International Journal of Engineering & Technology, 7 (3.6) (2018) 148-153

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Data Mining Models of High Dimensional Data Streams, and Contemporary Concept Drift Detection Methods: a Comprehensive Review

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.
[2]Assitant Professor, Department Of Computer Science And Engineering, JNTUCE, Anantapuramu, Andhra Pradesh, India.
[3]HOD And Associate Professor, Department Of Information Technology, AITS, Rajampet, Andhra Pradesh, India.
**Corresponding Author E-Mail:Sankaraprasannakumar@Gmail.Com*

Abstract

Concept drift is defined as the distributed data across multiple data streams that change over the time. Concept drift is visible only when the type of collected data changes after some stable period. The emergence of concept drift in data streams leads to increase misclassification and performing degradation of data streams. In order to obtain accurate results, identification of such concept drifts must be visible. This paper focused on a review of the issues related to identifying the changes occurred in the various multivariate high dimensional data streams. The insight of the manuscript is probing the inbuilt difficulties of existing contemporary change-detection methods when they encounter during data dimensions scales.

Keywords: *CUSUM, streaming ensemble algorithm, concept drift detection, dimensional data streams, change-detection tests, Hoteling's t-squared test, Bayesian Online Change Point Detection.*

1. Introduction

The advanced technology and its fast-growing nature have changed the world as the digital world. So, there is a continuous growth in the volume of automatic produced data. In general, the streams of generated digital data have three main complexities. These difficulties are categorized on the base of size, steady rate, and evolution in terms of time.

These issues are addressed by the computing system and computational intelligence systems. The traditional computing systems address the issue about the unbounded size of data and the arrival of high steady rate problems through the implementation of approximate and quick solutions promptly where the with advanced computational intelligence system resolve the complicated issue, data generated in the dynamic environment. This study of evolving data in the dynamic environment considerably has taken prominent attention over the past ten years. This study of resolving dynamism of the non-stationary data is familiar with the concept drift.

The conventional online learning algorithms are not able to control the dynamic evolution of data issue as they can't abandon the old and obsolete data. Hence the evolving learning algorithms which are adaptive can deal the concept-drift issue wisely as they can subsume unlearning mechanisms and discard the outdated concepts.

There are various definitions for different issues of the concept change such as a wish to alter the stock value during bear market but sudden bear market rally is not applicable. To surveillance, the lower limit will only decrease the false alarm of lower limit when monitoring the upper and lower limits of a control chart. To apply a heuristic method to an identified issue then the most significant data available will be trained data for the used method. The robust unsupervised approaches support to discuss the unfamiliar context to detect a change. There may be a swift or progressive change and a single or recurrent change. So, to detect a difference can be critical. The hypothesis on contextual properties will be more in different dimensions. Further, the change can evolve into a single feature, many features or different features. It is observed in the novelty detection literature [1] that unsupervised learning in data subspace usually performs better than unsupervised learning that considers entire data as single data.

The categorization literature states that the concept-drift [2] is considered when changes occurred in the distribution where detecting changes usually happen by supervising the evolving of errors over time in the scalar classification sequence [3], [4], [5].

2. The Taxonomy

Data Stream Challenges

Currently, researchers have paid attention to machine learning. And their prominent focus is on data stream categorization. Such works have dealt with several challenging issues [6] like:

- Arriving at online Data streams at high-speed rate steadily,
- Lacking the control on the order of incoming data
- Having potential limitless in size, caused to extract and store a precise information

Due to the above attributes, the machine learning algorithms can't deal with concept-drift as these algorithms need access to entire data to work offline.

Hence the process of online is appropriate to classify the data stream. So, the proposed model requires some prerequisites such as:

- Processing each example in specific groups and inspecting it for one time
- Using defined memory
- Working in the specified time,
- Readiness to forecast at whatever time.

