

Evolving Strategies in Machine Learning: A Systematic Review of Concept Drift Detection

Gurgen Hovakimyan ^{1,*} and Jorge Miguel Bravo ^{1,2,3,4,5}

¹ NOVA IMS—Information Management School, New University of Lisbon, 1070-312 Lisbon, Portugal; jbravo@novaims.unl.pt

² Department of Economics, Université Paris-Dauphine PSL, 75775 Paris, France

³ Centro de Investigação em Gestão de Informação (MagIC), 1070-312 Lisbon, Portugal

⁴ BRU-ISCTE-IUL, 1649-026 Lisbon, Portugal

⁵ CEFAGE-UE, 7000-809 Évora, Portugal

* Correspondence: 20231150@novaims.unl.pt

† These authors contributed equally to this work.

Abstract: In this comprehensive literature review, we rigorously adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for our process and reporting. This review employs an innovative method integrating the advanced natural language processing model T5 (Text-to-Text Transfer Transformer) to enhance the accuracy and efficiency of screening and data extraction processes. We assess strategies for handling the concept drift in machine learning using high-impact publications from notable databases that were made accessible via the IEEE and Science Direct APIs. The chronological analysis covering the past two decades provides a historical perspective on methodological advancements, recognizing their strengths and weaknesses through citation metrics and rankings. This review aims to trace the growth and evolution of concept drift mitigation strategies and to provide a valuable resource that guides future research and deepens our understanding of this rapidly changing field. Key findings highlight the effectiveness of diverse methodologies such as drift detection methods, window-based methods, unsupervised statistical methods, and neural network techniques. However, challenges remain, particularly with imbalanced data, computational efficiency, and the application of concept drift detection to non-tabular data like images. This review aims to trace the growth and evolution of concept drift mitigation strategies and provide a valuable resource that guides future research and deepens our understanding of this rapidly changing field.

Keywords: concept drift; systematic review; machine learning; types of concept drift; adaptive strategies; Science Direct API; IEEE API; streaming data; non-stationary environments; evolving data streams



Citation: Hovakimyan, G.; Bravo, J.M. Evolving Strategies in Machine Learning: A Systematic Review of Concept Drift Detection. *Information* **2024**, *15*, 786. <https://doi.org/10.3390/info15120786>

Academic Editors: Sunil Jha, Malgorzata Rataj and Xiaorui Zhang

Received: 15 October 2024

Revised: 21 November 2024

Accepted: 4 December 2024

Published: 7 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Rationale for a Literature Review on Concept Drift Detection

1.1. Introduction

In the era of big data and continuous information streams, machine learning models are widely employed to predict, classify, and analyze data in various domains, such as finance, healthcare, the semiconductor industry, and customer service. These models typically assume that the underlying data distribution remains static over time. However, this assumption often fails because of concept drift, where changes in the statistical properties of the target variable over time can degrade model performance. Concept drift research focuses on methodologies for detecting, understanding, and adapting to these changes, aiming to keep models accurate and reliable [1]. Traditional classifiers may struggle in such environments, leading to poor performance [2] as the fundamental patterns in training data evolve, causing model degradation and reduced accuracy over time [3].

1.2. Importance of Concept Drift Detection

Concept drift, characterized by changes in data patterns over time, necessitates continuous advancements in machine learning, especially in classification and regression tasks. Designing models that can adapt to these changes is crucial for maintaining performance [4]. The relevance of this research is underscored by its direct impact on model accuracy and decision-making processes.

The real-world implications of concept drift are evident across various sectors. For example, in fraud detection, systems that monitor credit card transactions must adapt to evolving customer behaviors and class imbalances, where genuine transactions vastly outnumber fraudulent ones [5]. Financial time series in stock market prediction are prone to concept drift, affecting forecasting accuracy as models become outdated [6].

In the semiconductor industry, monitoring the health of expensive equipment is critical. Traditional methods like statistical process control (SPC) often fail to detect drifts early enough for preventive action [7]. Similarly, healthcare monitoring involves tracking patient data for the early detection of health condition changes, which is crucial during pandemics like COVID-19 [8].

Other examples include recommendation systems in e-commerce, such as those used by Amazon and Netflix, which must adapt to changing user preferences to remain effective [9], and image classification, where long-term deployments face shifting visual environments and detecting concept drift is vital for maintaining model accuracy [10].

These examples highlight the pervasive and critical nature of concept drift across multiple domains, emphasizing the need for continuous research and development in drift detection methodologies.

1.3. Objectives of the Systematic Literature Review

The primary goal of this literature review is to illuminate the landscape of existing methods employed in detecting concept drift. We aim to comprehensively categorize the different detection techniques to understand the available tools. This effort involves classifying methods based on criteria such as their underlying algorithms, application domains, and performance metrics. Additionally, we endeavor to discern each method's strengths and weaknesses, identify the most effective strategies, and understand their limitations. This analysis will guide practitioners in selecting appropriate methods.

Furthermore, we aim to pinpoint and analyze emerging trends and relatively less explored topics for future research. We highlight these areas by focusing on promising avenues that warrant further investigation. An integral element of our endeavor is to update the information in existing literature review papers, serving as a timely and relevant resource in this fast-paced domain. By achieving these objectives, we intend to contribute a detailed and actionable synthesis of current knowledge in concept drift detection, facilitating further advancements in this field.

1.4. Scope of the Review

This review comprehensively examines various concept drift detection methods developed over the last two decades. It includes an in-depth analysis of statistical techniques, machine learning approaches, and deep learning methods. We also discuss the datasets, and their characteristics, widely used in concept drift detection. Furthermore, we evaluate the frameworks and tools commonly employed in this field, providing a holistic understanding of the landscape. This review aims to offer a thorough and up-to-date synthesis of the methodologies and resources pertinent to concept drift detection by covering these areas.

1.5. Theoretical Foundations of Concept Drift

Concept drift can be understood through the lens of well-established theoretical frameworks, including statistical learning theory, Bayesian inference, information theory, and online learning theory. These frameworks provide a structured basis for understanding the challenges and methodologies associated with drift detection.

1. Statistical Learning Theory

Statistical learning theory underpins the generalization capability of machine learning models. This theory assumes that the joint probability distribution $P(X, Y)$ remains static over time. However, concept drift violates this assumption, leading to degraded model performance. When $P(X)$ or $P(Y|X)$ changes, the empirical risk minimization principle no longer holds. Adaptive techniques are necessary to address these violations and restore model performance [11].

2. Bayesian Inference Bayesian inference provides a probabilistic framework for updating model beliefs as new data become available. Concept drift can be viewed as a continuous process of updating prior distributions to adapt to evolving evidence [12].

Example: In financial forecasting, Bayesian models dynamically update predictions to reflect changes in market conditions, ensuring more accurate risk assessments.

3. Information Theory Information-theoretic measures, such as entropy and Kullback–Leibler (KL) divergence, are commonly used to quantify changes in data distributions [13].

Example: KL divergence can be used to compare the statistical properties of incoming data streams with historical data, flagging significant deviations as potential drift.

4. Online Learning Theory

Online learning theory deals with incremental model updates as new data arrives. This framework emphasizes balancing stability (preserving past knowledge) with plasticity (adapting to new patterns) [14].

Example: Online learning models used in recommendation systems can adapt to changing user preferences without requiring a full retraining of the system.

Practical Implications The integration of these theoretical frameworks into concept drift detection methods enhances their adaptability and effectiveness:

- Statistical learning theory highlights the necessity of adaptive models to address changing distributions.
- Bayesian inference provides a natural mechanism for gradual drift adaptation.
- Information-theoretic measures enable precise quantification of virtual drift.
- Change detection theory offers robust tools for identifying abrupt changes.
- Online learning frameworks ensure scalability and real-time adaptability.

By grounding concept drift detection in these foundational theories, researchers can develop robust, adaptable models tailored to the complexities of dynamic environments.

2. Methodology

This literature review followed the PRISMA guidelines, ensuring a comprehensive and unbiased approach. The methodological process was structured within four crucial stages: the identification, screening, eligibility, and inclusion of studies.

2.1. Identification of Research

The initial research was conducted on two primary databases, IEEE and Science Direct, utilizing their inherent Application Programming Interfaces (APIs) and Python SDK. This led to the identification of 450 potential studies. Our search encompassed a comprehensive overview of all relevant topics, with search queries including terms such as the following:

- Concept drift;
- Change detection adaptive environment;
- Evolving data streams;
- Unstable environment;
- Drift detection;
- Distribution change;
- Online learning;
- Non-stationary environments.

As we delved deeper into the literature during our review process, we found additional key papers in the references sections of various articles. These relevant publications, curated

from platforms like Springer, ResearchGate, and ACM Digital Library, contributed an extra 40 studies. Therefore, our review process expanded, ultimately examining 490 studies, enhancing the comprehensiveness of our research base.

2.2. Inclusion and Exclusion Criteria

To ensure the relevance and quality of the collected literature, we established strict inclusion and exclusion criteria. Papers to be included had to address concept drift detection methodologies explicitly, provide empirical evaluations, and be published in reputable journals or conferences. Conversely, we excluded papers that focused solely on general machine learning concepts without specific relevance to concept drift detection.

From our initial pool of 490 studies, these criteria allowed us to sift down to 356 papers that satisfied all conditions. To further underscore the quality of our chosen literature, we conducted a citation analysis to identify the most influential articles on concept drift detection. This was calculated as the number of citations divided by the number of years since the article's publication [15], as given by the following formula:

$$\text{Adjusted Citation Rate} = \frac{\text{Total Number of Citations}}{\text{Years Since Publication}}.$$

2.3. Screening Process

Upon gathering the articles, we utilized the sophisticated Text-to-Text Transfer Transformer (T5) model [16] to summarize abstracts for relevance. This innovative natural language processing (NLP)-aided screening process was applied to the previously identified 356 studies. The aim was to identify articles that adequately met our inclusion criteria, which required the works to be peer-reviewed, written in English, and published in either research journals or conference proceedings. Each study needed to demonstrate applications or simulations related to concept drift and implement or evaluate techniques for detecting concept drift. Furthermore, the outcomes of the studies had to involve measuring or observing the effectiveness of the concept drift detection techniques. Adhering to these criteria, we narrowed our pool to 254 articles. This rigorous screening process yielded a final selection of articles that was efficient and accurate.

