is done, the clustering process displays the number of clusters that form as buffered windows.

Take the buffered windows projected as $BW = \{bw_1, bw_2, bw_3, \ldots bw_{|BW|}\}$ and mark the projected clusters as C clustering method, $C = \{cl_1, cl_2, cl_3, \ldots cl_{|C|}\}$

$dda = BW \cap C$ // The dda notation is now the exactness of the drift detection of both clusters and the buffered frame.

Consider the "drift detection accuracy ratio" is *ddar* which is *dda* over the buffered window *BW*

$$ddar = \frac{|dda|}{|BW|}$$

The failure rate of detection of drift is also taken in account and compared with later.

$ddf = BW \ / \ dda$ // ddf means "drift detection failure" of buffered window *BW* that present is the no. of tuples but not in dda

Take $ddfr$ as drift detection failure ratio for the *ddf* ratio over *BW* and expressed as:

$$ddfr = \frac{|ddf|}{|BW|}$$

Table 2, Table 3 reflects the metric values for comparative analyzes performed on both the AOCDD and SPC ExtreamModel on the same distribution of data sources and those tests are shown in pictographic format in Figure 1, Figure 2.

As the source dataset is divided into five separate sets, the experiment was carried out in five different sources, each of which was transferred between the five separate source data sets.

Table 2: AOCDD and SPC_Extream Model-Drift detection rates over diverse data sets

Drift Detection Rate		
Fold ID	**AOCDD**	**SPC_ExtreamModel**
1	0.983874	0.935488
2	0.982147	0.83929
3	0.95126	0.756102
4	0.977277	0.840913
5	0.963640	0.909095

Table 3: AOCDD and SPC_Extream Model Failure rates of drift detection over different data sets

Failure Rate of Drift Detection		
Fold ID	**AOCDD**	**SPC_ExtreamModel**
1	0.016133	0.064520
2	0.017861	0.160718
3	0.048712	0.243906
4	0.022731	0.159095
5	0.036368	0.090913

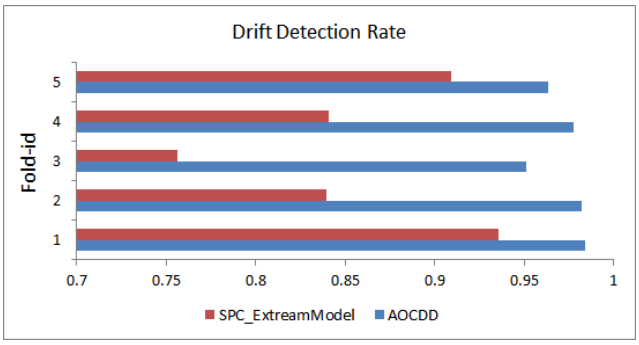

Figure 1: Drift detection accuracy ration of AOCDD and SPC_ExtreamModel

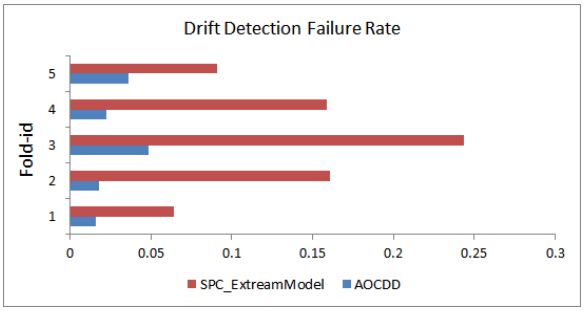

Figure 2: Failure drift detection ration of AOCDD and SPC_ExtreamModel

The experimental results show that in each case, the proposed AOCDD model works better than the SPC ExtreamModel.

5. CONCLUSION

In this paper the technique of detecting the "concept-drift" in data streams is suggested. Accordingly, the proposed model, AOCDD used a measurement scale for assessing the data distribution similarity called aspect pattern weights. Contrasting with the modern models, the model AOCDD processes independently in various practices of data stream mining. This also shows significant output in detecting the concept-drift of both gradual and sudden drifts, but contemporary models show significance in detecting drift in gradual drift or sudden changes. Using two-fold assessment i.e., distribution similarity at aspect level assessment and distribution similarity at aspect pattern level assessment, the model is discussed in the manuscript. On the AOCDD, and SPC_ExtreamModel [23], the experiment is performed with the use of KDD cup dataset. This data set is made of 5 unique sets and each set is utilized to conduct the five folds of the experiments. When compared to the results obtained in the experimental study of both models, the proposed model, AOCDD evinced significance performance for detecting the concept-drift. The contributions of distribution similarity assessment metrics in the manuscript can further be applied as fitness metrics for estimating the concept-drifts in the diverse environments.