Evolving and Adaptive Models

Continuous changing over a period of time/often times in concepts of streaming data is the powerful impact often lead by dynamic environments. Therefore, the main challenging issue is incoming data becomes non -stationary. For instance, the underneath distribution of any two points i and j may change to $D_i \neq D_j$. Consequently, the steady concept turn to concept drift and the traditional incremental models are not suitable to distribute the fresh incoming data. Hence to handle concept changes in some non-stationary environment adaptive and evolving algorithms perform successfully.

By assimilating forgetting mechanisms and discarding obsolete data, these adaptive learning algorithms deals the concept of drift appropriately. The core aim of this model is to supervise the process of learning, the adaption of the model, interpretation of changes encounter in response to brand new environment necessities. Framing such self-configuring learners is the task of scientific engineering. Hence there are three main issues are considered: (i) tracking the concept drift, (ii) filtering the data to be stored as well as discarded data (iii) adapting the learner parameters and structure for responding to these brand-new environment needs.

The enhanced version of incremental algorithms is the evolving learning models. These evolving models obviously deal the concept-drift which is a noteworthy feature. These models have an inbuilt mechanism to expand knowledge automatically mainly when the drift carries data distribution happens to an unknown space. Therefore, about such conditions the structure of evolving model develops and leads to the easy transition of existing model to new model.

Types of Concept Drift

According to Bayes theorem proposed the concept drift happens with any of the variables. The important three variables that may change the class members at posterior distribution, distribution of single or many classes, and class prior over a time period.

The change in the posterior probability is known as the Real concept drift. The posterior probability is characteristics changes of target concept values and presented in Fig. 1a. Such real concept drifts straightly impact the decision limitations leading to reduce the performance of learner. The techniques based on feedback prediction and learner's performance handle this Real Concept Drift.

Changes arise at the conditional probability of a class due to the influence applied on the posterior probability is called Virtual concept drift. That means, distributing the data in the Virtual concept-drift indicates the changes in the selfsame class barely impact the decision limitations (represented in Fig. 1b). Usually, several techniques heed on the distribution of input data and find changes in the conditional class deal this Virtual Concept Drift.

Class prior concept drift denotes the changes raised at the prior probability of a class. It shows an imbalance in the class, emerge from brand new class or, merge of current classes.

Here the important point is the Class Prior Concept Drift is not taken in the category of concepts drift when the surveys [2], [7] on various concept drift were conducted. However, it is viewed as both Virtual drift and Real drift based on its functions. That means, when class imbalance caused with hardly influencing the decision limitations it falls under the category of Virtual drift since prior class evolution occurred with impacting the decision limitations then it comes under real drift category.

At first in [8] the use of word Class prior concept drift mentioned. Classifying this drift into a separate category happens with relevance to the real-time applications in real. For example, the development of resistance in a few viruses to powerful antibiotics occurs in the biomedical domain. Later, in future, the resistant class viruses may evolve into a superior class. Then in such issue, the change happened with prior class distribution leads to Class prior concept drift.

Concept Drift Characteristics

The time taken during change occurred from an old concept to a new one, such drift duration is called drift width or drifting time. Research work [9] proclaims that speed is inversely proportionate to the drifting time which means high-speed is proportionate to lower time steps and visa-versa further based on speed, the drifts are classified into abrupt or gradual.

a) Within a less drifting time when new concepts supersede the old or existing concept, then the abrupt drift happens. The abrupt drift collapses the efficiency of the learner by abruptly replacing the old concept with the new concept in haste.
b) When the drifting time is proportionately high, then the Gradual drift occurs. The identification of the gradual drift is difficult because of unpredictable created between steady states. This drift is sorted into two types (1) probabilistic (2) continuous.

- The Gradual probabilistic drift is an active time period of both new as well as old concepts
- The Gradual, continuous drift is the concept alone undergoes a continuous change that means changing from the existing concept into a new concept changing at each time step

Recurrence

The recurrence concepts are another attribute of concept drift where the antecedent active concepts are appearing again after some period of time. The work [9] presents cyclic and acyclic behavior of the recurrent drifts.

The cyclic recurrent drift is a seasonal or periodic drift like the rise of electric products price based on the seasonal market demand for the product.

