JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
https://e-journal.uum.edu.my/index.php/jict

How to cite this article:
Palli, A, S., Jaafar, J., Gilal, A. R., Alsughayyir, A., Gomes, H. M., Alshanqiti, A., & Omar, M. (2024). Online machine learning from non-stationary data streams in the presence of concept drift and class imbalance: A systematic review. *Journal of Information and Communication Technology, 23*(1), 105-139. https://doi.org/10.32890/jict2024.23.1.5

Online Machine Learning from Non-stationary Data Streams in the Presence of Concept Drift and Class Imbalance: A Systematic Review

*[1]Abdul Sattar Palli, [2]Jafreezal Jaafar, [3]Abdul Rehman Gilal,
[4]Aeshah Alsughayyir, [5]Heitor Murilo Gomes,
[6]Abdullah Alshanqiti & [7]Mazni Omar
[1&2]Computer and Information Science Department,
University Technology Petronas, Malaysia
[1]Anti-Narcotics Force, Ministry of Narcotics Control, Pakistan
[2]Center for Research in Data Science (CERDAS),
University Technology Petronas, Malaysia
[3] Knight Foundation School of Computing and Information Sciences,
Florida International University, United States
[4]College of Computer Science and Engineering,
Taibah University, Saudi Arabia
[5]School of Engineering and Computer Science,
Victoria University of Wellington
[5]AI Institute, University of Waikato, Wellington, New Zealand
[6]Faculty of Computer and Information Systems,
Islamic University of Madinah, Saudi Arabia
[7]School of Computing, Universiti Utara Malaysia, Malaysia

*[1]abdul_20000263@utp.edu.my
[2]jafreez@utp.edu.my
[3]arehman@fiu.edu
[4]asughayyir@taibahu.edu.sa
[5]heitor.gomes@vuw.ac.nz
[6]amma@iu.edu.sa

[7]mazni@uum.edu.my
*Corresponding Author

Received: 28/8/2023 Revised: 23/1/2024 Accepted: 24/1/2024 Published: 30/1/2024

ABSTRACT

In IoT environment applications generate continuous non-stationary data streams with in-built problems of concept drift and class imbalance which cause classifier performance degradation. The imbalanced data affects the classifier during concept detection and concept adaptation. In general, for concept detection, a separate mechanism is added in parallel with the classifier to detect the concept drift called a drift detector. For concept adaptation, the classifier updates itself or trains a new classifier to replace the older one. In case, the data stream faces a class imbalance issue, the classifier may not properly adapt to the latest concept. In this survey, we study how the existing work addresses the issues of class imbalance and concept drift while learning from non-stationary data streams. We further highlight the limitation of existing work and challenges caused by other factors of class imbalance along with concept drift in data stream classification. Results of our survey found that, out of 1110 studies, by using our inclusion and exclusion criteria, we were able to narrow the pool of articles down to 35 that directly addressed our study objectives. The study found that issues such as multiple concept drift types, dynamic class imbalance ratio, and multi-class imbalance in presence of concept drift are still open for further research. We also observed that, while major research efforts have been dedicated to resolving concept drift and class imbalance, not much attention has been given to with-in-class imbalance, rear examples, and borderline instances when they exist with concept drift in multi-class data. This paper concludes with some suggested future directions.

Keywords: Concept Adaptation, Concept Drift, Class Imbalance, Data Streams, Non-stationary.

INTRODUCTION

In today's ever-expanding world, enormous amounts of data have been created at a breakneck speed generally known as data streams. A data stream is an endless sequence of data instances that continuously arrive over time (Margara & Rabl, 2019). The most difficult phase of this task is dealing with such a large volume of data and evaluating it. Sensor-based applications such as monitoring Remaining Useful Life in an industrial environment (Fauzi et al., 2019; Jayasinghe et al., 2018; Li et al., 2018), traffic congestion (Metz, 2018), weather prediction (Zhang et al., 2022), banking (Wang et al., 2021), the stock markets (Zhong & Enke, 2019), education (Pongpaichet et al., 2020), telecommunications (Amin et al., 2019), healthcare (Fazal & Alotaibi, 2020), network data monitoring (Abbasi et al., 2021) in computer-based distributed applications all are common examples of applications that generate data streams. The data stream is non-stationary which means its statistical properties change over time. This changing behaviour of data streams leads to a change in a concept known as concept drift; and a change in class distribution known as class imbalance (Ghazikhani et al., 2013; Khandekar & Srinath, 2020; Palli et al., 2022). In online machine learning environment where a machine learning model deals with continuous streaming data, when the concept of data $\boldsymbol{D}$ changes at a certain time $\boldsymbol{T}$ i.e., $\boldsymbol{D_T}$, the model $\boldsymbol{M}$ trained on old data $\boldsymbol{D_{T-1}}$ loses the classification performance for $\boldsymbol{D_T}$. To address this issue, either the existing model $\boldsymbol{M}$ need to be retrained on new data $\boldsymbol{D_T}$ or an ensemble approach is used where a new model is trained on the latest data to overcome the concept drift. In both cases, if the latest data is imbalanced, the new model may not be properly trained for all classes. Due to the imbalanced data, it becomes biased towards the majority classes (Branco et al., 2016; Spelmen & Porkodi, 2018).

A phenomenon known as concept drift describes a circumstance in which the statistical properties of a target class fluctuate randomly over time. These modifications may be the result of changes in elements that cannot be easily quantified or identified. Mainly two types of concept drift exist, known as virtual drift and real drift (Agrahari & Singh, 2021), as illustrated in Figure 1. Virtual drift may be thought of as a shift in the conditional probabilities **P(y|x),** which means the change occurs only in the data distribution without affecting the class boundary. In the context of evolving data streams, if a new feature emerges or becomes more prevalent within a specific class, it can introduce a shift in the data distribution for that class. This shift might affect the model's ability to accurately classify instances within that class based on the existing set of features. Whereas the real drift can be thought of as a shift in the unconditional probability distribution **P(x)** that can affect the relationships between features and labels, potentially impacting the class boundaries learned by a model. Real drift can occur when there are shifts in the environment, user behaviour, or other external factors. The performance of a model (classifier) may be affected by both drift types at the same time or in distinct ways. In terms of time, according to Krawczyk and Cano (2018) a change in concept occurs in four different ways known as (1) sudden drift (2) gradual drift (3) incremental drift, and (4) reoccurring drift.

In classification applications, the data is grouped into different categories generally called classes. An imbalance between the two classes occurs when one class, the minority group, contains a significantly a smaller number of samples than the other class, the majority group. For example, in a binary classification problem to predict whether an email is spam or not, if 90 percent of the emails are non-spam and only 10 percent are spam, there is a class imbalance ratio of 9:1 (majority class to minority class). Because of its greater prior probability, learners often overclassify the majority group when there is a class imbalance in the training data. As a result, incidents involving the minority group are misclassified more often than those involving the majority group (Branco et al., 2016; Spelmen & Porkodi, 2018). The application may have binary-class data such as anomaly detection by Meseguer et al. (2010) or can have multi-class data such as identifying the type of brain tumor (Kumar et al., 2021).

The classification of data can be simple if there are no overlapping samples, and no rare examples (samples) among the classes. In such datasets the class boundary or a hyperplane between the classes can easily be drawn. However, in case the data has issues of overlapping samples, rare examples, it become hard for a machine learning model to classify such data. The illustration of such simple and complex data scenarios is given in Figure 2. Compared to the binary class data the classification of an imbalanced data stream with more than two classes is a highly troublesome job. With the same amount of data, the computational cost of solving a multi-class problem is often higher than that of solving a binary classification problem (Palli et al., 2022).