2.4. Eligibility Assessment

The eligibility assessment phase aimed to ensure that the included studies adhered strictly to the pre-established quality and relevance criteria. From the 254 studies that passed the initial screening, we conducted a meticulous full-text review of each paper to ascertain its compliance with our inclusion criteria. This process emphasized the relevance of each paper to our research questions, the scope of the investigation, and its alignment with the objectives of this review.

During this phase, we delved deeper into the studies' methods, results, and conclusions. We performed this careful examination to include only high-quality and relevant studies in our review. We verified that each study utilized concept drift detection techniques and specifically applied them to the central topics of our review. Studies that used these techniques for unrelated tasks were excluded to maintain the focus and relevance of our review.

This diligent step refined our focus, resulting in a concentrated pool of 111 studies. The thoroughness employed at this stage underscores the trustworthiness and reliability of our review, maintaining the resonance and quality of the included research.

2.5. Quality Assessment

The objective of the quality assessment was to evaluate methodological rigor and potential sources of bias in the included studies, ensuring reliable and valid findings. We used an adapted Newcastle–Ottawa Scale (NOS) for methodological studies [17] and CASP (Critical Appraisal Skills Programme)-like checklists for empirical and application studies [18].

The key criteria used for the assessment included methodological clarity, empirical validation, reproducibility, and practical relevance. Methodological clarity involved assessing whether the concept drift problem was clearly defined and whether methods and algorithms were described in detail. Empirical validation included evaluating the relevance and quality of datasets; the clarity and relevance of performance metrics, such as accuracy and detection delay; and whether methods were compared with existing baseline methods. We assessed reproducibility by checking the availability of data, code, and steps to reproduce the study for replication and the transparency of the study's methodology and reporting. Finally, we determined the practical relevance based on whether the methods were applicable to real-world scenarios and the overall impact and contribution to the field. Each study was rated on a scale from 1 to 5 for each criterion, and the overall quality scores were calculated.

The studies considered high quality (a score of 4–5) had solid foundations with clear problem statements, detailed and rigorous methodologies, robust validations, and high transparency. Those of moderate quality (3–4) provided clear methodologies but had some limitations in data availability or comparative analysis—they offered useful insights but required careful interpretation. Those of low quality (below 3) had unclear problem statements and insufficient methodological detail—they had methodological weaknesses and were considered with caution. Table 1 lists the aggregated quality results.

Table 1. Summary of the quality assessment scores.

Quality Category	Score Range	Number of Studies	Percentage of Total
High	4–5	51	45%
Moderate	3–4	45	41%
Low	Below 3	15	14%
Total		111	100%

By systematically assessing the study quality and accounting for biases, this quality assessment enhanced the reliability of our review on concept drift detection.

2.6. Synthesis of Results

The included studies varied in study design, drift types, data types, and methods. Key characteristics included the following:

- Study Designs—supervised, semi-supervised, and unsupervised.
- Drift Types—pattern-based and distribution-based.
- Data Types—synthetic and real.
- Methods—drift detection mechanisms, window-based mechanisms, unsupervised and semi-supervised methods, ensemble methods, and neural networks.

The studies were categorized based on the type of concept drift and the methods used for detecting concept drift. Examining existing studies from the last two decades revealed that concept drift can be grouped into two primary categories: distribution- and pattern-based.

Distribution-based concept drift concerns changes in the statistical properties of data over time, impacting machine learning model performance. These alterations may be sudden, incremental, gradual, or recurrent. Conversely, pattern-based drift reflects changes within data relationships and patterns, which can involve modifications to associations between features, decision boundaries, or input and target variables.

Figures 1 and 2 illustrate these categories further. Notably, distribution-based drift can appear as virtual concept drift, real concept drift, or via the introduction of a novel class. Virtual drift involves changes in input feature distributions without impacting the target variable. Real concept drift, which is impactful across healthcare, economics, and financial markets, involves changes to the conditional probability $p(Y|X)$. Lastly, novel class appearance introduces new, previously unseen classes into the data stream.

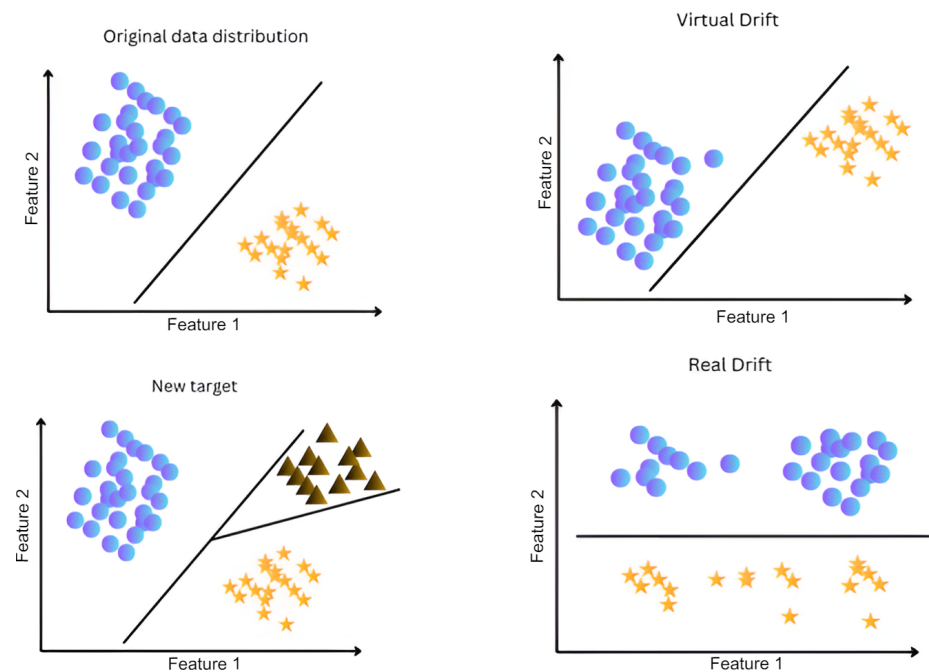


Figure 1. Distribution-based concept drift: The figure shows various concept drift scenarios, where different shapes represent different classes and changes in data distribution and class relationships.

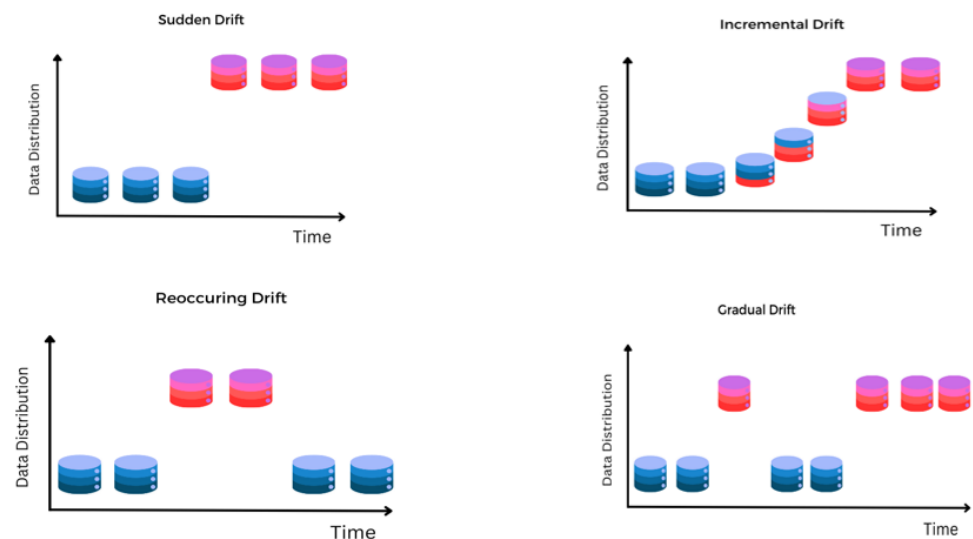


Figure 2. Pattern-based concept drift: The figure illustrates different types of concept drift over time, where changes in data distribution occur in sudden, incremental, reoccurring, and gradual patterns.

Pattern-based drift can be sudden, incremental, gradual, or recurrent. Sudden drift typically characterizes instantaneous data distribution changes, while incremental drift involves slow, continuous data changes. Meanwhile, gradual drift signifies consistent changes over an extended period, and recurrent drift exhibits periodical and cyclical shifts in data distribution.

The above described concept drift types are summarized in Table 2.

Table 2. Summary of the concept drift types.

Drift Type	Description	Real-World Example
Virtual Drift	Changes in $P(X)$ without altering $P(Y X)$.	Seasonal variations in customer transactions.
Real Drift	Changes in $P(Y X)$, affecting model accuracy.	Introduction of new fraud schemes in credit card transactions.
Sudden Drift	Abrupt, significant changes in data distribution.	Market crashes in financial forecasting.
Gradual Drift	Slow, continuous changes over time.	E-commerce trends shifting during holiday seasons.
Incremental Drift	Progressive evolution of data distribution.	Evolving user preferences in recommendation systems.
Recurrent Drift	Cyclical changes in data distribution.	Seasonal sales patterns in retail.

Most researchers concentrate on specific types of drifts. Additionally, we noticed significant similarities between several concept drift detection mechanism types. Consequently, we categorized drift detection mechanisms into five main types.

Drift detection mechanism. For drift detection mechanisms (DDMs), techniques such as statistical tests and control charts are commonly used. These methods show high accuracy in detecting distribution-based drifts.

Window-based mechanism (WBM): Methods employing a window-based mechanism (WBM), like sliding windows and time-based windows, are used to manage data streams. These approaches are effective for real-time drift detection.

Unsupervised and semi-supervised methods: Unsupervised and semi-supervised methods (USSMs) include clustering and semi-supervised learning. These methods are adaptable to changes in real-world data patterns and are often used to detect novel classes.

Ensemble method: Ensemble methods (EMs) combine multiple models to improve detection accuracy and robustness. Studies demonstrate high performance across various data types.