The acyclic recurrent drift is an uncertain period drift as there is no certainty when the concept reoccurs again. This acyclic drift is like a sudden rise and fall of electricity prices due to fluctuation in petrol prices.

It is important to know that the concept recurrence emerges as abrupt or gradual concept drift. Further, these drifts show impact an instance space locally or globally. Thus, in real-time, a regular mix of several drift attributes are learned in the transition time period.

Predictability

In work [9] the criterion predictability was mentioned primarily. This predictability points out the nature of the drift whether it is random in the time period or a pattern time indicator:

The drift predictability is rationalized for two purposes

(i) When there is a scope of drift predictability, then it becomes effortless to learn the origin of the drift and its future effects.
(ii) This predictable drift is easy to handle the concept with a little time lag to detect the change and false alarm rate. This is the most required attribute in several real-time applications.

3. Review of the Literature

The concept drift has influenced multidisciplinary domains grabbing the attention of the scholars of various disciplines. This multidisciplinary research has addressed various issues with multiple application areas. The works [10], [11] refer its application in the field of medicine. The works [12], [13] express the contribution in monitoring and control sector. [14], [15] depicts the research application on management and strategic planning domain. The authors of [16], [17] works state the use of application in personal assistance field. The authors of [18], [19] refers application in the ubiquitous environmental information domain. With increasing attention towards the development of approaches to deal drift concept, there has been an ample literature enhancement. Nevertheless, the issues stated in the literature vary. Further, the ideas associated with the individual concepts formulation refer with various terminologies due to that reason many current surveys [2], [7] are conducted.

In general, change detection has explained under the concept of supervised or unsupervised. The learning algorithm of error data process the target data or supervise the target data that doesn't have class labels for calculating the rate of error.

An extensive study was carried out to detect a change in scalar data named univariate data. The works [20] depicts that univariate case issue is the unsupervised detection of a change of an independent entity that has been thoroughly investigated for the past 50 years. The research has introduced popular approaches like CUSUM ("control-charts & cumulative-sum-chart"). The literature of various disciplines discusses different univariate methods. In 1993 a research on detectors of sudden change published [21] a monograph.

The works [22] contributes comprehensive reviews on the overlapping of the novelty detection, as well as outlier detection in work [23]. From the literature of classification, several approaches are proposed to supervise the incoming data error-rate along with discarding classifiers. The Massive Online Analysis framework is one of the most used open source tools for data mining of streams. It also supplies several approaches to detect univariate change. Possibly, the use of the univariate assumptions is because of its quality-control applications [21] included in the change-detection tests (CDTs). Moreover, a little research has done issues related to change detection of a multivariate data streams.

The referred works [24], [25], [26] propose that Multivariate approaches are assumed as extracting examples from a multivariate process where the feature is not independent. This Multivariate change detection tries to frame the multivariate process through a particular function which can check the incorporation of new data that may be an example or group of examples into the model. A few works [27] states solo monitoring of components and some works [1] indicates that the multivariate approach can't react to the changes occurred in the correlated components considered as a drawback based on the change and its context.

The work [28], proposed that using scalar CDT, every single component of the data stream is inspected to detect concept drift in a multivariate data when it is a straightforward case. Even though, it can't give a suitable multivariate solution to the straightforward case as the affected correlation among the components of the data components is not noticed while detecting the change. Depending on the model suitable for training a set of training stationary data, various CDTs compute log-likelihood for multivariate streams of data. The paper has studied various models that used CDTs such as [24] Gaussian mixtures, [29], [30] nonparametric density models. Furthermore, [24], [31], [32], [26] these works present some more CDTs which is built on unique multivariate statistics.

The work [24] proposes that Multivariate detectors view the features of multivariate as hidden components of the multivariate distribution. The mentioned above Multivariate detectors are considered as 'pure' detectors. These pure detectors outperform with a low volume of data. Moreover, the incoming data of the concept has to be in the sequence of $i.i.d$. This is observed very rarely in the real-time application. The work [27] claimed that in network intrusion detection field assuming example in an $i.i.d$ sequence is a restrictive practice.