Figure 1

Types of Drifts: Circles Represent Instances; Different Colours Represent Different Classes. (Gama et al., 2014)

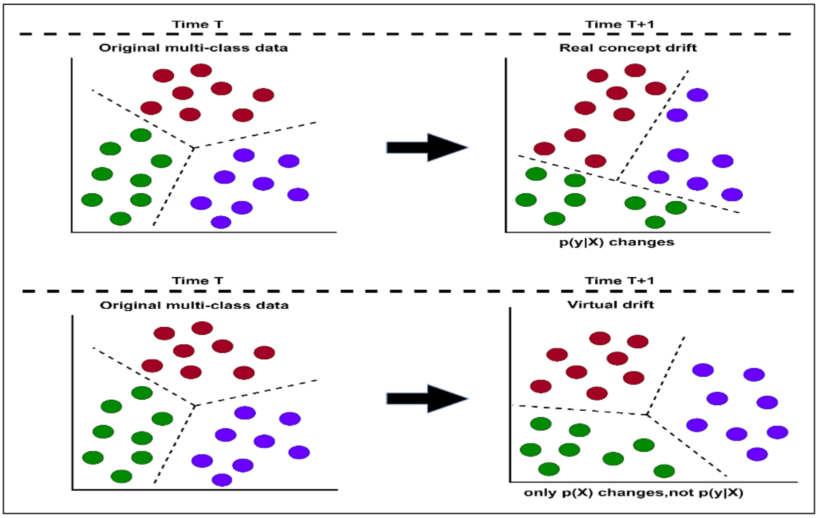
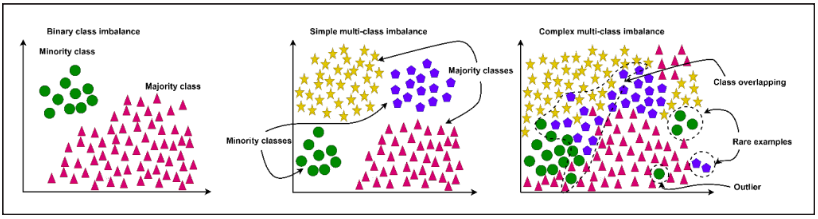


Figure 2

Class Imbalance Scenarios



Detecting and classifying time-changing concepts is very challenging for the classifier in the context of concept drift. A lack of predictability in the occurrence of concept drift results in a lack of capacity to train the classifier with such data, which in turn has an impact on performance (Karbasi, 2020; Palli et al., 2021). On the other hand, due to the class imbalance, it is very difficult for the classifier to learn tiny class events in cases of skew data or infrequent or uneven class distribution (Korycki & Krawczyk, 2021; Taha et al., 2021). In situations where both issues occur simultaneously, the classification performance is more affected (Lu et al., 2020). The complex multi-class imbalance with other factors such as class overlapping, and rare examples cause more difficulty in retraining the new classifier to address the changing behaviour of data (Brzezinski et al., 2021). During online learning, there are constraints of limited data, time, and memory for retaining the new model (Bahri et al., 2021).

Addressing the joint problem and keeping the model relevant to the most recent data many strategies are suggested in the literature. As the study proposed by Li et al. (2020) suggested Dynamic Updated Ensemble (DUE) which addresses class imbalance and concept drift in streaming data. Numerous potential classifiers are weighted piecewise for each block. The worst-performing classifiers are eliminated when the ensemble size exceeds a certain threshold. In situations where the minority and majority classes are switched, bagging is used to balance the training set for candidate classifiers. Each block has numerous candidate classifiers weighted piecewise. When the ensemble size reaches a predetermined threshold, the worst-performing classifiers are discarded. Bagging is used to balance the training set for candidate classifiers when the minority and majority classes are shifted. An offline classifier-based ensemble learning model was presented by Lin et al. (2019) for use in condition-based maintenance in manufacturing. The suggested paradigm consists of three stages: ensemble learning, concept drift identification, and concept drift adaptation. Before they started oversampling the minority class, they used dynamic methods like AdaBoost.NC and Synthetic Minority Over-sampling Technique (SMOTE) to train the model. Second, the concept drift was found using the Linear Four Rates (LFR) technique. Lastly, a new classifier is created to adapt to the new concept. Concept Drift Detector and Resampling Ensemble (CDRE) was proposed as a method for dealing with concept drift and multiclass imbalance in a study (Vasantha et al., 2019). In case the latest minority samples are not sufficient, the system keeps the old ones around for resampling. As a means of detecting new and missing classes, or concept drift, it compares the total number of classes present in the most recent and previous data streams. If none of the examples match the saved samples, it issues a warning. If there is concept drift, it saves those instances, and if more come later, it does ensemble classification and replaces the faulty classifier with a better one.

Keeping the importance of class imbalance and concept drift and their difficulty factors in view, this study is conducted to uncover the existing work to find the techniques used to address the joint issue of imbalanced data and concept drift in streaming data. The existing reviews are either limited to one problem only or mostly focused on concept drift and slightly discuss the class imbalance.

RELATED WORK

The class imbalance and concept drift are mainly studied separately as a result their standalone solution is proposed. In the context of online learning of non-stationary data streams where data is non-static, both issues can arrive together. Hence, in the last few years, this joint issue got intention from the research community and different solutions were proposed to address the joint issue while learning from the data stream. A survey conducted by Agrahari and Singh (2021) highlights the existing drift detection methods and their limitations. The study conducted by Lu et al. (2019) provides detailed insights about concept drift levels and the drift handling techniques for each level. The study also discusses potential research areas to work on, one among those is the issue of class imbalance along with concept drift. The systematic review conducted by Kaur et al. (2019) investigates the challenges faced by the machine learning model while learning from imbalanced data.

A critical review on the adverse effects of concept drift on machine learning was conducted by Jameel et al. (2020), the focus of their study was limited to concept drift only. Another systematic review was conducted in Kolajo et al. (2019) with a focus on big data stream analysis. Although the study was conducted on stream analysis in general, it was limited to technologies and methods used for big data stream learning and their limitations. A more recent survey was conducted by Han et al. (2022) on active and passive drift detection methods. The study further experimentally compared the performance of the drift detection methods. To the best of our knowledge, there is no other systematic review focused mainly on both the concept drift and class imbalance issues. Therefore, this study is focused to investigate how to detect concept drift in imbalance non-stationary data streams and adapt the machine learning model in presence of class imbalance. The summary of related work is presented in Table 1.

Table 1

Related Work Summary

Ref	Study Type	DS	CD	CI	Year
Kolajo et al. (2019)	Systematic review	√	x	√	2019
Lu et al. (2019)	Review	√	√	x	2019
Kaur et al. (2019)	Systematic review	x	x	√	2019
Jameel et al. (2020)	Review	√	√	x	2020
Agrahari and Singh (2021)	Review	√	√	x	2021
Han et al. (2022)	Survey	√	√	x	2022
This study	Systematic review	√	√	√	2024

DS= Data Stream, CD=Concept Drift, CI=Class Imbalance, Year=Year of Publication

It raises a fundamental question that what are effective methods for detecting and adapting to concept drift in imbalanced non-stationary data streams, considering the joint challenges of concept drift and class imbalance, and how do these methods compare in terms of performance and robustness? Therefore, this research question addresses the need to investigate methods that specifically target the simultaneous presence of concept drift and class imbalance in non-stationary data streams. It also emphasises the importance of evaluating the performance and robustness of these methods to provide insights into their effectiveness in real-world scenarios.

METHODOLOGY

The systematic literature review (SLR) undertaken in this study is based on the principles published by Keele (2007), with the objectives of summarising the existing literature, constructing a qualitative synthesis, and extracting the information.

Need for Systematic Review

Over a while, various techniques have been proposed in the literature for learning from data streams. Learning from data stream faces two main issues which are concept drift and class imbalance. Different types of drift have different effects on learning model similarly different factors of class imbalance cause also causes difficulty in concept detection and model adaptation. Most of the studies target one problem at a time. In recent years, different techniques are proposed to address the combined issue of class. This study investigates how to detect concept drift in imbalanced non-stationary data streams and adapt the machine learning model in presence of class imbalance.

Goals

The purpose of this research is to compile and analyse all high-quality and relevant prior research on learning from data streams when concept drift and class imbalance are present. In this study, we concentrate on finding publications that identify strategies for drift detection, class imbalance management, and drift adaptation. The studies designed for data stream mining with novel methodology examined. To achieve the overall goal of the survey, we investigate the latest literature to answer the following Research Questions (RQ).

- RQ1. Which methods are used to detect concept drift in imbalanced data streams?
- RQ2. Which methods are used to address the class imbalance in drifting data streams?
- RQ3. Which methods are used to adapt to the changes in data streams?

Information Sources

The research articles on the subject were collected from Scopus, ACM digital library, and Web-of-Science databases. The keyword-based search was performed on each database. The keywords were selected from the research questions. The core keywords used in our searching criteria were Data stream, non-stationary, online learning, drift, and imbalance.

Article Searching Mechanism

This section describes the process that was used to locate each article that was evaluated. The articles that were eventually published in journals were chosen from the electronic databases via the use of the conventional manual search strategy. As a result of this, the emphasis of our study is on three distinct areas, namely, data streams, drift detection, class imbalance, and drift adaptation. Therefore, conference papers were not included in our study. The rationale for this was that most of the time, conferences only consist of a small number of papers; as a result, these papers did not provide any new insights or specifics about the methodologies. Because data stream, concept drift, and class imbalance are the focus of our attention in this SLR research, our search string was designed to get the maximum number of full-length publications that were relevant to these topics as was practically feasible. We integrated several keyword combinations and other permutations by using the 'OR' and 'AND' operators in our search.