Neural networks: Advanced neural network (NN) models are employed for detecting complex drifts. Although relatively new in concept drift detection, these methods show high adaptability and accuracy.

3. Results

3.1. Study Selection

In our exploration of the concept drift carried out across a selected pool of 111 eligible studies, we embarked on a comprehensive analysis. This analysis resulted in a final ensemble of 65 high-impact studies. Our primary focus was on the different types of concept drift represented in these studies, forming a fundamental aspect of our forthcoming discussions. This included examining the specific strategies employed to mitigate concept drift, evaluating the strengths and weaknesses of these approaches, and assessing their practical applicability. We also spotlighted the commonly invoked comparison metrics, frameworks, and datasets supporting a comparative study of concept drift detection strategies. The meticulous findings from our exhaustive analysis of these 65 studies formed the backbone of our comprehensive review. In Figure 3, we provide an illustrative diagram to further elucidate our meticulous selection process and its resulting reductions at each stage. This PRISMA diagram summarizes the screening and eligibility assessment stages and the reasons for exclusion at each phase. By referring to this visual aid, readers can more easily comprehend our meticulous step-by-step approach and appreciate the depth of our systematic review process.

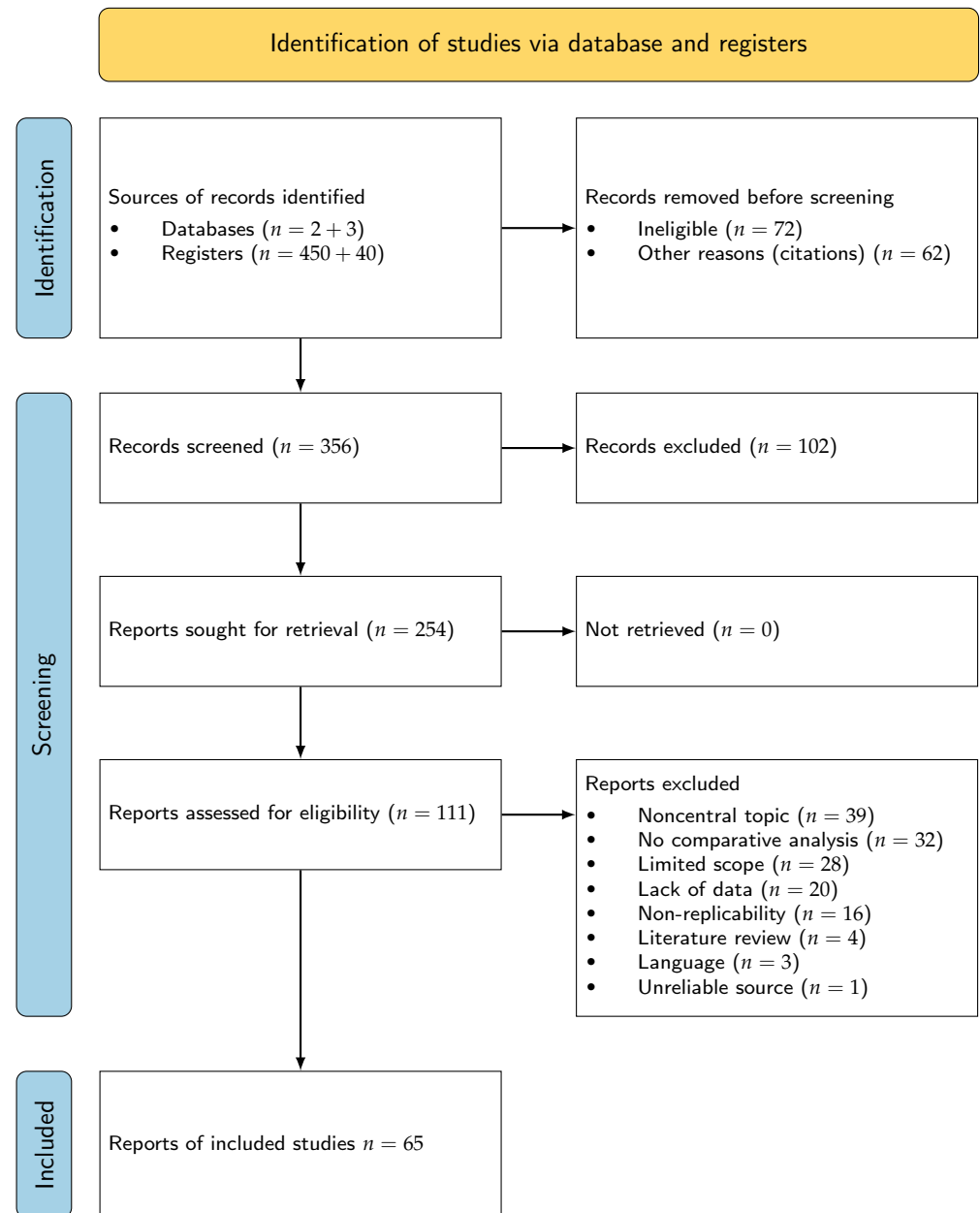


Figure 3. PRISMA flow diagram illustrating the selection process of the studies.

3.2. Study Characteristics

The included studies were diverse regarding the study design, drift types, and methods used for concept drift detection. Table 3 summarizes these characteristics.

Table 3. Summary of the characteristics of the included studies.

Study	Drift Type	Method	Findings
DDM [19]	Sudden+Gradual	DDM	Detects abrupt and gradual drifts by error rate changes.
Early DDM (EDDM) [20]	Sudden+Gradual	DDM	Improves gradual drift detection using classification error distances.
Hoeffding's DDM (HDDM) [21]	Sudden+Gradual	DDM	Non-parametric online detection for abrupt and gradual drifts.
Fast Hoeffding's DDM (FHDDM) [22]	Sudden+Gradual	DDM + WBM	Enhances HDDM speed and sensitivity for abrupt and gradual drifts.
Reactive DDM (RDDM) [23]	Sudden+Gradual+Reoccurring	DDM	Quickly detects abrupt and gradual drifts via the error rate.
Accurate DDM (ACDDM) [24]	Sudden+Gradual+Reoccurring	DDM	Identifies abrupt, gradual, and recurring drifts with high accuracy.
Diversity-measure DDM (DMDDM) [25]	Sudden+Gradual+Reoccurring	DDM	Uses classifier diversity to detect various drift types.
DDM with false positive rate (DDM-FP-M) [26]	Sudden+Gradual+Reoccurring	DDM	Controls false positives while detecting various drift types.
Noise-tolerant DDM [27]	Sudden+Gradual+Reoccurring	DDM	Distinguishes noise from drifts and detects gradual and abrupt changes.
Group DDM [28]	Sudden+Gradual+Reoccurring	DDM	Detects multiple drift types in data streams simultaneously.
Detecting concept drift using statistical testing (STEPD) [29]	Sudden+Gradual	WBM	Detects drift using hypothesis tests on data streams and balances false positives with detection sensitivity effectively.
Learning from time-changing data with adaptive windowing (ADWIN) [30]	Sudden+Gradual	WBM	Detects sudden and gradual drift efficiently; sensitive to noise.
Paired learners for concept drift (PL) [31]	Sudden+Gradual	WBM	Uses stable and reactive learners to detect concept drift efficiently and outperforms other methods with fewer resources.
Mining concept-drifting data streams with multiple semi-random decision trees (MSRTs) [32]	Sudden+Gradual+Reoccurring	WBM	Utilizes multiple semi-random decision trees for drift detection, adapting dynamically, and demonstrates improved performance in time and accuracy.
Double window-based classification algorithm for concept-drifting data streams [32]	Sudden+Gradual	WBM	Employs the double window mechanism
Dynamic financial distress prediction using instance selection for the disposal of concept drift [33]	-	WBM	Discusses four WBM types: full memory, no memory, fixed size, and adaptable size.
Classifier ensemble for text streams [34]	Sudden+Gradual	WBM	Experiments with textual data
EWMA charts [4]	Sudden+Gradual	WBM	Maintains a constant false positive rate without storing data points
iForestASD [35]	Sudden+Gradual+Reoccurring	WBM	Uses sliding with isolation forest windows for anomaly detection.
TRIO algorithm [36]	Gradual	WBM	Combines text categorization and SVM for the early detection of gradual drifts
SEED algorithm [37]	Sudden+Gradual	WBM	Integrates drift and volatility detection to handle high-volume and high-velocity data.

Table 3. Cont.

Study	Drift Type	Method	Findings
CEVOT [38]	Sudden+Gradual	WBM	Combines evolutionary algorithms with decision trees for incremental adaptation.
Dynamical systems tools [39]	Sudden+Gradual	WBM	Uses dynamical models for detecting subtle changes
Fisher’s exact test [40]	Sudden+Gradual	WBM	Adapts the Fisher test for efficient drift detection.
McDiarmid drift detection [41]	Sudden+Gradual	WBM	Novel approach to detect concept drift using McDiarmid’s inequality.
Bayesian approach [42]	Sudden	WBM	Bayesian method specifically for detecting abrupt concept drift.
Gaussian mixture models [43]	Sudden+Gradual+Reoccurring	WBM	Detects rapid changes in industrial settings.
Learn++ with sliding time window [44]	Sudden+Gradual+Reoccurring	WBM	Addresses imbalanced classification
Just-in-time learning (JITL) [45,46]	Sudden+Gradual+Reoccurring	WBM	Integrates with window mechanisms for soft sensor design.
Shifting window mechanism for SVMs [47,48]	Sudden+Gradual+Reoccurring	WBM	Integrates SVM into methodologies to manage concept drift effectively.
Spectral entropy and bernoulli map [49]	Sudden+Gradual	WBM	Employs spectral entropy combined with the Bernoulli map.
Multi-sliding windows for drift type identification [50]	-	WBM	Main focus is on the detection of drift type.
Classification and novel class detection in concept-drifting data streams under time constraints [51]	Novel Class	USSM	Addresses the novel class detection problem
SUN algorithm [52]	Sudden+Gradual	USSM	SUN employs k-modes and decision trees to identify concept drift effectively.
Detection of concept drift for learning from stream data [53]	Sudden+Gradual	USSM	Uses correlation information to detect concept drift in multi-stream data models.
Recognizing input space and target concept drifts with scarcely labeled and unlabeled instances [54]	Sudden+Gradual+Reoccurring	USSM	Uses active learning and classifier certainty to detect drifts with minimal labeled data.
Margin density drift detection (MD3) [55]	Sudden+Gradual	USSM	Tracks the number of samples in the uncertainty region as a metric to detect drift.
Handling adversarial concept drift in streaming data [56]	Sudden	USSM	With predict–detect framework, includes adversarial forethought and context for drift detection and mitigation.
DetectA: abrupt concept drift detection in non-stationary environments [57]	Sudden	USSM	As a proactive approach, uses unsupervised learning and statistical comparisons for efficient drift detection.
Concept drift robust adaptive novelty detection for data streams [58]	Novel Class	USSM	Includes ELBND, LE, and adaptive methods using NLMS metrics and excels in robust drift detection.