Many algorithms of change point detection depend on modeling of time series that needs preceding knowledge of the related domain. The work [33] presents that the commonly selected sequence models are Hidden Markov Models which have limitations of defining simple patterns. The referred work [34] states the utilization of online algorithm to detect the change of occurrence of change at any particular time point to study the data [35], [36], [37] by Bayesian Online Change Point Detection BOCPD [34]. However, these algorithms also need a prior model.

On the contrary, these Machine learning methods detect the patterns which are hard for a model. These learning methods are often used in non-stationary environments. In such environment over a time period, the concepts change erratically. There is a need for change detection methods to supervise an error-rate in an individual univariate stream from a learner when there is a chance of immediate accessibility of class labels. If an opportunity of detecting the change in the error stream then as per requirement a model can be retained or adapted. Nevertheless, if there is no availability of labels, the error rate is not considered as an indicator to measure performance. In such cases, an unsupervised approach is recommended.

The works [38], [39] presents supervised learning methods to solve these problems. However, there are not responsible for a gradual change. The research work [40] focused on detecting the gradual change points aside analyze error prediction of a learned model. The work [41] contributes to the extension of BOCPD for identifying abnormal segments which are against to the points as well as detecting the gradual changes. The wavelets are used to process time series in a multi-resolution way vividly. These effects the detection of the change point in scale-invariant.

The works [42], [43], [44], [45] express the use of Nonparametric methods to solve alike issues. These methods do not depend on data distribution assumptions and perform low if the volume of data is low.

The surveys [2], [7] contributed dissimilarities of virtual concept drift and real concept-drift where the former is a change with optimal decision boundary in class conditional probabilities class, and the latter is a change in data distribution or prior probabilities. The work refers definition of the latter as the unsupervised environment doesn't contain privilege real concept detection, which due to the lack of class labels. Henceforth, the detection of concept drifts in supervised learning environment carried through the process of detecting outliers, novelty, and variance in statistics of the data. These detection applications are unsupervised and are supposed to use right away with the data of any domain.

The detection of the change in multivariate methods needs two components. They are (I) estimation of incoming data distribution and (2) a test for calculating the probability new data points suits to the model. The clustering or multivariate distribution modeling estimates the streaming data distribution. For modeling, a multivariate process of novel uses a most common parametric like in the works of [25]. The contribution depicted in [46] estimating the kernel density to identify the drift in concept. The work [47] presents similar effects of construct KDQ trees. The process of tracking online clustering trajectories contributed by [48]. The use of cluster deviation leads to detect the data stream evolution proposed by [49]. The application of k clustering with input data as well as use of cluster populations to estimate the data distribution presented by [26].

The Hotelling's t-squared test [50] which is a Multivariate statistical test used to compare requires the adaption of a sequential form of windows data over time [24]. The Bespoke statistics works stated [26], [24] the undergoing continual improvisation for this reason. Further, it creates a log-likelihood ratio detectors family. These families compute the probability of distribution similarity between the pair of time-windows in sequence. The findings show that the input space usage is minimal in regard to log-likelihood centric detectors, where it is fully utilized to a univariate statistic by observing the ratio of the current univariate methods [51].

An ensemble method supports to monitor data streams evolution has increased the attention in the literature of change-detection. Of lately, surveys [52] and [2] are conducted on these areas. The work [52] states that small-scale research carried out on drift detection methods collaboration. The review of research publications [53], [54] on this subject presents work carried out on the collaboration of detectors of a univariate input data on the contrary to the formulation shown in the paper. The work [2] explored the nomenclature of the ensemble classification strategies applied on data streams. Further, the work [55] examined the ensemble concept drift detection process applied on univariate error stream of labeled records. The research in [56] explored a method that detects concept drift in the hierarchical order that enabled the fast detection of the sequential change that includes offline change detection.