Inclusion and Exclusion Criteria

Selection of relevant literature containing fundamental studies which assist in answering the research objectives stated in the SLR, the inclusion and exclusion criteria were employed. Papers are only considered for inclusion in the present study if they meet one or more of the research goals. While searching and downloading the research articles from databases, we also applied the language filter and downloaded the literature published in English language only. The numerous inclusion and exclusion criteria that were used in this SLR investigation are shown in Table 2.

Table 2

Inclusion and Exclusion Criteria

Inclusion criteria	Exclusion criteria
Articles are written in the English language	Articles are written in a language other than English
Published in the last five years	Falls outside of the period
Address concept drift and class imbalance in data streams	Addresses only concept drift in data streams
	Addresses only class imbalance in the data stream
Journal papers only	Research publications other than journal

Data Collection

We obtained research journal articles from three widely accessed databases, namely Scopus, Web of Science, and ACM Digital Library.

1110 articles were the total result of combining the findings from all the databases. Following the removal of duplicate entries, there were 887 articles left. After going through the titles, analysing the abstracts, and conclusions, we came up with 89 papers that were relevant to our research. Thereafter, we studied the whole paper and picked 35 studies, which were read completely. We used the search technique as a foundation to utilise the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) criteria that are shown in Figure 3. The PRISMA flow diagram in Figure 3 provides information about the number of publications that were included and eliminated from this research, as well as the number of studies that were taken from published databases and extracted. The present study has outlined all the criteria that were used to determine the inclusion and exclusion of the studies. Lastly, Table 3 emphasises the overall number of publications that were found, evaluated, and ultimately determined to be appropriate for inclusion in this SLR research.

Figure 3

PRISMA Flow Diagram

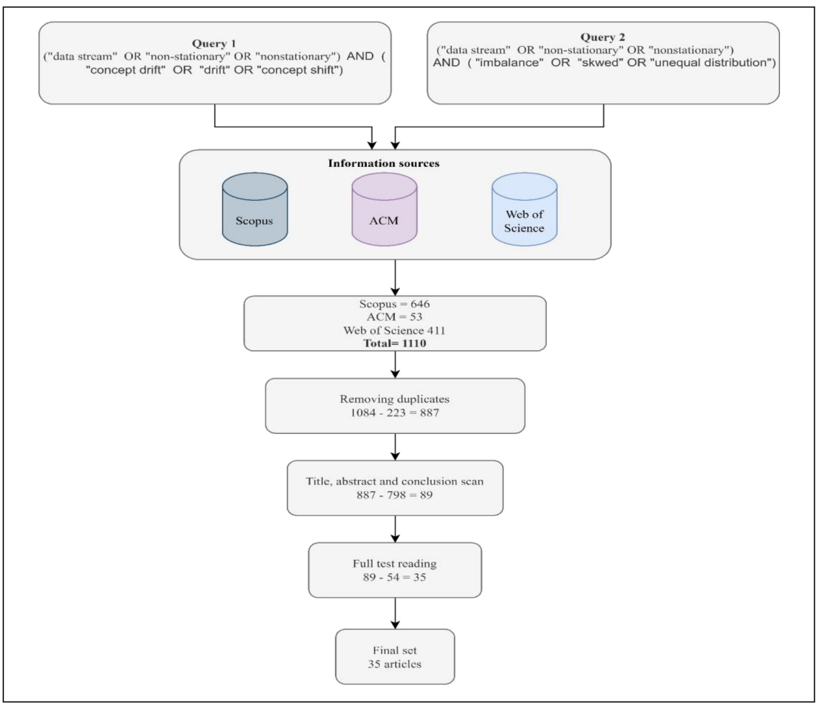


Table 3

List of final 35 Studies Selected for Analysis

Study ID	Ref	Approach for Drift Detection	Approach for Class Imbalance Handling	Approach for Concept Adaptation
S1	Khandekar and Shrinath (2022)	Error based	SMOTE	Ensemble Approach
S2	Mansour et al. (2021)	STEPD	SMOTE with Glowworm Swarm Optimization	Deep Learning – Bi-LSTM
S3	Li et al. (2021)	Data distribution based	-	Extreme Learning Machine
S4	Zhang et al. (2022)	Ensemble with time decay	Resampling with buffered minority samples	Ensemble Approach
S5	Jiao et al. (2022)	Ensemble with time decay	Adaptive nearest neighbor-SMOTE	Ensemble Approach
S6	Cano and Krawczyk (2022)	ADWIN	Separate window for minority and majority classes	Ensemble Approach
S7	Cabral and Minku (2022)	Moving the average of the output	Separate window for minority and majority classes	Classifier retraining on the latest data
S8	Zyblewski et al. (2021)	Ensemble with Dynamic Classifier Selection	Oversampling	Ensemble Approach
S9	Sun et al. (2021)	Monitor Feature Space	Cost-sensitive weights mechanism	Ensemble Approach
S10	Liu et al. (2021)	ADWIN	Higher weights on minority samples	Ensemble Approach
S11	Bernardo and Valle (2021)	ADWIN	SMOTE and SMOTE-Borderline	Single classifier, incremental training

(continued)

Study ID	Ref	Approach for Drift Detection	Approach for Class Imbalance Handling	Approach for Concept Adaptation
S12	Toor et al. (2020)	Sliding Window with Average Run Length based on a false positive outcome	-	-
S13	Toor et al. (2020)	Sliding Window	SMOTE	Updating the existing classifier
S14	Sun et al. (2020)	ADWIN- Kullback-Leibler divergence	Cost-sensitive approach	Ensemble Approach
S15	Prasad et al. (2020)	Ensemble-with Similarity range of Dimensionality	-	Ensemble Approach
S16	Lu et al. (2020)	Ensemble with time decay	Under-bagging	Ensemble Approach
S17	Li et al. (2020)	Sliding Window	SMOTE	Ensemble Approach
S18	Aydogdu and Ekinci (2020)	Output based on Drop in Accuracy	-	Extreme Learning Machine
S19	Ancy and Paulraj (2020)	Ensemble with time decay	Cost-sensitive feature selection	Ensemble Approach
S20	Zhang et al. (2019)	Sliding Window and Ensemble with time decay	Resampling with buffered minority samples	Ensemble Approach
S21	Ren et al. (2019)	Ensemble-Misclassification-based Weight Update	Minority samples' weight-based resampling	Ensemble Approach
S22	Zhang et al. (2018)	Paired Ensemble	Resampling with buffered minority samples	Ensemble Approach
S23	Ren et al. (2018)	Hybrid Ensemble	Minority selective resampling	Ensemble Approach
S24	Kulkarni et al. (2021)	Ensemble with time decay	Cost-sensitive boosting	Ensemble Approach

(continued)

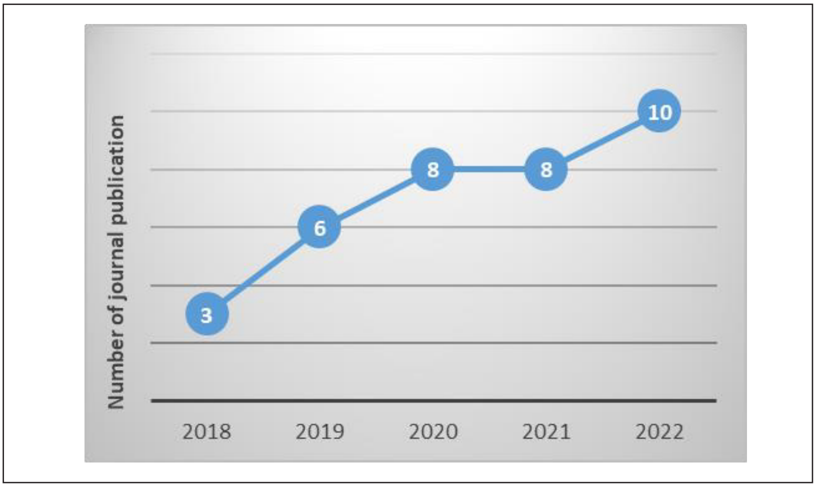
Study ID	Ref	Approach for Drift Detection	Approach for Class Imbalance Handling	Approach for Concept Adaptation
S25	Zhang and Liu (2019)	Monitors Change in Class Imbalance Ratio	Over and undersampling	Ensemble Approach
S26	Toor and Usman (2022)	CSCD	SMOTE	Ensemble Approach
S27	Hu and Jiang (2019)	Fuzzy Clustering	Imbalance Modified SMOTE	Deep Learning – Deep Neural Network
S28	Kulkarni et al. (2022)	Hybrid Ensemble	Random undersampling	Ensemble Approach
S29	Klikowski and Woźniak (2022)	Ensemble with a Weight-based Approach	Weighted-bagging	Ensemble Approach
S30	Priya and Uthra (2022)	Ensemble-based on Change in Class Imbalance Ratio	KNN minority oversampling	Ensemble Approach
S31	Dal et al. (2018)	Sliding Window	Random undersampling	Ensemble Approach
S32	Chowdhury et al. (2022)	Error based on Confusion Matrix	Distribution preserving SMOTE	Deep Learning – Deep Neural Network
S33	Roseberry et al. (2019)	Monitors Instances with High Error	Adaptive window with removing misclassified majority class instance	Single classifier- KNN
S34	Jani et al. (2019)	DDM and EDDM	SMOTE	Update the classifier
S35	Malialis et al. (2021)	Sliding Window	All minority class samples with a subset of majority class samples	Artificial Neural Network