REFERENCES

1. Gama J, Zliobaite I, Bifet A, Pechenizkiy M, Bouchachia A. **A survey on concept-drift adaptation**. *ACM Computing Surveys.* 2014;46(4):44-1. https://doi.org/10.1145/2523813
2. Widmer G, Kubat M. **Learning in the presence of concept-drift and hidden contexts**. *Machine learning.* 1996 Apr 1;23(1):69-101. https://doi.org/10.1007/BF00116900
3. Žliobaitė I. **Learning under concept-drift: an overview**. *arXiv preprint arXiv:*1010.4784. 2010 Oct 22.
4. Hoens TR, Polikar R, Chawla NV. **Learning from streaming data with concept-drift and imbalance: an overview.** *Progress in Artificial Intelligence*. 2012 Apr 1;1(1):89-101.
5. Moreno-Torres JG, Raeder T, Alaiz-RodríGuez R, Chawla NV, Herrera F. **A unifying view on dataset shift in classification**. *Pattern Recognition.* 2012 Jan 1;45(1):521-30. https://doi.org/10.1016/j.patcog.2011.06.019
6. Kolter JZ, Maloof MA. **Dynamic weighted majority: An ensemble method for drifting concepts.** *Journal of Machine Learning Research.* 2007;8(Dec):2755-90.
7. Hulten G, Spencer L, Domingos P. **Mining time-changing data streams.** In Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining 2001 Aug 26 (pp. 97-106). ACM. https://doi.org/10.1145/502512.502529
8. Bifet A, Gavalda R. **Learning from time-changing data with adaptive windowing.** *InProceedings of the 2007 SIAM international conference on data mining* 2007 Apr 26 (pp. 443-448). Society for Industrial and Applied Mathematics.
9. Gama J, Medas P, Castillo G, Rodrigues P. **Learning with drift detection**. *InBrazilian symposium on artificial intelligence* 2004 Sep 29 (pp. 286-295). Springer, Berlin, Heidelberg.
10. Baena-García M, del Campo-Ávila J, Fidalgo R, Bifet A, Gavaldà R, Morales-Bueno R**. Early drift detection method.** 2006.
11. Minku LL, Yao X. **DDD: A new ensemble approach for dealing with concept-drift.** *IEEE transactions on knowledge and data engineering.* 2012 Apr;24(4):619-33. https://doi.org/10.1109/TKDE.2011.58
12. Bartlett PL, Ben-David S, Kulkarni SR. Learning changing concepts by exploiting the structure of change. Machine Learning. 2000 Nov 1;41(2):153-74.
13. Minku FL, Yao X. **Using diversity to handle concept-drift in on-line learning**. *In Neural Networks, 2009. IJCNN 2009. International Joint Conference on* 2009 Jun 14 (pp. 2125-2132). IEEE.
14. Kosina P, Gama J, Sebastiao R. **Drift Severity Metric**. *In ECAI* 2010 Aug 4 (pp. 1119-1120).

15. Kullback S, Leibler RA. **On information and sufficiency.** *The annals of mathematical statistics*. 1951 Mar 1;22(1):79-86.
16. Hoens TR, Chawla NV, Polikar R. **Heuristic updatable weighted random subspaces for non-stationary environments**. *In2011 11th IEEE International Conference on Data Mining* 2011 Dec 11 (pp. 241-250). IEEE. https://doi.org/10.1109/ICDM.2011.75
17. Tsymbal A. **The problem of concept-drift: definitions and related work**. *Computer Science Department, Trinity College Dublin.* 2004 Apr 29;106(2).
18. Bose RJ, van der Aalst WM, Žliobaitė I, Pechenizkiy M. **Handling concept-drift in process mining**. *In International Conference on Advanced Information Systems Engineering* 2011 Jun 20 (pp. 391-405). Springer, Berlin, Heidelberg.
19. Huang DT, Koh YS, Dobbie G, Pears R. **Tracking drift types in changing data streams.** *In International Conference on Advanced Data Mining and Applications* 2013 Dec 14 (pp. 72-83). Springer, Berlin, Heidelberg.
20. Minku LL, White AP, Yao X. **The impact of diversity on online ensemble learning in the presence of concept-drift**. *IEEE Transactions on Knowledge and Data Engineering.* 2010 May;22(5):730-42.
21. Dongre PB, Malik LG. **A review on real time data stream classification and adapting to various concept-drift scenarios.** *In Advance Computing Conference (IACC), 2014 IEEE International* 2014 Feb 21 (pp. 533-537). IEEE. https://doi.org/10.1109/IAdCC.2014.6779381
22. Brzeziński D. **Mining data streams with concept-drift**. Cs. Put. Poznan. Pl. 2010:89.
23. Demšar, Jaka, and Zoran Bosnić. **Detecting concept drift in data streams using model explanation.** *Expert Systems with Applications 92* (2018): 546-559.
24. Liu A, Lu J, Liu F, Zhang G. Accumulating regional density dissimilarity for concept-drift detection in data streams. *Pattern Recognition.* 2018 Apr 1;76:256-72.
25. Jech T. ***Set theo*ry**. Springer Science & Business Media; 2013 Jun 29.
26. Borevich ZI, Shafarevich IR. ***Number theory***. Academic press; 1986 May 5.
27. Yang XS, Deb S. **Engineering optimisation by cuckoo search**. arXiv preprint arXiv:1005.2908. 2010 May 17. https://doi.org/10.1504/IJMMNO.2010.035430
28. Stolfo SJ, Fan W, Prodromidis A, Chan PK, Lee W. **Cost-sensitive modelling for fraud and intrusion detection: Results from the JAM project**. *In Proceedings of the 2000 DARPA Information Survivability Conference and Exposition* 2000.
29. K. Prasanna, M. Seetha and A. P. S. Kumar, "CApriori: Conviction based Apriori algorithm for discovering frequent determinant patterns from high dimensional datasets," 2014 International Conference on Science Engineering and Management Research (ICSEMR), Chennai, 2014, pp. 1-6. https://doi.org/10.1109/ICSEMR.2014.7043622
30. Lippmann RP, Fried DJ, Graf I, Haines JW, Kendall KR, McClung D, Weber D, Webster SE, Wyschogrod D, Cunningham RK, Zissman MA. **Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation.** *In DARPA Information Survivability Conference and Exposition, 2000. DISCEX'00. Proceedings* 2000 (Vol. 2, pp. 12-26). IEEE.
31. Pelleg D, Moore AW. X-means: Extending k-means with efficient estimation of the number of clusters. In Icml 2000 Jun 29 (Vol. 1, pp. 727-734).