Table 3. Cont.

Study	Drift Type	Method	Findings
OLINDDA technique [59]	Novel Class	USSM	Presents a cluster-based approach for detecting novelty and concept drift
Online reliable semi-supervised learning on evolving data streams [60]	Sudden+Gradual+Reoccurring	USSM	As a micro-cluster-based technique, ensures high performance in semi-supervised environments with limited labeled data.
Semi-supervised classification on data streams with recurring concept drift and concept evolution [61]	Sudden+Gradual+Reoccurring	USSM	Uses an ESCR framework employs Jensen–Shannon divergence and clustering for efficient and accurate drift detection.
Conformal prediction for semi-supervised classification on data streams (CPSSDS) algorithm [62]	Reoccurring	USSM	Proposes a semi-supervised framework with conformal prediction and self-training for efficient drift detection in evolving data streams.
Novel semi-supervised classification approach for evolving data streams [63]	Sudden+Gradual+Reoccurring	USSM	As a semi-supervised framework, uses dynamic micro-clusters and online updates for efficient concept drift detection.
Robust ensemble learning for mining noisy data streams [64]	Sudden+Gradual	EM	Employs AE framework that combines varied learning algorithm-based classifiers, enhancing performance on noisy data streams.
Adaptive ensemble classifier for mining concept-drifting data streams [65]	Sudden+Gradual	EM	Merges decision trees and clustering to enhance classification accuracy and adaptability.
Online accuracy updated ensemble (OAUE) [3]	Sudden+Gradual+Reoccurring	EM	Transforms block-based ensembles for efficient reactions to sudden and gradual concept drifts.
Ensemble of subset online sequential extreme learning machine (ESOS-ELM) [66]	Sudden+Gradual	EM	Balances class samples using dynamic weighted voting, excelling in both stationary and non-stationary environments.
DEVE-AT and ADASVM-TW algorithms [67]	Sudden+Gradual+Reoccurring	EM	Use time weighting and Adaboost-SVM for superior financial distress prediction.
Gradual resampling ensemble (GRE) algorithm [68]	Gradual	EM	Uses selective resampling, clustering, and dynamic updates to achieve superior performance across various drift types.
Selection-based resampling ensemble (SRE)	Sudden+Gradual+Reoccurring	EM	Balances class distributions and periodically updates members for robust performance across concept drifts.
Handling imbalanced data with concept drift (HIDC) [69]	Sudden+Gradual+Reoccurring	EM	Combines dynamic sampling and ensemble classification, adapting to ensure accuracy despite concept drift and class imbalance.
Preprocessed dynamic classifier ensemble selection for highly imbalanced drifted data streams [70]	Sudden+Gradual	EM	Integrates preprocessing and dynamic selection, enhancing performance in imbalanced and non-stationary data.
AESAKNNS [71]	Gradual	EM	Employs adaptive subspaces and dynamic training to achieve superior adaptation to concept drift in multi-label datasets.

Table 3. Cont.

Study	Drift Type	Method	Findings
Extreme learning machine (ELM) [66]	Sudden+Gradual+Reoccurring	NN	With a quick learning speed, uses a single-step LSE method for training SLFN, bypassing iterative gradient descent methods.
Fast incremental ELM algorithm [72]	Sudden+Gradual+Reoccurring	NN	Determines neuron quantity and randomly selects activation functions to achieve improved accuracy in classification.
FP-ELM [73]	Sudden+Gradual+Reoccurring	NN	Performs better than traditional ensemble approaches on several regression and classification problems with concept drift.
GPU-accelerated ELM [74]	Sudden+Gradual+Reoccurring	NN	Balances classes and detect drift, reducing processing time, providing effective real-time big data mining.
Meta-cognitive online sequential ELM (MOS-ELM) [75]	Sudden+Gradual+Reoccurring	NN	Outperforms several existing methods on various multi-class datasets, effectively handling both imbalanced data and concept drift
Dynamic ELM (DELM) [76]	Sudden+Gradual+Reoccurring	NN	Uses ELM and a drift detection threshold to ensure superior accuracy and adaptability in various concept drift scenarios.
Evolving spiking NN (eSNN) [77]	Gradual+Reoccurring	NN	Employs data reduction and repository size limitation, providing higher accuracy in online learning environments.
Meta-RKOS-ELM-DDM [78]	Sudden+Gradual+Reoccurring	NN	Integrates a DDM and ALD kernel filter, enhancing prediction accuracy and learning efficiency.
Meta-RRKOS-ELM [79]	Sudden+Gradual+Reoccurring	NN	Merges recurrent kernel OS-ELM with DDM and the recursive kernel method to achieve enhanced performance and efficiency in handling drift.
I-LSTM [80]	Gradual+Reoccurring	NN	For IoT services, uses smooth activation and a time-based forgetting mechanism to achieve superior anomaly detection in time-series data.
SEOA [81]	Sudden+Gradual+Reoccurring	NN	Uses adaptive depth units in an ensemble to achieve superior performance, balance, speed, and generalization in handling concept drift.

3.3. Findings on Concept Drift Detection Methods

The studies were categorized based on the methods used for detecting concept drift. This section details the findings for each category.

DDMs. The concept of drift detection was significantly advanced by the work of Gama et al. in 2004 with their introduction of the DDM [19]. This method has since become a benchmark, inspiring numerous subsequent techniques that either build upon or modify the original algorithm. The main structure of the DDM is provided in Algorithm 1.

Algorithm 1: General algorithm for drift detection mechanisms.

Input: Data stream continuously entering the system

Output: Updated model, responsive to the concept drift

Function DetectDrift(*data*):

while *new data available* **do**

 Calculate error rate p_t

 Compute standard deviation s_t :

$$s_t = \sqrt{\frac{p_t(1-p_t)}{t}}$$

 // Set the minimum error rate and standard deviation based
 on historical data or the initial stable period

 Set p_{\min} as the minimum observed error rate

 Set s_{\min} as the standard deviation corresponding to p_{\min}

 Set warning level threshold:

$$p_t + s_t \geq p_{\min} + 2 \times s_{\min}$$

 Set drift level threshold:

$$p_t + s_t \geq p_{\min} + 3 \times s_{\min}$$

if *warning level detected* **then**

 Initiate a new model

end

if *drift level detected* **then**

 Replace current model

end

end

The DDM operates based on Statistical Process Controls (SPCs) [82] by monitoring the error rate of a predictive model over time. As new data points are processed, the algorithm calculates the error rate p_t and its standard deviation s_t . The method sets two critical thresholds: a warning level and a drift level. The warning level threshold is defined as $p_t + s_t \geq p_{\min} + 2 \times s_{\min}$, indicating potential concept drift and prompting the system to initiate a new model. The drift level threshold, defined as $p_t + s_t \geq p_{\min} + 3 \times s_{\min}$, confirms the drift, leading to the replacement of the current model.

This approach allows the system to dynamically adapt to changes in the data stream, maintaining model accuracy and performance. The pioneering work by Gama et al. laid the foundation for many advanced drift detection methods that incorporate various enhancements and modifications to improve detection accuracy and responsiveness to different types of drifts. While DDM is efficient for real-time applications, it can produce false positives in noisy environments. Its limitations in detecting gradual drift have led to the development of extensions like EDDM [20].

WBMs. WBMs are pivotal in the detection of concept drift within data streams. By utilizing a systematic approach that compares the statistical characteristics of historical data with newly observed data, these methods effectively monitor and identify shifts in data distribution. Typically, these approaches involve the use of two distinct data windows, one static (historical) and one dynamic (adjusting with new data), allowing for the continuous assessment and identification of changes over time.

Several seminal papers, though published over a decade ago, continue to serve as benchmarks in the field of concept drift detection. Notably, the research paper “Detecting Concept Drift Using Statistical Testing” by Kyosuke Nishida from 2007 [29] introduced the STEP algorithm, which remains influential. Meanwhile, Albert Bifet’s 2007 work “Learning from Time-changing Data with Adaptive Windowing” (ADWIN) [30] has made substantial contributions to adaptive learning strategies. Based on these papers, many new advanced techniques have been developed.

The general structure of the window-based concept DDMs is described in Algorithm 2.

Algorithm 2: WBM for concept drift detection

Input: Data stream continuously entering the system

Output: Updated model, responsive to the concept drift

Function DetectDrift(*data*):

 Define window types (Sliding, Fixed, Adaptive)

 Initialize model with initial training set

while new data available **do**

 Update window with incoming data

if window type is adaptive **then**

 Adjust window size based on drift detection

else

 Maintain fixed or slide window

end

 Detect potential drift using statistical tests or error monitoring

if drift detected **then**

 Retrain or update the model with data within the current window

 Optionally adjust ensemble models

 Evaluate model performance

 Adjust window strategy based on performance feedback

 Trigger alerts or automated responses

end

 Provide updated predictions

end

WBMs like ADWIN are mostly responsive to sudden drifts, but are computationally intensive. They may struggle with gradual drift due to their focus on distinct statistical shifts. Another challenge for learning is that the feedback (the ground truth of mass flow) is not available at all; it can only be approximately estimated by retrospectively inspecting the historical data. An additional challenge is to deal with specific one-sided outliers that can be easily mistaken for changes [83].