RESULTS AND ANALYSIS

In this study, literature addressing the joint issue of concept drift and the class imbalance was examined. This study is focused on more recent work in the field and restricted to work published in the last five years i.e., from 2018 to 2022. After the systematic selection of papers, 35 studies were identified. Although S3, S12, and S18 do not address the class imbalance, still included because these studies address concept drift in streaming data. The year-wise details are given in Figure 4. Only 3 research articles for the year 2018 were found from the selected databases. 6 for 2019, 8 for 2020, and 2021, whereas 10 articles were found for the year 2022 till 15th August. It can be seen in Figure 4; the joint issue of concept drift and class imbalance is getting more attention in recent years.

Figure 4

Number of Research Articles per year



RQ 1: Which methods are used to detect concept drift in imbalanced data streams?

The existing approaches which deal with concept drift are generally divided into two type types: Active approach and Passive approach (Heusinger et al., 2022). The active approach uses an explicit drift detection mechanism to detect the concept drift, while the passive approach continuously updates the existing model regardless of the occurrence of the drift. The traditional drift detection approaches consider that data is balanced, whereas in on-stationary environment data changes frequently so as a class imbalance. Therefore, the idea behind this question is to identify the drift detection methods which are proposed for data streams having class imbalance issues along with concept drift.

Among the selected studies 34.28 percent used an ensemble approach to address the concept drift. An ensemble approach is a passive approach that mostly creates a new classifier for each new chunk of data and makes it part of the ensemble. It removes the classifier which is trained on old data using different approaches such as time decay. 25.71 percent of studies monitor the input streams to detect the drift and the most common method was the use of sliding windows. Only 14.28 percent of studies addressed the concept drift based on the error rate of the classifier. Whereas 22.86 percent of the studies used existing drift detection methods. From the existing drift detection methods, ADWIN proposed by Bifet and Gavalda (2007) was the most used method among the studies analysed in this work. Only one study, which is 2.85 percent of the studies selected for the review, used the change in class imbalance to monitor the concept drift. The techniques used for drift detection in the studies analysed in this work are listed in Figure 5. The same are discussed in detail one by one in the sub sections below.

a. Ensemble approach

The ensemble-based approach was applied to address concept drift by studies S4, S5, S16, S19, S20, S21, S22, S23, S24, and S28. In general, the ensemble approach works as a passive approach to deal with concept drift which means it does not use an explicit mechanism to detect the drift. It continuously updates the existing classifiers or trains a new classifier on the latest data to address the concept drift. Therefore, it is considered a drift adaptation technique, hence we will discuss it in more detail in the drift adaptation section.

b. Window-based approaches

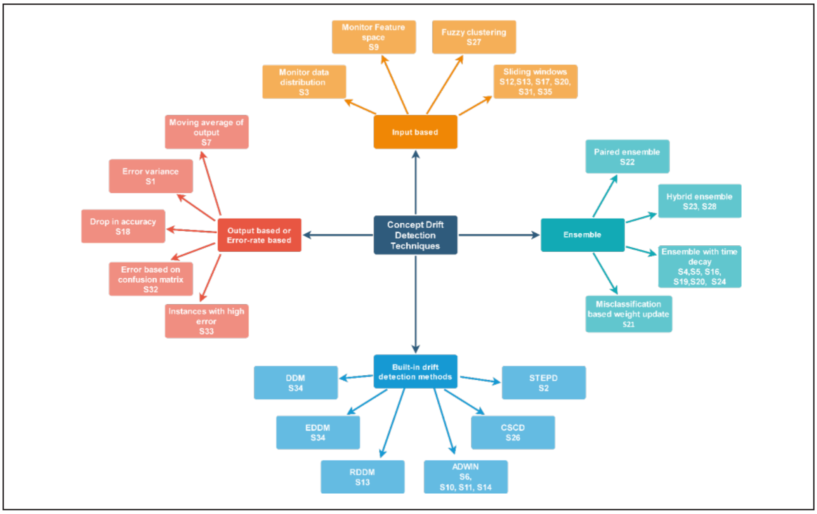
Store the latest streams for drift detection such as S12, S13, S17, S31, S20, and S35. These include a single window, and two parallel windows to track old and new data. The window can be of fixed size, or its size can be changed dynamically. The resizable or adaptive window can increase and decrease its size based on the speed of the data. The main advantage of using a window-based approach is, it helps to buffer the latest data which may represent a new concept. In case a new concept does not fit into a single window, two parallel fixed or adaptive windows are applied. The window-based approach can be useful in detecting sudden drift. The drawback of the window-based approach is, in case the data examples of the new concept are appearing slow i.e., the concept is changing gradually, this approach may not produce good results ref. Also, it requires more memory in case multiple adaptive windows or single long windows are used. The study S3 monitors the change in data distribution whereas S9 monitors the changes in feature space.

c. Output-based approaches

Monitor the performance of the classifier, either in terms of a decrease in accuracy S18 or an increase in error rate S1. The output error is monitored using different approaches like tracking the simple average of the error S32, moving the average of the error S7, or identifying the data examples with high error. In case the performance of the classifier drops (or the error of the classifier is increased) to a certain level, it is considered that it is because of the concept change. The studies S2, S6, S10, S11, S13, S14, and S34 used existing drift detection methods. The ADWIN is the most frequently used method compared to other existing drift detection methods. Another important conclusion we draw from RQ1 is that most of the existing work is focused in detecting sudden and gradual drift type. Very little focus has been given to detect the reoccurring and incremental drift. The summary of the work in terms of addressing each drift type is given in Figure 6.

Figure 5

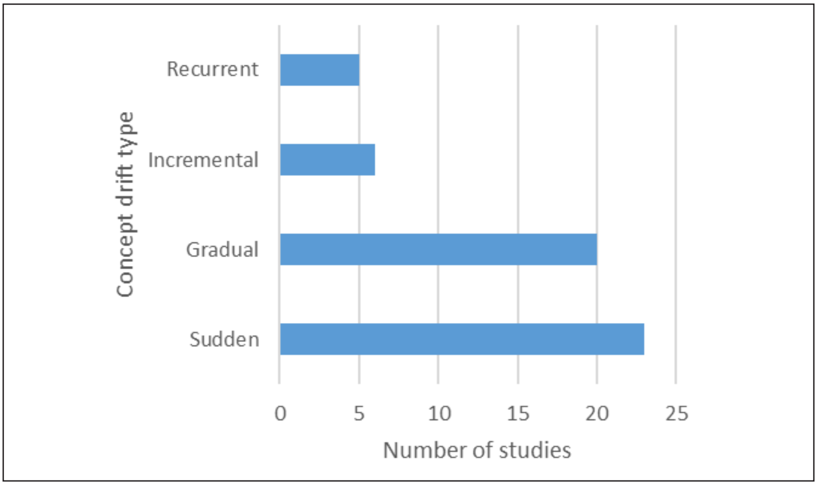
Drift Detection Methods Used in the Studies



DDM=Drift Detection Method	EDDM=Early Drift Detection Method	RDDM=Reactive Drift Detection Method
ADWIN=Adaptive Window	CSCD=Cumulative Sum Change Detector	STEPD=Statistical Test of Equal Proportions Detector

Figure 6

Studies Addressing Different Concept Drift Types



RQ 2: Which methods are used to address the class imbalance in drifting data streams?

