ISSN: 2959-6386 (Online), Vol. 2, Issue 3

Journal of Knowledge Learning and Science Technology

journal homepage: https://jklst.org/index.php/home

Cross-Domain Applications of MLOps: From Healthcare to Finance

Prachi Tembhekar[1], Jesu Narkarunai Arasu Malaiyappan[2], Lavanya Shanmugam[3]

[1]Amazon Web Services, USA
[2]Meta Platforms Inc, USA
[3]Tata Consultancy Services, USA

Abstract

In today's digital era, the significance of data cannot be overstated. It embodies the factual and numerical essence of our everyday transactions, arriving not just statically but dynamically, in the form of data streams. These streams constitute an influx of limitless, continuous, and swift information, particularly prominent in sectors like healthcare. However, navigating this torrent of data presents formidable challenges. The sheer volume, pace, and variety make processing data streams exceedingly complex. Moreover, the task of classifying data streams is compounded by the phenomenon of concept drift, where the underlying statistical characteristics of the target variable undergo unexpected changes, especially noticeable in supervised learning scenarios.Addressing these challenges head-on, our research delves into various manifestations of concept drift within healthcare data streams. We offer an overview of established statistical and machine learning techniques tailored to tackle concept drift. Furthermore, we underscore the efficacy of deep learning algorithms in detecting concept drift and elucidate the diverse healthcare datasets employed in this endeavor.

KEYWORDS: concept drift, data stream, drift detection methods, unsupervised learning, feature (interest) point selection

ArticleInformation:
Article history: *Received:*01/12/2023 *Accepted:*10/12/2023 *Online:*30/12/2023 *Published*: 30/12/2023
DOI: https://doi.org/10.60087/jklst.vol2.n3.p598

Introduction

Machine learning (ML) encompasses a variety of methods, techniques, and tools aimed at diagnosing and prognosing medical conditions (Kralj and Kuka, 1998). It facilitates forecasting illness progression, extracting medical insights for outcome analysis, aiding in therapy planning and support, and managing patient care. ML also plays a pivotal role in data analysis, including the identification of patterns in data, effective handling of incomplete or faulty data, interpretation of continuous data from sources like Intensive Care Units, and intelligent alert systems, thus enhancing monitoring capabilities (Strausberg and Person, 1999). Successful ML approaches can streamline the integration of computer-based systems into healthcare, ultimately easing the workload of medical professionals and improving efficiency and quality of care.

Healthcare services are funded through healthcare finance, which encompasses accounting and financial management. Accounting quantifies a business's activities and finances in monetary terms, while financial management (corporate finance) utilizes theories and concepts to aid managerial decision-making. AI/ML models often predict financial costs associated with various disease severities, with concept drift occurring as illness severity changes and treatment costs evolve. Concept drift refers to the adaptation of a model due to shifts in data patterns. In this study, we explore concept drift in the context of healthcare and financial data, addressing its definition, types, and strategies for mitigation.

Our paper will cover the following:

- Defining concept drift in the healthcare domain and exploring different types of drift.
- Reviewing existing strategies for handling concept drift in the healthcare sector.
- Discussing classification techniques tailored to address concept drift effectively.

Concept drift, commonly encountered in online supervised learning scenarios, occurs when the relationship between input data and the target variable evolves over time (Gama et al., 2014). This leads to an increase in model error rates and degradation of prediction accuracy, a phenomenon often referred to as model drift or decay in AI parlance. Consequently, the model may misclassify data during classification tasks (Liu et al., 2017b), as illustrated in Figure 1.

For instance, in the healthcare domain, a model initially designed to predict treatment costs or insurance claims may undergo changes as the complexity of diseases increases over time.

Types of concept drift are depicted in Figure 2:

- Sudden: Abrupt changes in health parameter values.
- Gradual: Slow erosion of one concept by another, such as increased blood pressure leading to heart disease.
- Recurring: Changes that reappear periodically, like diabetic complications resurfacing with changes in dietary habits.
- Incremental: Gradual changes occurring over time.

Various methods can monitor the occurrence of concept drift in data streams, as illustrated in Figure 3:

- Monitoring changes in data distribution.
- Observing feature changes to predict concept drift.
- Monitoring classifier predictions.
- Tracking label changes in data generated over time.

Papers were selected based on specific inclusion criteria for applicable techniques:

- Innovation in drift detection methods or integration of drift detectors into prediction systems.
- Identification of papers related to drift detection across diverse categories from reputable journal databases.
- Inclusion of papers addressing concept drift in the healthcare sector.