USSMs: USSMs detect concept drift by leveraging clustering, density estimation, and other techniques to monitor changes in data distributions. These methods are particularly effective in scenarios with sparse labeled data or mixed data streams containing both categorical and numerical values. USSMs have gained significant attention over the past two decades due to their ability to handle complex, real-world datasets and detect novel classes as they emerge [84]. USSMs focus on identifying changes in data patterns without relying heavily on labels. For instance, clustering-based approaches monitor the formation and evolution of clusters over time, flagging new or significantly altered clusters as potential

drifts. This makes USSMs especially valuable in applications like fraud detection and cybersecurity, where data streams may exhibit unexpected patterns or novel behaviors [59].

USSMs often rely on clustering techniques to group similar data points into clusters, representing stable patterns in the data. When new data points deviate significantly from these clusters, they are flagged as drift events. Novel class detection, a key application of USSMs, identifies previously unseen categories in the data stream, making these methods highly adaptable in dynamic environments such as network security or e-commerce [59]. Density-based methods within USSMs provide probabilistic interpretations for drift detection. These methods calculate the density of data points in a given space, with significant density changes indicating potential drift [85]. USSMs exhibit several advantages. These methods can detect novel patterns and categories, making them ideal for dynamic environments with minimal labeled data. However, USSMs also present notable disadvantages. Clustering and density estimation methods are computationally expensive, especially with high-dimensional data streams. Furthermore, USSMs are prone to false positives in noisy datasets as outliers can be misclassified as drift events [86].

EMs: EMs have become exceedingly popular for concept drift detection due to their ability to combine multiple models, improving predictive performance and robustness. By leveraging the diversity of individual models, EMs can effectively adapt to changes in data distribution, making them highly suitable for dynamic environments where concept drift is prevalent. This adaptability is particularly valuable in addressing imbalanced datasets, where ensemble algorithms provide superior performance by emphasizing minority classes and improving the overall detection rate [3].

Ensemble techniques like the online accuracy updated ensemble (OAUE) adapt model weights based on recent performance, ensuring that the ensemble remains responsive to current data patterns. This dynamic adjustment aligns with decision theory, where weighting models based on their suitability to the current data distribution enhances robustness [3]. OAUE, for instance, continuously evaluates model performance and re-weights or replaces individual models as needed, maintaining high accuracy in non-stationary environments. EMs are also effective in handling recurring drifts by preserving previously trained models that can quickly be reactivated when similar patterns reappear.

The theoretical foundation of EMs lies in ensemble learning theory, which emphasizes the use of diverse classifiers to reduce bias and variance. By combining predictions from multiple models, ensemble methods mitigate overfitting and improve generalization [3]. Techniques like bagging and boosting are commonly employed to train diverse classifiers, while weighting strategies prioritize models most attuned to current data distributions. These methods enable EMs to handle both abrupt and gradual drifts, as well as recurring patterns, with minimal loss in accuracy.

While EMs exhibit several advantages, they also come with notable drawbacks. EMs provide high accuracy and adaptability, making them effective across a variety of application domains such as fraud detection, predictive maintenance, and real-time recommendation systems. However, their computational demands are significant due to the overhead of maintaining multiple models and dynamically adjusting their weights. This complexity requires sophisticated management strategies to ensure that the ensemble optimally adapts without incurring excessive resource costs. Additionally, EMs may face scalability challenges in high-velocity data streams where quick responses are critical.

Despite these challenges, ensemble methods continue to be a cornerstone in concept drift detection research. Their versatility and robustness make them indispensable for real-world applications, especially in domains where data distributions are highly volatile or imbalanced.

NNs: Neural networks (NNs) have attracted significant attention in the field of concept drift detection because of their powerful learning capabilities and adaptability to changing data distributions. By leveraging deep learning techniques, NNs can effectively capture complex patterns and relationships within data streams, making them well suited for dynamic environments. Among the various neural network approaches, those from the

extreme learning machine (ELM) family have become particularly popular. ELMs are widely used in large datasets and online learning applications due to their fast learning speed. Unlike iterative gradient descent methods, such as backpropagation, ELMs utilize a single-step least square estimation (LSE) method for training a single-hidden-layer feed-forward network (SLFN), making them highly efficient [66].

Neural networks, particularly deep architectures, are capable of modeling complex, non-linear relationships in high-dimensional and unstructured data. Incremental learning methods, such as ELMs, enhance adaptability for data streams with frequent drift [66]. LSTM (long short-term memory) networks are another popular approach for handling sequential data. Their ability to retain temporal dependencies makes them highly effective in detecting concept drift in time-series data. To maintain stability in dynamic environments, regularization techniques like elastic weight consolidation (EWC) have been developed. EWC minimizes catastrophic forgetting by preserving critical parameters while adapting to new data, ensuring that previously learned knowledge is not overwritten [77].

The theoretical foundations of NNs in drift detection are grounded in statistical learning and neural computation, emphasizing adaptability and generalization. ELMs, in particular, are designed for efficient drift detection, achieving high accuracy while reducing computational overhead. Techniques like dropout and EWC further improve the robustness of neural networks by preventing overfitting and mitigating catastrophic forgetting [6].

While NNs excel in adaptability and accuracy, they also have notable disadvantages. NNs are computationally intensive, requiring significant resources for training and inference, which can limit their scalability in real-time applications. Additionally, they are prone to catastrophic forgetting, where the model's performance on previously seen data degrades as it learns new patterns. Regularization techniques like EWC provide solutions to this issue but add to the complexity of implementation. Despite these challenges, the flexibility and robustness of NNs make them invaluable in domains such as financial time-series analysis, fraud detection, and dynamic pricing, where data distributions evolve continuously [6].

3.4. Comparison of Concept Drift Detection Methods

To objectively assess the performance of concept drift detection methods, this section incorporates findings from comparative studies that evaluate accuracy, computational cost, and applicability across various data types and drift scenarios.

Recent studies have systematically compared these methods, providing deeper insights into their relative performance across various contexts. Barros et al. presented two papers in 2018 and 2019 where researchers evaluated different methods across multiple datasets and drift types. Their studies highlighted the robustness of ensemble techniques, particularly their adaptability to diverse drift scenarios, though computational cost remained a challenge [87,88]. Other important findings were provided by Poenaru-Olaru et al. (2022). The researchers analyzed the reliability of concept drift detectors in real-time applications. Their research compared error rate-based and data distribution-based detectors, revealing trade-offs between detection delay and false alarm rates [89]. Hinder et al. (2023) provided a survey and standardized experiments to benchmark unsupervised drift detection methods. Their study underscored the effectiveness of USSMs in scenarios with sparse labels but noted their susceptibility to noise [90]. Last, but not least, LSTM networks [91] are particularly effective in handling sequential data as they retain temporal dependencies crucial for detecting concept drift in time-series applications. For instance, Lobo et al. (2018) demonstrated the effectiveness of evolving spiking neural networks, inspired by LSTMs, for online learning over drifting data streams [77].

Key Observations

1. **Accuracy vs. Computational Cost:** While neural networks (NNs) and ensemble methods (EMs) provide the highest accuracy, their computational cost limits their real-time applicability. In contrast, drift detection mechanisms (DDMs) and window-based mechanisms (WBMs) offer a balance between accuracy and efficiency.

2. Specialized Use Cases: Unsupervised and semi-supervised methods (USSMs) excel in novel class detection, and WBM are ideal for streaming environments. EMs and NNs are better suited for complex, evolving data distributions.
3. Emerging Trends: Hybrid approaches combining lightweight methods (e.g., DDMs) with adaptive techniques (e.g., NNs) show promise in balancing computational efficiency with detection accuracy.

The findings are summarized in Table 4.

Table 4. Comparison of the concept drift detection methods.

Method	Accuracy	Computational Cost	Applicability
DDM	High	Low	Cost effective and easy to apply in real time
WBM	Medium	Medium	Works better with gradual drift than DDM and easy to apply in real time
USSM	Medium	Medium	Works best with novel class detection
EM	Very High	High	Can be applied to various data types but with higher computational cost
NN	Very High	High	Mostly have higher cost but higher accuracy

3.5. Other Notable Findings on Concept Drift Detection Methods

In reviewing numerous papers on concept drift detection using different methods, several key comparison criteria emerged that are commonly used to evaluate the performance of these approaches. These criteria can be used to comprehensively assess a model's effectiveness and efficiency in handling dynamic data streams. The main comparison criteria included the following.

Prequential error. The prequential error measurement method developed by Dawid [92] is mostly used to compute a model's accuracy. The prequential error is often calculated using common evaluation metrics such as accuracy. The incremental calculation for accuracy (Acc) in a prequential context can be expressed as follows:

$$Acc_{t+1} = \frac{t \times Acc_t + \delta_{t+1}}{t + 1},$$

where Acc_{t+1} is the accuracy after observing the $(t + 1)$ th instance, Acc_t is the accuracy after observing the t th instance, and δ_{t+1} is the binary indicator of whether the prediction for the $(t + 1)$ th instance is correct (1 if correct, 0 if incorrect).

Handling imbalanced classification. Many real-world data streams are imbalanced, and some classes are significantly under-represented. Evaluating how models handle imbalanced classification is essential as it affects a model's ability to learn from minority classes and maintain high overall accuracy. Common strategies to address imbalanced classification include the following:

- Resampling techniques—methods such as SMOTE (synthetic minority over-sampling technique) and its variants are widely used to balance datasets. Recent advancements, such as adaptive oversampling techniques, focus on regions prone to classification errors [93].
- Cost-sensitive learning—adjusting the learning algorithm to penalize the misclassifications of a minority class more heavily [94].
- Synthetic data generation—creating synthetic examples of a minority class to balance the dataset [95].

The effectiveness of these strategies is typically assessed using metrics like F1 score, precision–recall curves, and area under the curve (AUC) for receiver operating characteristic (ROC) analysis.

Comparison of speed: The speed of model training and prediction is a critical factor, especially in real-time applications. Various metrics are used to measure and compare the computational efficiency of different approaches:

- Training time—incremental learning models like Hoeffding trees are well suited for reducing the training overhead in data streams. One of the most popular approaches in concept drift is considered Hoeffding trees [96].
- Prediction time—the time taken to make predictions on new data points.
- Throughput—the number of data points processed per unit time. DDM's maintain high throughput with minimal computational overhead, making them suitable for high-speed applications.
- Latency—the delay before the system starts to output predictions after receiving new data [97].