In online learning, once the drift is identified, the model needs to be updated to adapt to the new concept. In case the data related to the new concept is imbalanced, the updated classifier will become biased towards the majority class. Therefore, a good approach is to address the class imbalance before updating the classifier. In this context, the research question aimed to identify the methods which are applied to address the class imbalance in the drifting data streams.

Among the studies analysed in this work, most of the studies used a window-based approach to store the incoming data stream in a buffer and apply different techniques such as SMOTE and or its modified version to make a balanced training set.

a. Window-based approach

The window-based approaches store recent data to monitor the concept drift. The same approaches are used to store the data related to the minority class or create a window for each class. After the identification of drift, to update the existing model or create a new classifier for the latest concept, most of the studies use the data which is available in buffer memory (window) to create a training set. In this systematic review, 10 studies specifically used a window-based approach to store the most recent data stream and then applied a weigh-based approach to balance the data such as S2, S9, and S10. In the weight-based approach, higher weights are applied to minority class samples and small weights are applied to majority class samples to address the class imbalance issue. In some studies, the window is only used to store minority class samples which will be used to create the balanced training set for model retraining when drift is identified in S4, S20, and S22. Study S33 uses a single sliding window to store data related to all classes, whereas study S35 creates two fixed-sized windows to keep samples for minority classes and majority classes separately. S35 cannot be applied for multi-class data because it is purely designed for two-class data. The S33 can work for multi-class data, but storing data related to all classes on a single window, the data may not be equally represented for every class in a single window. Because for some classes, data may arrive frequently and for others may not. Another version of the window-based approach proposed by S6 and S7 is where a separate window is created for each class to store the samples. Although it is a simple way to address the class imbalance issue but managing multiple windows required lots of resources and computations.

b. SMOTE and its variants

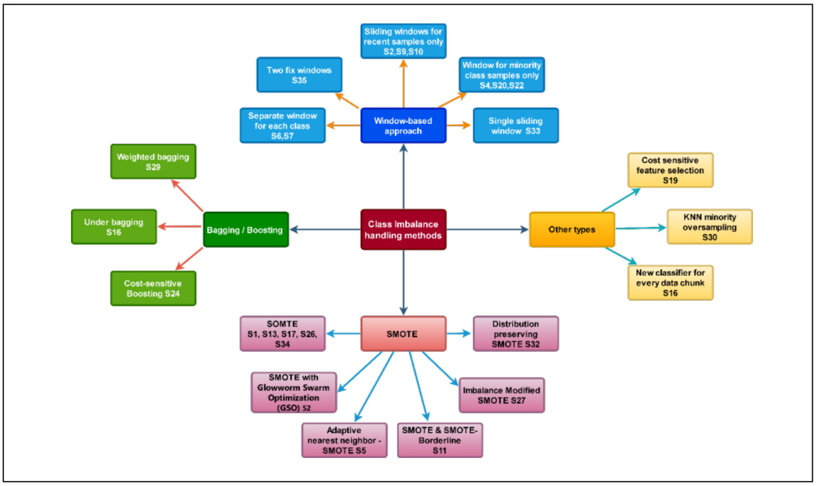
Most of the studies either used SMOTE or proposed a modified version of SMOTE to balance the data stream. Although these studies did not specifically mention the use of memory windows in their work, somehow, they are using buffer memory to keep the latest data chunk and then applying SMOTE or its modified version to oversample the minority class data. Studies S1, S13, S17, S26, and S34 directly applied SMOTE to address without any modification. S2, S5, S11, S27, and S32 proposed modified SOMTE and then applied it to generate synthetic data for the minority class.

c. Bagging and boosting

To address the class imbalance in online learning, few studies used a bagging-based approach to address the class imbalance issue by creating subsets of the data such as S16 and S29. The bagging-based approach works as an ensemble learning approach, it creates multiple subsets of the actual data and trains a separate classifier on each subset of data. The final output is selected based on a majority vote. Study S24 applies boosting approach to address the class imbalance. In both bagging and boosting approaches, the training set is divided into multiple subsets. These approaches can be good in batch processing where all data is available at the time of model training, but in an online learning environment where data is a continuous stream and can store very limited data in a buffer or window, it is hard to create multiple subsets from limited data. The techniques used for addressing the class imbalance in the studies analysed in this work are listed in Figure 7.

Figure 7

Class Imbalance Handling Methods used in the Studies



RQ 3: Which methods are used to adapt to the changes in data streams?

The drift adaptation techniques are broadly categorised into two groups that are Blind and Informed (Gama et al., 2014). The blind drift adaptation methods are also called proactive approaches. The blind approach continuously updates the model on the latest data without using any explicit method to detect drift. In the blind proactive approach, the model is updated even if there is no change in the concept. The blind drift adaptation uses a fixed-size sliding window to keep the latest data examples for retraining the model periodically, such as S12, S13, S17, S20, S31, and S35. The informed drift adaptation approaches use explicit mechanisms to detect the drift. They react only when the explicit drift detection method signal for the drift. Once the drift detection method triggers a drift signal, either the existing model is fully replaced with a newly trained model, or it is updated on the latest data but keeps the old knowledge too such as S1, S6, S7, S10, S11, S14, S3, S2, S18, S26, S32, S33. Following are the approaches which were used in the studies surveyed in this work to adapt the new concept.

a. Ensemble approach

Among the studies analysed in this work, 65.71 percent use an ensemble approach to adapt to the new concept. An ensemble-based approach is consisting of multiple trained classifiers (Arabmakki et al., 2017; Khandekar & Shrinath, 2022; Liu et al., 2021; Palli et al., 2023; Vafaie et al., 2020; Vasantha et al., 2019) when it receives a drift signal, the ensemble-based approach creates a new classifier and trains it on the latest data and adds it to the ensemble. It removes the classifier with poor performance. The common approach for discarding the classifier form ensemble is the weight-based approach. Some studies apply higher weights to the best performing classifier and reduce the weights of the classifier with poor performance such as S21. Other studies assign higher weights to the classifier which is trained on the latest data and reduce the weights of the classifier, which is trained on older data like S4, S5, S16, S19, S20, and S28. In such a situation the final output is also achieved by either a voting-based approach or the output of the classifier with the highest weights is considered.

b. Single classifier approach

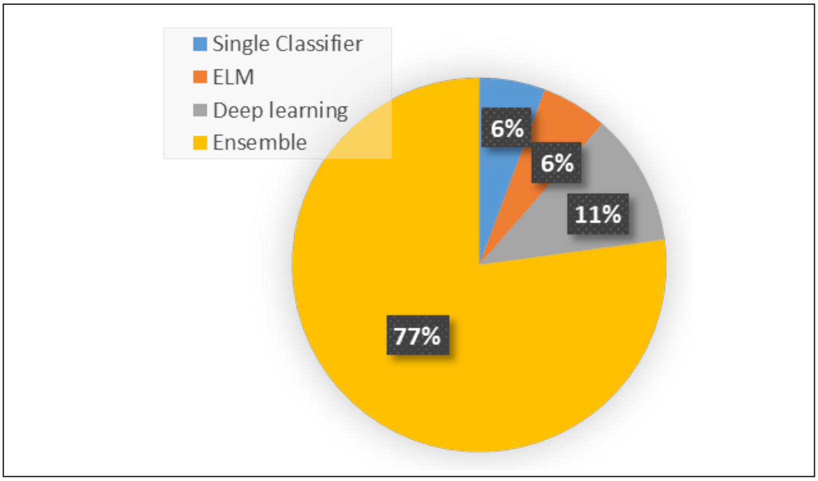
The single classifier approach retrains the existing classifier on the latest data to adapt the new concept like S11 and S33. These approaches get the knowledge on the latest data and forget the knowledge learned from old data. The single classifier-based approach can be useful in addressing the sudden. This kind of approach struggles in adapting incremental drift where new concept gradually appears, and old concept gradually disappears.

c. Deep learning

Deep learning is a subclass of machine learning that comprises neural networks (NN) with three or more layers. Deep learning does away with some of the conventionally required data pre-processing that is necessary for machine learning. These algorithms can take in and interpret unstructured data such as text and pictures, and they can also automate the extraction of features, which helps reduce the reliance on human specialists. Few of the studies analysed in this work used a deep learning-based model to adapt the concept drift. The study S35 used a fully connected Artificial Neural Network (ANN) as the classifier to learn from the data streams. Whereas S32 and S27 applied deep neural networks (DNN) for the prediction of the input streams. DNN is a multilayer feed-forward artificial neural network generally with multiple hidden layers, the model is updated on misclassified data when it receives a drift signal. The overall summary of these techniques is also presented in Figure 8.