The [57], [58] classification literature discuss ensemble change detection use of these techniques for monitoring classifiers accuracy rate to decide the retention and replacement of the classifier in an ensemble. Here in [3], [59] the work, most commonly univariate methods used to detect change are updated for the supervised approach that provides a discrete error stream. Algorithm referred by the work [60] is the streaming ensemble algorithm (SEA) which is found as foremost of several ensembles approaches for addressing the streaming supervised learning issues. Nevertheless, SEA introduces a robust adaptive classifier of concept drift that doesn't depend on a change detection method. In work [1] contributes the use of a subspace ensemble in novelty context detection with a single class support vector machine classifier. There is a division of input space into three random subspaces where an individual ensemble member monitors every subspace. The work Kuncheva [61] expressed the utilization of an ensembles classifier right way for detecting concept-drift in data with no class labels. The tabular representation of the contemporary models reviewed in this section follows (see Table 1).

Table 1: Properties of the Contributions Reviewed from the Contemporary Literature

Domain Specific		
[10], [11]		Field-of-medicine;
[12], [13]		Monitoring-and-control sector;
[14], [15]		Management-and-strategic-planning domain;
[16], [17]		Personal-assistance-field;
[18], [19]		Ubiquitous-environmental-information-domain.
[2], [7], [22]		Surveys
[27]		Network-intrusion-detection field
Unsupervised		
[23]	Univariate	Outlier-detection
[24], [25], [26]	Multivariate	By example models, can't react to the changes occurred in the correlated components considered
[27], [28]	Multivariate	solo monitoring of components using ensemble scalar CDT, every single component of the data stream is inspected to detect concept drift in a multivariate data, can't react to the changes occurred in the correlated components considered
[29], [30]	Multivariate	nonparametric density models
[31], [32], [26]	Multivariate	'Pure' detectors, outperform with a low volume of data
Supervised		
[33]	Machine learning	Uses commonly selected sequences as Hidden Markov Models, limits to define simple patterns.
[34], [35], [36], [37]	Machine learning	Uses online algorithm, performs low on high dimensional data
[34]	Machine learning	Bayesian Online Change Point Detection BOCPD, performs low on high dimensional data
[41]	Machine learning	Extension of BOCPD, detects gradual changes
[38], [39]	Machine learning	Not responsible for a gradual change.
[40]	Machine learning	Detecting the gradual change points, analyzes error prediction of a learned model.
[42], [43], [44], [45]	Nonparametric methods	Not depends on data distribution assumptions
[46], [47]	Multivariate	Estimating the kernel density to identify the drift in concept.
[48]	Multivariate	Tracking online clustering trajectories;
[49]	Multivariate	Cluster deviation
[26]	Multivariate	k-clustering with input data, using cluster populations to estimate the data distribution.
[50]	Multivariate	t-squared test, computes the probability of distribution similarity between the pair of time-windows in sequence.
[53], [54]	Ensemble	The collaboration of detectors of a univariate input data
[55]	Ensemble	Ensemble concept drift detection, univariate error stream of labeled records.
[56]	Ensemble	Detects concept drift in hierarchical order, enabled the fast detection of sequential change, also performs offline change detection
[57], [58]	Ensemble	Ensemble change detection methods, monitoring classifiers accuracy rate to switch from one classifier to other
[3], [59]	Ensemble	Ensembles updated univariate methods, performs drift detection in supervised learning, provides discrete error stream
[60]	Ensemble	Streaming ensemble algorithm (SEA), addressing the streaming supervised learning issues, doesn't depend on a change detection method.
[1]	Ensemble	Subspace ensemble method, detects novelty context relies on single class SVM
[61]	Ensemble	Ensembles classifier, detects concept-drift in data with no class labels.

4. Conclusion

The manuscript endeavored to insight the concept drift and its impact on mining models over data streams. In addition, explore the precise review on the nomenclature of the concept drift and concept drift detection strategies available in the contemporary literature. From the insights of the contributions, it is learned that the concept drift detection is challenging and existing models have not fulfilled the need of concept drift detection in all possible dimensions of the streaming data context. In particular, the concept drift of domain-specific requirements, and concept drift due multivariate elements of high dimensional data streams are considerable challenges of the emerging drift detection strategies. Hence, it is obvious to conclude the research scope towards proposing new and novel concept drift detection strategies to challenge the constraints such as domain-specific concept drift and concept drift due to multivariate elements.