Evaluating these metrics helps us understand the trade-offs between model complexity, accuracy, and computational demands such that the chosen approach can meet the real-time requirements of the application.

Dataset characteristics: Researchers commonly utilize a range of benchmark datasets to thoroughly evaluate the performance and robustness of various concept drift detection methods. These datasets represent diverse real-world scenarios, including different types and magnitudes of concept drift, class imbalances, and varying data distributions. In Table 5, we summarize the main datasets and their characteristics used for the concept drift detection problem.

Many of these datasets are implemented and tested within the MOA (Massive On-line Analysis, ver. 24.07.0) framework, a widely used open-source software for data stream mining.

The THU-Concept-Drift-Datasets are relatively new. These datasets are free and integrated with convenient interfaces for data stream generation and manipulation, making them an excellent resource for testing and comparing concept drift detection methods [98].

Table 5. Summary of the datasets used for concept drift detection.

Dataset Type	Name	Characteristics	Short Description
Synthetic	SEA Concepts	Abrupt drift	Data points generated from three attributes, with concept drift introduced by changing the decision boundary.
Synthetic	Hyperplane	Gradual drift	Represents a moving hyperplane in a multi-dimensional space, with drift simulated by changing the hyperplane's position or orientation over time.
Synthetic	Rotating Hyperplane	Continuous drift	A variant of the Hyperplane dataset where the hyperplane rotates over time, introducing smooth concept drift.
Synthetic	Random RBF	Sudden and gradual drift	Instances generated based on several centroids, with drift introduced by moving the centroids or changing their standard deviations.
Real-world	Electricity Market (Elec2)	Abrupt and gradual drift	Contains electricity price data, with the task of predicting price changes under varying market conditions.
Real-world	Airlines	Concept drift due to changing factors	Flight delay information, used to predict delays based on historical data, reflecting drift from changing flight patterns and weather conditions.
Real-world	KDD Cup 1999	Various attack patterns	Network intrusion detection dataset with different types of attacks, used to evaluate robustness against varying attack patterns.
Real-world	Weather Data	Gradual and abrupt drift	Weather prediction dataset incorporating both gradual and abrupt drift due to seasonal variations and unexpected events.
Real-world (imbalanced)	Credit Card Fraud Detection	Class imbalance	Credit card transaction dataset used to detect fraudulent activities, characterized by a significant imbalance between fraud and non-fraud cases.
Imbalanced	Medical Data	Rare events	Used for disease prediction tasks, where the incidence of the condition is much lower than that of the non-condition, testing the algorithm's ability to detect rare events.

Table 5. Cont.

Dataset Type	Name	Characteristics	Short Description
THU-Concept-Drift-Datasets	Linear	Gradual, sudden, recurrent, abrupt drift	Decision boundary is a straight line, with drift simulated by rotating the line.
THU-Concept-Drift-Datasets	CakeRotation	Rotational drift	Samples with odd/even angle areas belong to different classes, with drift simulated by rotating the disk.
THU-Concept-Drift-Datasets	ChocolateRotation	Rotational drift	Samples with odd/even $x + y$ areas belong to different classes, with drift introduced by rotating the chocolate plate.
THU-Concept-Drift-Datasets	RollingTorus	Overlapping categories	Features two toruses where samples in different toruses belong to different classes, with drift simulated by rolling a third torus over the first two.

4. Conclusions and Future Work Directions

This systematic literature review explored the evolving strategies in concept drift detection over the past two decades. By analyzing a wide range of methodologies, including DDMs, WBM, USSMs, EMs, and NN techniques, we provide a comprehensive overview of how the field has progressed and adapted to the dynamic nature of data streams. Our review highlights the significant contributions of various algorithms and frameworks, each addressing specific challenges associated with concept drift.

The use of synthetic and real-world datasets has been crucial in evaluating the performance of these methods. Datasets such as SEA Concepts, Hyperplane, Electricity Market, and Credit Card Fraud Detection have been instrumental in testing the adaptability and robustness of concept drift detection algorithms. Additionally, the THU-Concept-Drift-Datasets and the MOA (Massive Online Analysis) framework have emerged as valuable resources for researchers, offering diverse scenarios to rigorously test and compare different methods. However, further efforts should focus on curating benchmark datasets that include real-world complexities, such as noisy environments, mixed drift types, and high-dimensional data, to better reflect practical applications.

Despite the extensive research and numerous advancements in the field, several weak points remain. One major limitation is the handling of imbalanced data, which continues to pose significant challenges for many algorithms. While some methods have been specifically designed to address class imbalance, further research is needed to develop more effective and generalizable solutions. For example, integrating dynamic ensemble methods with adaptive resampling techniques could address imbalances in streaming environments. *Recommendation:* Investigate hybrid frameworks combining oversampling with adaptive cost-sensitive algorithms to improve performance on imbalanced datasets.

Another area that requires attention is the computational efficiency of concept drift detection methods. As the volume and velocity of data streams increase, the need for fast and scalable algorithms becomes more critical. Many current approaches still struggle with high computational costs, which can limit their applicability in real-time environments. *Recommendation:* Explore parallel computing and hardware acceleration, such as GPU-optimized neural networks, and investigate lightweight, low-latency detection models for deployment in real-time scenarios.

Additionally, concept drift detection in regression tasks has not been as thoroughly analyzed as in classification tasks. Many existing studies focus on classification, leaving a gap in understanding and effectively addressing concept drift in regression scenarios. More research is needed to develop robust methods for regression tasks, where the continuous nature of the target variable presents unique challenges. *Recommendation:* Develop regression-specific drift detection metrics and algorithms that focus on subtle shifts in continuous relationships, such as changes in correlation structures or error variance.

The application of concept drift detection methods to non-tabular datasets, such as image and time-series data, remains underexplored. Most current approaches are designed for tabular data, and adapting these methods to handle image data's high-dimensional and complex nature presents a significant challenge. Techniques such as convolutional neural

networks (CNNs) and graph neural networks (GNNs) could be adapted to directly integrate drift detection into their architectures. *Recommendation*: Investigate the application of unsupervised feature extraction techniques, such as autoencoders, combined with domain-specific neural architectures to handle non-tabular data effectively.

Another pressing issue is the limited focus on unsupervised drift detection methods, which are critical for scenarios where labeled data are scarce or unavailable. Advancing clustering-based and density-based techniques to handle high-dimensional, noisy data is a promising direction. *Recommendation*: Incorporate self-supervised learning techniques to generate pseudo-labels, enabling more effective drift detection in unlabeled datasets. Develop approaches that require less computational cost.

Moreover, the majority of existing studies focus on synthetic and controlled datasets, which may not fully capture the complexities of real-world data streams. More research is needed on diverse and representative datasets to ensure that the proposed methods can be effectively generalized to practical applications. *Recommendation*: Curate real-world benchmark datasets with annotations for known drift types, including contextual metadata, to improve the practical evaluation of drift detection methods.

Finally, the evaluation of concept drift detection methods would benefit from the development of standardized protocols. Existing evaluation criteria often emphasize accuracy while neglecting practical metrics such as detection delay, computational throughput, and memory footprint. *Recommendation*: Establish a unified evaluation framework incorporating detection latency, scalability metrics, and resource usage alongside traditional accuracy measures to promote practical applicability.

In summary, while significant progress has been made in the field of concept drift detection, ongoing research is essential to address the existing limitations. By leveraging advanced machine learning techniques and incorporating diverse datasets, future studies can further enhance the robustness and efficiency of concept drift detection methods, ensuring their applicability in a wide range of dynamic and evolving data environments.

Author Contributions: Conceptualization, J.M.B. and G.H.; methodology, G.H. and J.M.B.; software, G.H.; validation, G.H.; formal analysis, G.H.; investigation, G.H.; resources, G.H.; data curation, G.H.; writing—original draft preparation, G.H.; writing—review and editing, J.M.B.; visualization, G.H.; supervision, J.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by national funds through the FCT—Fundação para a Ciência e a Tecnologia, I.P., grants UIDB/04152/2020—Centro de Investigação em Gestão de Informação (MagIC) and UIDB/00315/2020—BRU-ISCTE-IUL.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lu, J.; Liu, A.; Dong, F.; Gu, F.; Gama, J.; Zhang, G. Learning under Concept Drift: A Review. *IEEE Trans. Knowl. Data Eng.* **2019**, *31*, 2346–2363. [\[CrossRef\]](#)
2. Iwashita, A.S.; Papa, J.P. An Overview on Concept Drift Learning. *IEEE Access* **2019**, *7*, 1532–1547. [\[CrossRef\]](#)
3. Brzezinski, D.; Stefanowski, J. Combining block-based and online methods in learning ensembles from concept drifting data streams. *Inf. Sci.* **2014**, *265*, 50–67. [\[CrossRef\]](#)
4. Ross, G.J.; Adams, N.M.; Tasoulis, D.K.; Hand, D.J. Exponentially weighted moving average charts for detecting concept drift. *Pattern Recognit. Lett.* **2012**, *33*, 191–198. [\[CrossRef\]](#)
5. Dal Pozzolo, A.; Boracchi, G.; Caelen, O.; Alippi, C.; Bontempi, G. Credit card fraud detection and concept-drift adaptation with delayed supervised information. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–8. [\[CrossRef\]](#)
6. Cavalcante, R.C.; Oliveira, A.L.I. An approach to handle concept drift in financial time series based on Extreme Learning Machines and explicit Drift Detection. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–8. [\[CrossRef\]](#)
7. Munirathinam, S. Drift detection analytics for iot sensors. *Procedia Comput. Sci.* **2021**, *180*, 903–912. [\[CrossRef\]](#)
8. Susnjak, T.; Maddigan, P. Forecasting patient flows with pandemic induced concept drift using explainable machine learning. *EPJ Data Sci.* **2023**, *12*, 11. [\[CrossRef\]](#)