Figure 8

Summary of Approaches used for Drift Adaptation in the Studies



DISCUSSION

In this survey, we investigated the studies which address the concept drift and class imbalance in online learning and adapt the model to the latest concept. Our focus was on identifying the techniques being applied to address the joint issue and keep the model updated. For RQ1, we saw that 28.57 percent of the studies use the ensemble approach, 25.71 percent of studies use window-based approaches, 22.86 percent of studies use existing drift detection methods and 14.28 percent of studies addressed the concept of drift based on the error rate of the classifier. Only 2.85 percent of the studies selected for the review, used the change in class imbalance to monitor the concept drift. Most of the studies are focused on the only sudden and gradual drift. It is considered that the ensemble approach can cope with most of the drift types as it trains new classifiers for the latest data and keeps the classifiers that were trained on old data for a certain time. But once the old classifiers are removed, the recurrent concept cannot be identified. The ensemble size also has not been standardised meaning that we are not sure how many classifiers should be in the ensemble. In window-based approaches, the most critical thing is the size of the window regardless of whether it is fixed or sliding. The behaviour of the data stream is always different from one application to the other.

Therefore, identifying window size is not an easy task. As in window-based approaches, concept drift is identified based on the statistical properties of the data distribution or dissimilarity of the new and old data. Although these approaches are good for the identification of virtual drift complex data streams with many features and multiple classes' dissimilarity measurement become a challenging task. The pre-processing approaches such as feature selection can be useful in such cases. The feature selection should also be adaptive because the importance of the feature may change over time. The window-based approaches also are unable to identify the real drift where data distribution remains the same but class boundary changes. The error-based approaches are good for the identification of real drift. The decrease in performance or increase in error rate monitors the classifier's performance. This is the most traditional method for drift detection. These approaches are reactive, they adapt the classifier when the performance is decreased to a certain level. These approaches cannot be suitable for applications where we have less margin of error. Another common issue with these approaches is the selection of performance measures.

There are different viewpoints on the selection of performance measures to detect the drift. S18 uses an accuracy metric to trigger drift alerts in case accuracy drops from a certain level, monitoring accuracy only cannot be a good choice in case data is imbalanced. S7 uses the moving average of output, S1 uses error variance, S33 uses instances with high error, and S32 uses a confusion matrix to detect the concept drift. There is no consensus on the selection of performance measurement metrics for drift detection. However, we believe monitoring recall and precision for each class could give better results in identifying concept drift as well as change in imbalance ratio in case data is imbalanced. It will also help in identifying drift in multiple classes.

Much of the research examined in this article employs SMOTE to solve the issue of class imbalance in non-stationary data streams by producing synthetic data samples for the minatory class. SMOTE reduces overfitting caused by random oversampling by generating synthetic examples rather than replicating data examples. SMOTE also has the benefit of not losing any information when used, unlike undersampling methods. It is well known that SMOTE typically over-generalises the minority class, resulting in misclassifications for the majority class and affecting the model's overall balance. When generating synthetic instances by taking each minority class sample and injecting synthetic examples, the SMOTE algorithm ignores the distribution of minority classes and latent noises in the data set.

The bagging and boosting techniques applied by Klikowski and Woźniak (2022) are a kind of ensemble approach. In online learning, generating multiple copies of the continuous data stream is quite challenging. The bagging and boosting-based classifiers may fail to adapt to the incremental and gradual drifts. The reason is both the concepts grow gradually, hence the data samples related to the new concept will be limited. If these limited samples of the new concept are divided into multiple subsets and a separate classifier is trained on each of the training subsets, none of the classifiers will adapt to the new concept. In such a situation it is better to use a simple ensemble approach such as training a single classifier on the latest data. The class imbalance issue should be addressed in a way that does not affect the learning ability of the classifier.

CONCLUSION

The concept drift and class imbalance are inbuilt problems of most real-time applications which generate non-stationary data streams. In this study, we performed a survey on the literature which addresses the joint problem. Our result shows that a total of 1110 research articles related to the joint issue were found on the keyword-based search. Among those only 35 studies were focused on addressing both issues. Our results show that the ensemble approach is considered the most common approach to address the challenges of non-stationary data streams. In the active learning approach, window-based approach is mostly used to detect the concept drift. Moreover, the ensemble approach is frequently used for drift adaptation once the drift is detected, because of its flexibility of adding a new classifier to adapt the new concept wherever a change is identified. Whereas in passive learning approach, where there is no external mechanism for drift detection, again ensemble approach is widely applied to continuously update the classifier using the time decay mechanism to keep the new knowledge and forget the old one. Whereas for addressing the class imbalance issue, mostly existing data balancing approaches are used such as SMOTE or its variant. The non-stationary streaming data has different nature compared to the static or stationary data, therefore traditional methods which are designed for static data cannot produce satisfactory results. Although significant work is being done in this area still there are some challenges that need to be addressed.

Challenges and Future Direction

Although there quite some studies which are proposed to address the joint issue of concept drift and class imbalance for non-stationary data stream, still there are some challenges that need to be addressed.

- In the ensemble approach, every time a new classifier is created an old classifier is removed based on the performance of the weights-based method. The calculation of weights in changing environment may mislead and cause the removal of best performing classifier.
- Some studies apply weights on data instead of classifiers, these apply higher weights on the latest data samples and decrease the weight values of older data to give importance to the latest data. In such a scenario only the time factor is considered. The latest data may not always be useful it may have noise or corrupted data. Instead of relying on only time, other factors such as identifying useful features or valuable data samples may be considered.
- Defining window size for monitoring the latest data streams is still a challenge in case minority class samples do not frequently appear. This will affect the drift detection method if drift occurs in the minority class. In such a situation class imbalance addressing method will also be affected if the current window only contains the majority of class samples.

- Error-rate-based drift detection methods monitor the performance of the classifier and trigger drift alerts in case of a drop in accuracy or an increase in error. The decrease in performance could be because of class imbalance without the concept change.
- Multi-class imbalance in presence of concept drift is still a great challenge in online data stream learning.
- Defining the size (number of classifiers) of an ensemble is also not consistent.
- Selecting relevant features for measuring the similarity and dissimilarity among data items for drift detection remains a challenging task, especially when dealing with evolving high-dimensional data streams.
- In online machine learning environment, how a classifier can learn in case the data labels are delayed is still a challenge.
- For non-stationary data stream classification, not only the classifier should be adaptive, but also the drift detector, as well as the class balancing technique, should also be adaptive so that they keep learning from the new data and updating themselves. We also believe that more attention is required to address the concept drift in the multi-class imbalance problem. It is also suggested that, while addressing the multi-class imbalance in streaming data, other class imbalance factors such as borderline data, rare examples, and outliers should also be considered.

ACKNOWLEDGEMENT

This research is supported by Universiti Teknologi PETRONAS (UTP), under the Research Graduate Scheme.

REFERENCES

Abbasi, M., Shahraki, A., & Taherkordi, A. (2021). Deep learning for network traffic monitoring and analysis (NTMA): A survey. *Computer Communications*, 170, 19-41. https://doi.org/10.1016/j.comcom.2021.01.021

Agrahari, S., & Singh, A. K. (2021). Concept Drift Detection in Data Stream Mining: A literature review. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2021.11.006

Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, 94, 290-301.

Ancy, S., & Paulraj, D. (2020). Handling imbalanced data with concept drift by applying dynamic sampling and ensemble classification model. *Computer Communications*, 153, 553-560. https://doi.org/10.1016/j.comcom.2020.01.061

Arabmakki, E., Kantardzic, M., & Sethi, T. S. (2017, 14-19 May 2017). A partial labeling framework for multi-class imbalanced streaming data. Paper presented at the 2017 International Joint Conference on Neural Networks (IJCNN).

Aydogdu, O., & Ekinci, M. (2020). A new approach for data stream classification: Unsupervised feature representational online sequential extreme learning machine. *Multimedia Tools and Applications*, *79*(37-38), 27205-27227. https://doi.org/10.1007/s11042-020-09300-y

Bahri, M., Bifet, A., Gama, J., Gomes, H. M., & Maniu, S. (2021). Data stream analysis: Foundations, major tasks and tools. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 11*(3), e1405. https://doi.org/10.1002/widm.1405

Bernardo, A., & Della Valle, E. (2021). VFC-SMOTE: Very fast continuous synthetic minority oversampling for evolving data streams. *Data Mining and Knowledge Discovery, 35*(6), 2679-2713. https://doi.org/10.1007/s10618-021-00786-0

Bifet, A., & Gavalda, R. (2007). Learning from time-changing data with adaptive windowing. Paper presented at the Proceedings of the 2007 SIAM international conference on data mining.

Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modeling on imbalanced domains. *ACM Computing Surveys* (CSUR), *49*(2), 1-50.

Brzezinski, D., Minku, L. L., Pewinski, T., Stefanowski, J., & Szumaczuk, A. (2021). The impact of data difficulty factors on classification of imbalanced and concept drifting data streams. *Knowledge and Information Systems, 63*(6), 1429-1469. https://doi.org/10.1007/s10115-021-01560-w

Cabral, G. G., & Minku, L. L. (2022). Towards Reliable Online Just-in-time Software Defect Prediction. IEEE Transactions on Software Engineering, 1-1. 10.1109/TSE.2022.3175789

Cano, A., & Krawczyk, B. (2022). ROSE: Robust online self-adjusting ensemble for continual learning on imbalanced drifting data streams. *Machine Learning, 111*(7), 2561-2599. https://doi.org/10.1007/s10994-022-06168-x

Chowdhury, A., Raut, S., & Pal, A. (2022). Internet of Things resource monitoring through proactive fault prediction. *Computers and Industrial Engineering*, 169. https://doi.org/10.1016/j.cie.2022.108265

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE transactions on neural networks and learning systems*, *29*(8), 3784-3797. https://doi.org/10.1109/TNNLS.2017.2736643

Fauzi, M. F. A. M., Aziz, I. A., & Amiruddin, A. (2019). The Prediction of Remaining Useful Life (RUL) in Oil and Gas Industry using Artificial Neural Network (ANN) Algorithm. Paper presented at the 2019 IEEE Conference on Big Data and Analytics (ICBDA).

Fazal, K., & Alotaibi, S. (2020). Applications of artificial intelligence and big data analytics in m-health: A healthcare system perspective. *Journal of Healthcare Engineering*, 2020, https://doi.org/10.1155/2020/8894694

Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys, 46*(4). https://doi.org/10.1145/2523813

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys* (CSUR), *46*(4), 1-37.

Ghazikhani, A., Monsefi, R., & Yazdi, H. S. (2013). Ensemble of online neural networks for non-stationary and imbalanced data streams. *Neurocomputing*, 122, 535-544.

Han, M., Chen, Z., Li, M., Wu, H., & Zhang, X. (2022). A survey of active and passive concept drift handling methods. *Computational Intelligence. 38*(4), 1492-1535. https://doi.org/10.1111/coin.12520

Heusinger, M., Raab, C., & Schleif, F.-M. (2022). Passive concept drift handling via variations of learning vector quantisation. *Neural Computing and Applications, 34*(1), 89-100. https://doi.org/10.1007/s00521-020-05242-6

Hu, Z. X., & Jiang, P. (2019). An Imbalance Modified Deep Neural Network With Dynamical Incremental Learning for Chemical Fault Diagnosis. *IEEE Transactions on Industrial Electronics, 66*(1), 540-550. https://doi.org/10.1109/tie.2018.2798633

Jameel, S. M., Hashmani, M. A., Alhussain, H., Rehman, M., & Budiman, A. (2020). A critical review on adverse effects of concept drift over machine learning classification models. *International Journal of Advanced Computer Science and Applications* (IJACSA), *11*(1).

Jani, R., Bhatt, N., & Shah, C. (2019). Handling concept drift in data stream classification. *International Journal of Innovative Technology and Exploring Engineering, 8*(10), 548-550. https://doi.org/10.35940/ijitee.J8857.0881019

Jayasinghe, L., Samarasinghe, T., Yuen, C., Low, J. C. N., & Ge, S. S. (2018). Temporal convolutional memory networks for remaining useful life estimation of industrial machinery. In *2019 IEEE International Conference on Industrial Technology* (ICIT) (pp. 915-920).

Jiao, B., Guo, Y., Gong, D., & Chen, Q. (2022). Dynamic ensemble selection for imbalanced data streams with concept drift. *IEEE transactions on neural networks and learning systems*, 1-14. https://doi.org/10.1109/TNNLS.2022.3183120

Karbasi, S. (2020). Fast and accurate concept drift detection from event logs. *Journal of Information and Communication Technology, 43*(43), 105.

Kaur, H., Pannu, H. S., & Malhi, A. K. (2019). A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Computing Surveys, 52*(4). https://doi.org/10.1145/3343440

Keele, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Retrieved from https://www.researchgate.net/profile/Barbara-Kitchenham/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering/links/61712932766c4a211c03a6f7/Guidelines-for-performing-Systematic-Literature-Reviews-in-Software-Engineering.pdf

Khandekar, V. S., & Shrinath, P. (2022). Hybrid dynamic chunk ensemble model for multi-class data streams. *Indonesian Journal of Electrical Engineering and Computer Science*, *25*(2), 1115-1122. https://doi.org/10.11591/ijeecs.v25.i2.pp1115-1122

Khandekar, V. S., & Srinath, P. (2020). Non-stationary Data Stream Analysis: State-of-the-Art Challenges and Solutions. In *Proceeding of International Conference on Computational Science and Applications: ICCSA 2019* (pp. 67-80). Springer Singapore.

Klikowski, J., & Woźniak, M. (2022). Deterministic Sampling Classifier with weighted Bagging for drifted imbalanced data stream classification. *Applied Soft Computing*, 122. https://doi.org/10.1016/j.asoc.2022.108855

Kolajo, T., Daramola, O., & Adebiyi, A. (2019). Big data stream analysis: A systematic literature review. *Journal of Big Data*, *6*(1), 47.

Korycki, L., & Krawczyk, B. (2021). Concept drift detection from multi-class imbalanced data streams. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)* (pp. 1068-1079). IEEE.

Krawczyk, B., & Cano, A. (2018). Online ensemble learning with abstaining classifiers for drifting and noisy data streams. *Applied Soft Computing, 68*, 677-692.

Kulkarni, R., Revathy, S., & Patil, S. (2021). A Novel Approach to Maximise G-mean in Nonstationary Data with Recurrent Imbalance Shifts. *International Arab Journal of Information Technology, 18*(1), 103-113. https://doi.org/10.34028/iajit/18/1/12

Kulkarni, R. V., Revathy, S., & Patil, S. H. (2022). Smart pools of data with ensembles for adaptive learning in dynamic data streams with class imbalance. *IAES International Journal of Artificial Intelligence, 11*(1), 310-318. https://doi.org/10.11591/ijai.v11.i1.pp310-318

Kumar, R. L., Kakarla, J., Isunuri, B. V., & Singh, M. (2021). Multi-class brain tumor classification using residual network and global average pooling. *Multimedia Tools and Applications*, 80, 13429-13438.

Li, Q. D., Xiong, Q. Y., Ji, S. F., Yu, Y., Wu, C., & Gao, M. (2021). Incremental semi-supervised Extreme Learning Machine for Mixed data stream classification. *Expert Systems with Applications,* 185. https://doi.org/10.1016/j.eswa.2021.115591

Li, X., Ding, Q., & Sun, J.-Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering & System Safety*, 172, 1-11.

Li, Z., Huang, W. C., Xiong, Y., Ren, S. Q., & Zhu, T. F. (2020). Incremental learning imbalanced data streams with concept drift: The dynamic updated ensemble algorithm. *Knowledge-Based Systems,* 195. https://doi.org/10.1016/j.knosys.2020.105694

Lin, C.-C., Deng, D.-J., Kuo, C.-H., & Chen, L. (2019). Concept drift detection and adaption in big imbalance industrial IoT data using an ensemble learning method of offline classifiers. *IEEE Access*, 7, 56198-56207.

Liu, W., Zhang, H., Ding, Z., Liu, Q., & Zhu, C. (2021). A comprehensive active learning method for multiclass imbalanced data streams with concept drift. *Knowledge-Based Systems*, 215, 106778. https://doi.org/10.1016/j.knosys.2021.106778

Liu, W. K., Zhang, H., Ding, Z. Y., Liu, Q. B., & Zhu, C. (2021). A comprehensive active learning method for multiclass imbalanced data streams with concept drift. *Knowledge-Based Systems,* 215. https://doi.org/10.1016/j.knosys.2021.106778

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2019). Learning under Concept Drift: A Review. *IEEE Transactions on Knowledge and Data Engineering, 31*(12), 2346-2363. https://doi.org/10.1109/TKDE.2018.2876857

Lu, Y., Cheung, Y. M., & Yan Tang, Y. (2020). Adaptive chunk-based dynamic weighted majority for imbalanced data streams with concept drift. *IEEE transactions on neural networks and learning systems, 31*(8), 2764-2778. https://doi.org/10.1109/TNNLS.2019.2951814

Malialis, K., Panayiotou, C. G., & Polycarpou, M. M. (2021). Online Learning with Adaptive Rebalancing in Nonstationary Environments. *IEEE transactions on neural networks and learning systems, 32*(10), 4445-4459. https://doi.org/10.1109/TNNLS.2020.3017863

Mansour, R. F., Al-Otaibi, S., Al-Rasheed, A., Aljuaid, H., Pustokhina, I. V., & Pustokhin, D. A. (2021). An optimal big data analytics with concept drift detection on high-dimensional streaming data. *Computers, Materials and Continua, 68*(3), 2843-2858. https://doi.org/10.32604/cmc.2021.016626

Margara, A., & Rabl, T. (2019). Definition of Data Streams. In S. Sakr & A. Y. Zomaya (Eds.), Encyclopedia of Big Data Technologies (pp. 648-652). Cham: Springer International Publishing.