References

[1] Evangelista PF, Embrechts MJ & Szymanski BK, "Taming the curse of dimensionality in kernels and novelty detection", Applied soft computing technologies: The challenge of complexity, (2006), pp.425-438.

[2] Gama J, Žliobaitė I, Bifet A, Pechenizkiy M & Bouchachia A, "A survey on concept drift adaptation", ACM computing surveys (CSUR), (2014).

[3] Gama J, Medas P, Castillo G & Rodrigues P, "Learning with drift detection", Brazilian symposium on artificial intelligence, (2004) pp.286-295.

[4] Alippi C, Boracchi G & Roveri M, "Just-in-time classifiers for recurrent concepts", IEEE transactions on neural networks and learning systems, Vol.24, No.4,(2013), pp.620-34.

[5] Ross GJ, Adams NM, Tasoulis DK & Hand DJ, "Exponentially weighted moving average charts for detecting concept drift", Pattern recognition letters, Vol.33, No.2, (2012), pp.191-8.

[6] Gama J, Knowledge discovery from data streams, CRC Press, (2010).

[7] Ditzler G, Roveri M, Alippi C & Polikar R, "Learning in nonstationary environments: A survey", IEEE Computational Intelligence Magazine, Vol.10, No.4,(2015), pp.12-25.

[8] Khamassi I, Sayed-Mouchaweh M, Hammami M & Ghédira K, "Self-adaptive windowing approach for handling complex concept drift", Cognitive Computation, Vol.7, No.6,(2015), pp.772-90.

[9] 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, Vol.22, No.5,(2010), pp.730-42.

[10] Tsymbal A, Pechenizkiy M, Cunningham P & Puuronen S, "Handling local concept drift with dynamic integration of classifiers: Domain of antibiotic resistance in nosocomial infections", 19th IEEE International Symposium on Computer-Based Medical Systems, (2006), pp. 679-684.

[11] Sebastião R, Silva MM, Rabiço R, Gama J & Mendonça T, "Real-time algorithm for changes detection in depth of anesthesia signals", Evolving Systems, Vol.4, No.1,(2013), pp.3-12.

[12] Toubakh H & Sayed-Mouchaweh M, "Hybrid dynamic data-driven approach for drift-like fault detection in wind turbines", Evolving Systems, Vol.6, No.2, (2015), pp.115-29.

[13] Navarro-Gonzalez JL, Lopez-Juarez I, Ordaz-Hernandez K & Rios-Cabrera R, "On-line incremental learning for unknown conditions during assembly operations with industrial robots", Evolving Systems, Vol.6, No.2, (2015), pp.101-14.

[14] Sun J, Li H & Adeli H, "Concept drift-oriented adaptive and dynamic support vector machine ensemble with time window in corporate financial risk prediction", IEEE Transactions on Systems, Man, and Cybernetics: Systems, Vol.43, No.4,(2013), pp.801-13.

[15] Aloraini A, "Penalized ensemble feature selection methods for hidden associations in time series environments case study: equities companies in saudi stock exchange market", Evolving Systems, Vol.6, No.2,(2015), pp.93-100.

[16] Wang S, Minku LL & Yao X, "Online class imbalance learning and its applications in fault detection", International Journal of Computational Intelligence and Applications, Vol.12, No.04,(2013).

[17] AlZoubi O, Fossati D, D'Mello S & Calvo RA, "Affect detection from non-stationary physiological data using ensemble classifiers", Evolving Systems, Vol.6, No.2,(2015), pp.79-92.

[18] Tünnermann J & Mertsching B, "Region-based artificial visual attention in space and time", Cognitive computation, Vol.6, No.1, (2014), pp.125-43.

[19] Amiribesheli M, Benmansour A & Bouchachia A, "A review of smart homes in healthcare", Journal of Ambient Intelligence and Humanized Computing, Vol.6, No.4,(2015), pp.495-517.

[20] Wald A, Sequential analysis, Courier Corporation, (1973).