9. Žliobaitė, I.; Pechenizkiy, M.; Gama, J. An overview of concept drift applications. In *Big Data Analysis: New Algorithms for a New Society*; Springer: Cham, Switzerland, 2016; pp. 91–114.
10. Langenkämper, D.; Van Kevelaer, R.; Purser, A.; Nattkemper, T.W. Gear-induced concept drift in marine images and its effect on deep learning classification. *Front. Mar. Sci.* **2020**, *7*, 506. [\[CrossRef\]](#)
11. Vapnik, V.N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **1999**, *10*, 988–999. [\[CrossRef\]](#)
12. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
13. Hershey, J.R.; Olsen, P.A. Approximating the Kullback Leibler divergence between Gaussian mixture models. In Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, Honolulu, HI, USA, 15–20 April 2007.
14. Shalev-Shwartz, S. Online learning and online convex optimization. *Found. Trends Mach. Learn.* **2012**, *4*, 107–194. [\[CrossRef\]](#)
15. Paul, J.; Feliciano-Cestero, M.M. Five decades of research on foreign direct investment by MNEs: An overview and research agenda. *J. Bus. Res.* **2021**, *124*, 800–812. [\[CrossRef\]](#)
16. Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; Liu, P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *J. Mach. Learn. Res.* **2020**, *21*, 1–67.
17. Wells, G.A.; Shea, B.; O'Connell, D.; Peterson, J.; Welch, V.; Losos, M.; Tugwell, P. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses*; Ottawa Hospital Research Institute: Ottawa, ON, Canada, 2000.
18. Critical Appraisal Skills Programme. *CASP Qualitative Research Checklist*; Critical Appraisal Skills Programme: Oxford, UK, 2017.
19. Gama, J.; Medas, P.; Castillo, G.; Rodrigues, P. Learning with drift detection. In Proceedings of the Advances in Artificial Intelligence—SBIA 2004: 17th Brazilian Symposium on Artificial Intelligence, Sao Luis, Maranhao, Brazil, 29 September–1 October 2004; Proceedings 17; Springer: Berlin/Heidelberg, Germany, 2004; pp. 286–295.
20. Baena-Garcia, M.; del Campo-Ávila, J.; Fidalgo, R.; Bifet, A.; Gavaldá, R.; Morales-Bueno, R. Early drift detection method. In Proceedings of the Fourth International Workshop on Knowledge Discovery from Data Streams, Philadelphia, PA, USA, 20 August 2006; Citeseer: Princeton, NJ, USA, 2006; Volume 6, pp. 77–86.
21. Frías-Blanco, I.; Campo-Ávila, J.d.; 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 Trans. Knowl. Data Eng.* **2015**, *27*, 810–823. [\[CrossRef\]](#)
22. Pesaranghader, A.; Viktor, H.L. Fast hoeffding drift detection method for evolving data streams. In Proceedings of the Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2016, Riva del Garda, Italy, 19–23 September 2016; Proceedings, Part II 16; Springer: Cham, Switzerland, 2016; pp. 96–111.
23. Barros, R.S.; Cabral, D.R.; Gonçalves, P.M., Jr.; Santos, S.G. RDDM: Reactive drift detection method. *Expert Syst. Appl.* **2017**, *90*, 344–355. [\[CrossRef\]](#)
24. Yan, M.M.W. Accurate detecting concept drift in evolving data streams. *ICT Express* **2020**, *6*, 332–338. [\[CrossRef\]](#)
25. Mahdi, O.A.; Pardede, E.; Ali, N.; Cao, J. Diversity measure as a new drift detection method in data streaming. *Knowl.-Based Syst.* **2020**, *191*, 105227. [\[CrossRef\]](#)
26. Wang, P.; Jin, N.; Fehringer, G. Concept drift detection with False Positive rate for multi-label classification in IoT data stream. In Proceedings of the 2020 International Conference on UK-China Emerging Technologies (UCET), Glasgow, UK, 20–21 August 2020; pp. 1–4. [\[CrossRef\]](#)
27. Wang, P.; Jin, N.; Woo, W.L.; Woodward, J.R.; Davies, D. Noise tolerant drift detection method for data stream mining. *Inf. Sci.* **2022**, *609*, 1318–1333. [\[CrossRef\]](#)
28. Yu, H.; Liu, W.; Lu, J.; Wen, Y.; Luo, X.; Zhang, G. Detecting group concept drift from multiple data streams. *Pattern Recognit.* **2023**, *134*, 109113. [\[CrossRef\]](#)
29. Nishida, K.; Yamauchi, K. Detecting concept drift using statistical testing. In Proceedings of the International Conference on Discovery Science, Sendai, Japan, 1–4 October 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 264–269.
30. Bifet, A.; Gavaldá, R. Learning from time-changing data with adaptive windowing. In Proceedings of the 2007 SIAM International Conference on Data Mining, Minneapolis, MN, USA, 26–28 April 2007; SIAM: Philadelphia, PA, USA, 2007; pp. 443–448.
31. Bach, S.H.; Maloof, M.A. Paired Learners for Concept Drift. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 5–19 December 2008; pp. 23–32. [\[CrossRef\]](#)
32. Li, P.; Hu, X.; Wu, X. Mining concept-drifting data streams with multiple semi-random decision trees. In Proceedings of the International Conference on Advanced Data Mining and Applications, Chengdu, China, 8–10 October 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 733–740.
33. Sun, J.; Li, H. Dynamic financial distress prediction using instance selection for the disposal of concept drift. *Expert Syst. Appl.* **2011**, *38*, 2566–2576. [\[CrossRef\]](#)
34. Yang, B.; Zhang, Y.; Li, X. Classifying text streams by keywords using classifier ensemble. *Data Knowl. Eng.* **2011**, *70*, 775–793. [\[CrossRef\]](#)
35. Ding, Z.; Fei, M. An anomaly detection approach based on isolation forest algorithm for streaming data using sliding window. *IFAC Proc. Vol.* **2013**, *46*, 12–17. [\[CrossRef\]](#)
36. Marseguerra, M. Early detection of gradual concept drifts by text categorization and Support Vector Machine techniques: The TRIO algorithm. *Reliab. Eng. Syst. Saf.* **2014**, *129*, 1–9. [\[CrossRef\]](#)

37. Huang, D.T.J.; Koh, Y.S.; Dobbie, G.; Pears, R. Detecting Volatility Shift in Data Streams. In Proceedings of the 2014 IEEE International Conference on Data Mining, Shenzhen, China, 14–17 December 2014; pp. 863–868. [\[CrossRef\]](#)
38. Jankowski, D.; Jackowski, K.; Cyganek, B. Learning decision trees from data streams with concept drift. *Procedia Comput. Sci.* **2016**, *80*, 1682–1691. [\[CrossRef\]](#)
39. da Costa, F.G.; Rios, R.A.; de Mello, R.F. Using dynamical systems tools to detect concept drift in data streams. *Expert Syst. Appl.* **2016**, *60*, 39–50. [\[CrossRef\]](#)
40. de Lima Cabral, D.R.; de Barros, R.S.M. Concept drift detection based on Fisher’s Exact test. *Inf. Sci.* **2018**, *442*, 220–234. [\[CrossRef\]](#)
41. Pesaranghader, A.; Viktor, H.L.; Paquet, E. McDiarmid drift detection methods for evolving data streams. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–9.
42. Cano, A.; Gómez-Olmedo, M.; Moral, S. A Bayesian approach to abrupt concept drift. *Knowl.-Based Syst.* **2019**, *185*, 104909. [\[CrossRef\]](#)
43. Diaz-Rozo, J.; Bielza, C.; Larrañaga, P. Machine-tool condition monitoring with Gaussian mixture models-based dynamic probabilistic clustering. *Eng. Appl. Artif. Intell.* **2020**, *89*, 103434. [\[CrossRef\]](#)
44. Wang, L.; Wu, C. Dynamic imbalanced business credit evaluation based on Learn++ with sliding time window and weight sampling and FCM with multiple kernels. *Inf. Sci.* **2020**, *520*, 305–323. [\[CrossRef\]](#)
45. Alakent, B. Soft sensor design using transductive moving window learner. *Comput. Chem. Eng.* **2020**, *140*, 106941. [\[CrossRef\]](#)
46. Urhan, A.; Alakent, B. Integrating adaptive moving window and just-in-time learning paradigms for soft-sensor design. *Neurocomputing* **2020**, *392*, 23–37. [\[CrossRef\]](#)
47. Gálmeanu, H.; Andonie, R. Weighted Incremental–Decremental Support Vector Machines for concept drift with shifting window. *Neural Netw.* **2022**, *152*, 528–541. [\[CrossRef\]](#)
48. Jain, M.; Kaur, G.; Saxena, V. A K-Means clustering and SVM based hybrid concept drift detection technique for network anomaly detection. *Expert Syst. Appl.* **2022**, *193*, 116510. [\[CrossRef\]](#)
49. Chikushi, R.T.M.; de Barros, R.S.M.; da Silva, M.G.N.M.; Maciel, B.I.F. Using spectral entropy and bernoulli map to handle concept drift. *Expert Syst. Appl.* **2021**, *167*, 114114. [\[CrossRef\]](#)
50. Husheng, G.; Hai, L.; Qiaoyan, R.; Wenjian, W. Concept Drift Type Identification Based on Multi-Sliding Windows. *Inf. Sci.* **2021**, *585*, 1–23.
51. Masud, M.; Gao, J.; Khan, L.; Han, J.; Thuraisingham, B.M. Classification and novel class detection in concept-drifting data streams under time constraints. *IEEE Trans. Knowl. Data Eng.* **2010**, *23*, 859–874. [\[CrossRef\]](#)
52. Wu, X.; Li, P.; Hu, X. Learning from concept drifting data streams with unlabeled data. *Neurocomputing* **2012**, *92*, 145–155. [\[CrossRef\]](#)
53. Lee, J.; Magoules, F. Detection of concept drift for learning from stream data. In Proceedings of the 2012 IEEE 14th International Conference on High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems, Liverpool, UK, 25–27 June 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 241–245.
54. Lughofer, E.; Weigl, E.; Heidl, W.; Eitzinger, C.; Radauer, T. Recognizing input space and target concept drifts in data streams with scarcely labeled and unlabelled instances. *Inf. Sci.* **2016**, *355*, 127–151. [\[CrossRef\]](#)
55. Sethi, T.S.; Kantardzic, M. On the reliable detection of concept drift from streaming unlabeled data. *Expert Syst. Appl.* **2017**, *82*, 77–99. [\[CrossRef\]](#)
56. Sethi, T.S.; Kantardzic, M. Handling adversarial concept drift in streaming data. *Expert Syst. Appl.* **2018**, *97*, 18–40. [\[CrossRef\]](#)
57. Escovedo, T.; Koshiyama, A.; da Cruz, A.A.; Vellasco, M. DetectA: Abrupt concept drift detection in non-stationary environments. *Appl. Soft Comput.* **2018**, *62*, 119–133. [\[CrossRef\]](#)
58. Cejnek, M.; Bukovsky, I. Concept drift robust adaptive novelty detection for data streams. *Neurocomputing* **2018**, *309*, 46–53. [\[CrossRef\]](#)
59. Spinosa, E.J.; de Leon F. de Carvalho, A.P.; Gama, J. Olindda: A cluster-based approach for detecting novelty and concept drift in data streams. In Proceedings of the 2007 ACM Symposium on Applied Computing, Seoul, Republic of Korea, 11–15 March 2007; pp. 448–452.
60. Din, S.U.; Shao, J.; Kumar, J.; Ali, W.; Liu, J.; Ye, Y. Online reliable semi-supervised learning on evolving data streams. *Inf. Sci.* **2020**, *525*, 153–171. [\[CrossRef\]](#)
61. Zheng, X.; Li, P.; Hu, X.; Yu, K. Semi-supervised classification on data streams with recurring concept drift and concept evolution. *Knowl.-Based Syst.* **2021**, *215*, 106749. [\[CrossRef\]](#)
62. Tanha, J.; Samadi, N.; Abdi, Y.; Razzaghi-Asl, N. CPSSDS: Conformal prediction for semi-supervised classification on data streams. *Inf. Sci.* **2022**, *584*, 212–234. [\[CrossRef\]](#)
63. Liao, G.; Zhang, P.; Yin, H.; Deng, X.; Li, Y.; Zhou, H.; Zhao, D. A novel semi-supervised classification approach for evolving data streams. *Expert Syst. Appl.* **2023**, *215*, 119273. [\[CrossRef\]](#)
64. Zhang, P.; Zhu, X.; Shi, Y.; Guo, L.; Wu, X. Robust ensemble learning for mining noisy data streams. *Decis. Support Syst.* **2011**, *50*, 469–479. [\[CrossRef\]](#)
65. Farid, D.M.; Zhang, L.; Hossain, A.; Rahman, C.M.; Strachan, R.; Sexton, G.; Dahal, K. An adaptive ensemble classifier for mining concept drifting data streams. *Expert Syst. Appl.* **2013**, *40*, 5895–5906. [\[CrossRef\]](#)