Meseguer, J., Puig, V., & Escobet, T. (2010). Fault diagnosis using a timed discrete-event approach based on interval observers: Application to sewer networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 40*(5), 900-916.

Metz, D. (2018). Tackling urban traffic congestion: The experience of London, Stockholm and Singapore. *Case Studies on Transport Policy, 6*(4), 494-498.

Palli, A.-S., Jaafar, J., Hashmani, M.-A., Gomes, H.-M., Alsughayyir, A., & Gilal, A.-R. (2023). *Combined Effect of Concept Drift and Class Imbalance on Model Performance During Stream Classification, 75*(1), 1827-1845.

Palli, A. S., Jaafar, J., Gomes, H. M., Hashmani, M. A., & Gilal, A. R. (2022). An Experimental Analysis of Drift Detection Methods on Multi-Class Imbalanced Data Streams. *Applied Sciences, 12*(22), 11688. https://doi.org/10.3390/app122211688

Palli, A. S., Jaafar, J., & Hashmani, M. A. (2021). Performance Degradation of Multi-class Classification Model Due to Continuous Evolving Data Streams. In: Vol. 72. Lecture Notes on Data Engineering and Communications Technologies (pp. 696-706): Springer Science and Business Media Deutschland GmbH.

Palli, A. S., Jaafar, J., Hashmani, M. A., Gomes, H. M., & Gilal, A. R. (2022). A Hybrid Sampling Approach for Imbalanced Binary and Multi-Class Data Using Clustering Analysis. *IEEE Access*, 10, 118639-118653. https://doi.org/10.1109/ACCESS.2022.3218463

Pongpaichet, S., Jankapor, S., Janchai, S., & Tongsanit, T. (2020). Early detection at-risk students using machine learning. In *2020 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 283-287)*. IEEE.

Prasad, K. S. N., Rao, A. S., & Ramana, A. V. (2020). Ensemble framework for concept-drift detection in multidimensional streaming data. *International Journal of Computers and Applications, 44*(12), 1193-1200. https://doi.org/10.1080/1206212X.2020.1711617

Priya, S., & Uthra, A. (2022). Adaptive Synthetic Oversampling Algorithm for Handling Class Imbalance in Multi-Class Data Stream Classification. *Journal of Computer Science, 18*(7), 650-664. https://doi.org/10.3844/jcssp.2022.650.664

Ren, S., Liao, B., Zhu, W., Li, Z., Liu, W., & Li, K. (2018). The Gradual Resampling Ensemble for mining imbalanced data streams with concept drift. Neurocomputing, 286, 150-166. https://doi.org/10.1016/j.neucom.2018.01.063

Ren, S. Q., Zhu, W., Liao, B., Li, Z., Wang, P., Li, K. Q., . . . Li, Z. J. (2019). Selection-based resampling ensemble algorithm for nonstationary imbalanced stream data learning. *Knowledge-Based Systems*, 163, 705-722. https://doi.org/10.1016/j.knosys.2018.09.032

Roseberry, M., Krawczyk, B., & Cano, A. (2019). Multi-label punitive kNN with self-adjusting memory for drifting data streams. *ACM Transactions on Knowledge Discovery from Data, 13*(6). https://doi.org/10.1145/3363573

Sainin, M. S., Alfred, R., & Ahmad, F. (2021). Ensemble meta classifier with sampling and feature selection for data with imbalance multiclass problem. *Journal of Information and Communication Technology, 20*(2), 103-133.

Soares, R. G. F., & Minku, L. L. (2023). OSNN: An Online Semisupervised Neural Network for Nonstationary Data Streams. *IEEE transactions on neural networks and learning systems, 34*(9), 6029-6041. https://doi.org/10.1109/TNNLS.2021.3132584

Spelmen, V. S., & Porkodi, R. (2018). A review on handling imbalanced data. In *2018 international conference on current trends towards converging technologies (ICCTCT) (pp. 1-11)*. IEEE.

Sun, Y., Sun, Y., & Dai, H. (2020). Two-stage cost-sensitive learning for data streams with concept drift and class imbalance. *IEEE Access*, 8, 191942-191955. https://doi.org/10.1109/ACCESS.2020.3031603

Sun, Y. G., Li, M., Li, L., Shao, H., & Sun, Y. (2021). Cost-Sensitive Classification for Evolving Data Streams with Concept Drift and Class Imbalance. *Comput Intell Neurosci*, 2021. https://doi.org/10.1155/2021/8813806

Taha, A. Y., Tiun, S., Abd Rahman, A. H., & Sabah, A. (2021). Multilabel over-sampling and under-sampling with class alignment for imbalanced multilabel text classification. *Journal of Information and Communication Technology, 20*(3), 423-456.

Toor, A. A., & Usman, M. (2022). Adaptive telecom churn prediction for concept-sensitive imbalance data streams. *Journal of Supercomputing, 78*(3), 3746-3774. https://doi.org/10.1007/s11227-021-04021-x

Toor, A. A., Usman, M., Younas, F., Fong, A. C. M., Khan, S. A., & Fong, S. (2020). Mining Massive E-Health Data Streams for IoMT Enabled Healthcare Systems. Sensors (Basel), *20*(7). https://doi.org/10.3390/s20072131

Toor, A. A., Usman, M., Younas, F., Habib, M., & Fong, A. C. M. (2020). Efficient mining of IoT based data streams for advanced computer vision systems. *Multimedia Tools and Applications*, 1-16. https://doi.org/10.1007/s11042-020-09175-z

Vafaie, P., Viktor, H., & Michalowski, W. (2020). Multi-class imbalanced semi-supervised learning from streams through online ensembles. In *2020 International Conference on Data Mining Workshops (ICDMW) (pp. 867-874)*. IEEE

Vasantha Kokilam, K., Ponmary Pushpa Latha, D., & Joseph Pushpa Raj, D. (2019). Learning of concept drift and multi class imbalanced dataset using resampling ensemble methods. *International Journal of Recent Technology and Engineering, 8*(1), 1332-1340.

Wang, T., Zhao, S., Zhu, G., & Zheng, H. J. A. e. (2021). A machine learning-based early warning system for systemic banking crises, *53*(26), 2974-2992.

Zhang, G., Yang, D., Galanis, G., & Androulakis. (2022). Solar forecasting with hourly updated numerical weather prediction. *Renewable and Sustainable Energy Reviews*, 154, 111768.

Zhang, H., & Liu, Q. B. (2019). Online Learning Method for Drift and Imbalance Problem in Client Credit Assessment. *Symmetry, 11*(7), 890. https://doi.org/10.3390/sym11070890

Zhang, H., Liu, W., & Liu, Q. (2022). Reinforcement Online Active Learning Ensemble for Drifting Imbalanced Data Streams. *IEEE Transactions on Knowledge and Data Engineering, 34*(8), 3971-3983. https://doi.org/10.1109/TKDE.2020.3026196

Zhang, H., Liu, W., Shan, J., & Liu, Q. (2018). Online Active Learning Paired Ensemble for Concept Drift and Class Imbalance. *IEEE Access*, 6, 73815-73828. https://doi.org/10.1109/ACCESS.2018.2882872

Zhang, H., Liu, W., Wang, S., Shan, J., & Liu, Q. (2019). Resample-based ensemble framework for drifting imbalanced data streams. *IEEE Access*, 7, 65103-65115. https://doi.org/10.1109/ACCESS.2019.2914725

Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation, 5*(1), 1-20.

Zyblewski, P., Sabourin, R., & Woźniak, M. (2021). Preprocessed dynamic classifier ensemble selection for highly imbalanced drifted data streams. *Information Fusion*, 66, 138-154. https://doi.org/10.1016/j.inffus.2020.09.004