[21] Basseville M & Nikiforov IV, Detection of abrupt changes: theory and application, Englewood Cliffs: Prentice Hall, (1993).

[22] Pimentel MA, Clifton DA, Clifton L & Tarassenko L, "A review of novelty detection", Signal Processing, Vol.99,(2014), pp.215-49.

[23] Ben-Gal I, "Outlier detection", Data mining and knowledge discovery handbook, (2005), pp.131-146.

[24] Kuncheva LI, "Change detection in streaming multivariate data using likelihood detectors", IEEE Transactions on Knowledge and Data Engineering, Vol.25, No.5,(2013), pp.1175-80.

[25] Zorriassatine F, Al-Habaibeh A, Parkin RM, Jackson MR & Coy J, "Novelty detection for practical pattern recognition in condition monitoring of multivariate processes: a case study", The International Journal of Advanced Manufacturing Technology, Vol.25, No.9-10, (2005),pp. 954-63.

[26] Nguyen TD, Du Plessis MC, Kanamori T & Sugiyama M, "Constrained least-squares density-difference estimation", IEICE TRANSACTIONS on Information and Systems, Vol.97, No.7,(2014), pp.1822-9.

[27] Tartakovsky AG, Rozovskii BL, Blažek RB & Kim H, "Detection of intrusions in information systems by sequential change-point methods", Statistical methodology, Vol.3, No.3,(2006), pp.252-93.

[28] Tartakovsky AG, Rozovskii BL, Blazek RB & Kim H, "A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods", IEEE Transactions on Signal Processing, Vol.54, No.9, (2006), pp.3372-82.

[29] Krempl G, "The algorithm APT to classify in concurrence of latency and drift", International Symposium on Intelligent Data Analysis, (2011), pp.222-233.

[30] Dyer KB, Capo R & Polikar R, "Compose: A semisupervised learning framework for initially labeled nonstationary streaming data", IEEE transactions on neural networks and learning systems, Vol.25, No.1,(2014), pp.12-26.

[31] Lung-Yut-Fong A, Lévy-Leduc C & Cappé O, "Robust change-point detection based on multivariate rank statistics", A IEEE International Conference on coustics, Speech and Signal Processing (ICASSP), (2011), pp.3608-3611.

[32] Ditzler G & Polikar R, "Hellinger distance-based drift detection for nonstationary environments", IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE), (2011), pp.41-48.

[33] Montanez GD, Amizadeh S & Laptev N, "Inertial Hidden Markov Models: Modeling Change in Multivariate Time Series", AAAI, (2015), pp.1819-1825

[34] Adams RP & MacKay DJ, "Bayesian online change point detection", arXiv preprint arXiv:0710.3742, (2007).

[35] Cavalcante RC, Minku LL & Oliveira AL, "Fedd: Feature extraction for explicit concept drift detection in time series", International Joint Conference on Neural Networks (IJCNN), (2016), pp.740-747.

[36] Barnett I & Onnela JP, "Change point detection in correlation networks", Scientific reports, (2016).

[37] Idé T, Phan DT & Kalagnanam J, "Change Detection Using Directional Statistics", IJCAI, (2016), pp.1613-1619.

[38] Yamada M, Kimura A, Naya F & Sawada H, "Change-Point Detection with Feature Selection in High-Dimensional Time-Series Data", IJCAI, (2013), pp.1827-1833.

[39] Hocking T, Rigaill G, Vert JP & Bach F, "Learning sparse penalties for change-point detection using max margin interval regression", International conference on machine learning, (2013), pp.172-180.

[40] Harel M, Mannor S, El-Yaniv R & Crammer K, "Concept drift detection through resampling", International Conference on Machine Learning, (2014), pp.1009-1017.

[41] Bardwell L & Fearnhead P, "Bayesian detection of abnormal segments in multiple time series", Bayesian Analysis, Vol.12, No.1, (2017), pp.193-218.

[42] Cabrieto J, Tuerlinckx F, Kuppens P, Grassmann M & Ceulemans E, "Detecting correlation changes in multivariate time series: A comparison of four non-parametric change point detection methods", Behavior research methods, Vol.49, No.3,(2017), pp.988-1005.