66. Mirza, B.; Lin, Z.; Liu, N. Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift. *Neurocomputing* **2015**, *149*, 316–329. [\[CrossRef\]](#)
67. Sun, J.; Fujita, H.; Chen, P.; Li, H. Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble. *Knowl.-Based Syst.* **2017**, *120*, 4–14. [\[CrossRef\]](#)
68. Ren, S.; Liao, B.; Zhu, W.; Li, Z.; Liu, W.; Li, K. The gradual resampling ensemble for mining imbalanced data streams with concept drift. *Neurocomputing* **2018**, *286*, 150–166. [\[CrossRef\]](#)
69. Ancy, S.; Paulraj, D. Handling imbalanced data with concept drift by applying dynamic sampling and ensemble classification model. *Comput. Commun.* **2020**, *153*, 553–560. [\[CrossRef\]](#)
70. Zyblewski, P.; Sabourin, R.; Woźniak, M. Preprocessed dynamic classifier ensemble selection for highly imbalanced drifted data streams. *Inf. Fusion* **2021**, *66*, 138–154. [\[CrossRef\]](#)
71. Alberghini, G.; Junior, S.B.; Cano, A. Adaptive ensemble of self-adjusting nearest neighbor subspaces for multi-label drifting data streams. *Neurocomputing* **2022**, *481*, 228–248. [\[CrossRef\]](#)
72. Xu, S.; Wang, J. A fast incremental extreme learning machine algorithm for data streams classification. *Expert Syst. Appl.* **2016**, *65*, 332–344. [\[CrossRef\]](#)
73. Liu, D.; Wu, Y.; Jiang, H. FP-ELM: An online sequential learning algorithm for dealing with concept drift. *Neurocomputing* **2016**, *207*, 322–334. [\[CrossRef\]](#)
74. Krawczyk, B. GPU-accelerated extreme learning machines for imbalanced data streams with concept drift. *Procedia Comput. Sci.* **2016**, *80*, 1692–1701. [\[CrossRef\]](#)
75. Mirza, B.; Lin, Z. Meta-cognitive online sequential extreme learning machine for imbalanced and concept-drifting data classification. *Neural Netw.* **2016**, *80*, 79–94. [\[CrossRef\]](#)
76. Xu, S.; Wang, J. Dynamic extreme learning machine for data stream classification. *Neurocomputing* **2017**, *238*, 433–449. [\[CrossRef\]](#)
77. Lobo, J.L.; Laña, I.; Del Ser, J.; Bilbao, M.N.; Kasabov, N. Evolving spiking neural networks for online learning over drifting data streams. *Neural Netw.* **2018**, *108*, 1–19. [\[CrossRef\]](#)
78. Liu, Z.; Loo, C.K.; Seera, M. Meta-cognitive recurrent kernel OS-ELM for concept drift handling. *Appl. Soft Comput.* **2019**, *75*, 494–507. [\[CrossRef\]](#)
79. Liu, Z.; Loo, C.K.; Pasupa, K.; Seera, M. Meta-cognitive recurrent kernel online sequential extreme learning machine with kernel adaptive filter for concept drift handling. *Eng. Appl. Artif. Intell.* **2020**, *88*, 103327. [\[CrossRef\]](#)
80. Xu, R.; Cheng, Y.; Liu, Z.; Xie, Y.; Yang, Y. Improved Long Short-Term Memory based anomaly detection with concept drift adaptive method for supporting IoT services. *Future Gener. Comput. Syst.* **2020**, *112*, 228–242. [\[CrossRef\]](#)
81. Guo, H.; Zhang, S.; Wang, W. Selective ensemble-based online adaptive deep neural networks for streaming data with concept drift. *Neural Netw.* **2021**, *142*, 437–456. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Gama, J.; Rodrigues, P.P.; Spinosa, E.; Carvalho, A. Knowledge discovery from data streams. In *Web Intelligence and Security*; IOS Press: Amsterdam, The Netherlands, 2010; pp. 125–138.
83. Gama, J.; Žliobaitė, I.; Bifet, A.; Pechenizkiy, M.; Bouchachia, A. A survey on concept drift adaptation. *ACM Comput. Surv.* **2014**, *46*, 1–37. [\[CrossRef\]](#)
84. Žliobaitė, I. Learning under concept drift: An overview. *arXiv* **2010**, arXiv:1010.4784.
85. Cui, Z.; Tian, H.; Shen, H. Effective Density-Based Concept Drift Detection for Evolving Data Streams. In Proceedings of the International Conference on Parallel and Distributed Computing: Applications and Technologies, Jeju, Republic of Korea, 16 August 2023; Springer Nature: Singapore, 2023; pp. 190–201.
86. Li, Q.; Xiong, Q.; Ji, S.; Yu, Y.; Wu, C.; Gao, M. Incremental semi-supervised extreme learning machine for mixed data stream classification. *Expert Syst. Appl.* **2021**, *185*, 115591. [\[CrossRef\]](#)
87. Barros, R.S.; Santos, S.G. A large-scale comparison of concept drift detectors. *Inf. Sci.* **2018**, *451*, 348–370. [\[CrossRef\]](#)
88. de Barros, R.S.; de Carvalho Santos, S.G.T. An overview and comprehensive comparison of ensembles for concept drift. *Inf. Fusion* **2019**, *52*, 213–244. [\[CrossRef\]](#)
89. Poenaru-Olaru, L.; Cruz, L.; van Deursen, A.; Rellermeyer, J.S. Are concept drift detectors reliable alarming systems?—A comparative study. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 3364–3373.
90. Hinder, F.; Vaquet, V.; Hammer, B. One or two things we know about concept drift—A survey on monitoring in evolving environments. Part A: Detecting concept drift. *Front. Artif. Intell.* **2024**, *7*, 1330257. [\[CrossRef\]](#)
91. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [\[CrossRef\]](#)
92. Dawid, A.P. Present position and potential developments: Some personal views statistical theory the prequential approach. *J. R. Stat. Soc. Ser. A (Gen.)* **1984**, *147*, 278–290. [\[CrossRef\]](#)
93. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
94. Liu, W.; Zhang, H.; Ding, Z.; Liu, Q.; Zhu, C. A comprehensive active learning method for multiclass imbalanced data streams with concept drift. *Knowl.-Based Syst.* **2021**, *215*, 106778. [\[CrossRef\]](#)
95. Tang, B.; He, H. KernelADASYN: Kernel based adaptive synthetic data generation for imbalanced learning. In Proceedings of the 2015 IEEE Congress on Evolutionary Computation (CEC), Sendai, Japan, 25–28 May 2015; pp. 664–671.

96. Domingos, P.; Hulten, G. Mining high-speed data streams. In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston, MA, USA, 1 August 2000; pp. 71–80.
97. Marrs, G.R.; Hickey, R.J.; Black, M.M. The impact of latency on online classification learning with concept drift. In Proceedings of the InKnowledge Science, Engineering and Management: 4th International Conference, KSEM 2010, Belfast, Northern Ireland, UK, 1–3 September 2010; Proceedings 4; Springer: Berlin/Heidelberg, Germany, 2010; pp. 459–469.
98. Liu, Z.; Hu, S.; He, X. Real-Time Safety Assessment of Dynamic Systems in Non-Stationary Environments: A Review of Methods and Techniques. In Proceedings of the 2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), Yibin, China, 22–24 September 2023; pp. 1–6. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.