[43] Jones M, Nikovski D, Imamura M & Hirata T, "Exemplar learning for extremely efficient anomaly detection in real-valued time series", Data Mining and Knowledge Discovery, Vol.30, No.6, (2016), pp.1427-54.

[44] Qahtan AA, Alharbi B, Wang S & Zhang X, "A pca-based change detection framework for multidimensional data streams: Change detection in multidimensional data streams", Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2015), pp.935-944.

[45] Guha S, Mishra N, Roy G & Schrijvers O, "Robust random cut forest based anomaly detection on streams", International Conference on Machine Learning, (2016), pp.2712-2721.

[46] Song X, Wu M, Jermaine C & Ranka S, "Statistical change detection for multi-dimensional data", Proceedings of the 13th ACM

SIGKDD international conference on Knowledge discovery and data mining, (2007), pp.667-676.
[47] Dasu T, Krishnan S, Venkatasubramanian S & Yi K, "An information-theoretic approach to detecting changes in multi-dimensional data streams", Proc. Symp. on the Interface of Statistics, Computing Science, and Applications, (2006).
[48] Krempl G, Siddiqui ZF & Spiliopoulou M, "Online clustering of high-dimensional trajectories under concept drift", Proceedings of the European conference on Machine learning and knowledge discovery in databases-Volume Part II, (2011), pp.261-276.
[49] Gaber MM & Yu PS, "Classification of changes in evolving data streams using online clustering result deviation", Proc. Of International Workshop on Knowledge Discovery in Data Streams, (2006).
[50] Hunt KM & Turner AG, "The effect of soil moisture perturbations on Indian monsoon depressions in a numerical weather prediction model", Journal of Climate, Vol.30, No.21,(2017), pp.8811-23.
[51] Faithfull WJ & Kuncheva LI, "On Optimum Thresholding of Multivariate Change Detectors", Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), (2014), pp.364-373.
[52] Krawczyk B, Minku LL, Gama J, Stefanowski J & Woźniak M, "Ensemble learning for data stream analysis: A survey", Information Fusion, Vol.37, (2017), pp.132-56.
[53] Maciel BI, Santos SG & Barros RS, "A lightweight concept drift detection ensemble", IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI), (2015), pp.1061-1068.
[54] Woźniak M, Ksieniewicz P, Cyganek B & Walkowiak K, "Ensembles of Heterogeneous Concept Drift Detectors-Experimental Study", 15th IFIP International Conference on Computer Information Systems and Industrial Management, (2016), pp.538-549.
[55] Du L, Song Q, Zhu L & Zhu X, "A selective detector ensemble for concept drift detection", The Computer Journal, Vol.58, No.3, (2014), pp.457-71.
[56] Alippi C, Boracchi G & Roveri M, "Hierarchical change-detection tests. IEEE transactions on neural networks and learning systems", Vol.28, No.2, (2017), pp.246-58.
[57] Bifet A, Frank E, Holmes G & Pfahringer B, "Ensembles of restricted hoeffding trees", ACM Transactions on Intelligent Systems and Technology (TIST), Vol.3, No.2, (2012).
[58] Frías-Blanco I, del Campo-Ávila J, Ramos-Jiménez G, Morales-Bueno R, Ortiz-Díaz A & Caballero-Mota Y, "Online and non-parametric drift detection methods based on Hoeffding's bounds", IEEE Transactions on Knowledge and Data Engineering, Vol.27, No.3, (2015), pp.810-23.
[59] Baena-García M, del Campo-Ávila J, Fidalgo R, Bifet A, Gavaldà R & Morales-Bueno R, "Early drift detection method", (2006).
[60] Street WN & Kim Y, "A streaming ensemble algorithm (SEA) for large-scale classification", Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, (2001), pp.377-382.
[61] Kuncheva LI, "Classifier ensembles for detecting concept change in streaming data: Overview and perspectives", 2nd Workshop SUEMA, (2008), pp.5-10.