The manuscript's review process for detecting concept drift in healthcare applications is detailed in Figure 4.

Background Review: Generic Framework for Concept Drift Detection

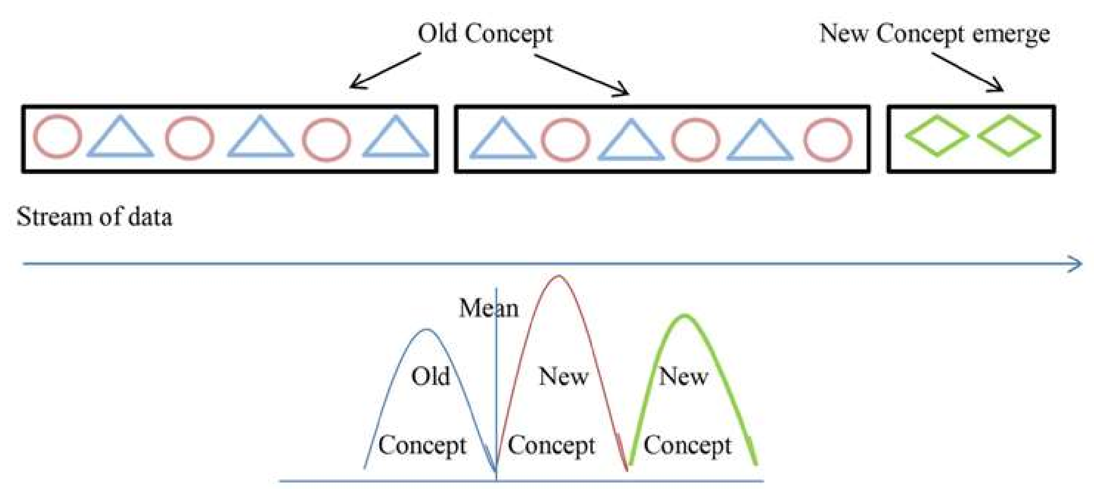

FIGURE 1
Concept drift.

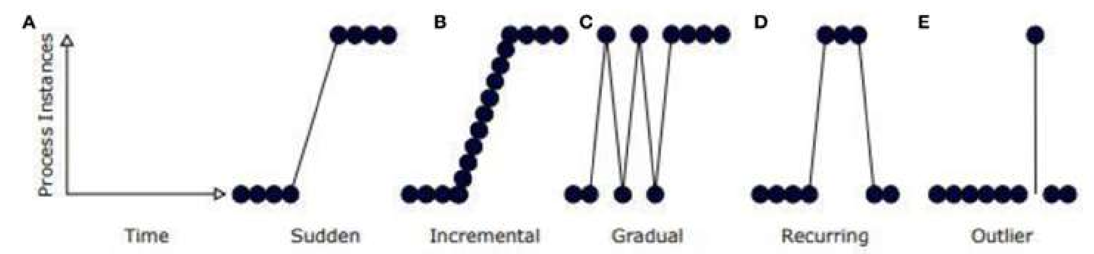

FIGURE 2
(A–E) Types of concept drift.

A generic framework for detecting concept drift typically consists of four stages, illustrated in Figure 5:

Stage 1 (Data Retrieval): In this stage, data retrieval involves extracting data chunks from data streams. Given that a single data instance may not adequately represent the overall distribution, it is crucial to organize data pieces into meaningful patterns for analysis in data stream tasks (Lu et al., 2016; Ramirez-Gallego et al., 2017).

Stage 2 (Data Modeling): Data modeling abstracts the retrieved data and identifies the key features that significantly impact a system if they undergo drift. Optionally, sample size reduction or dimensionality reduction techniques may be employed to accommodate storage and online processing speed requirements (Liu et al., 2017a).

Stage 3 (Test Statistics Calculation): This stage involves estimating the distance or dissimilarity between data instances (Test Statistics Calculation). The severity of drift is evaluated, and hypothesis test statistics are computed. Defining an accurate dissimilarity assessment remains challenging in concept drift detection. Various dissimilarity measurements can be utilized, including those for clustering evaluation and comparing sample sets (Silva et al., 2013; Dries and Ruckert, 2009).

Stage 4 (Hypothesis Test): In Stage 4 (Hypothesis Test), the statistical significance of the change identified in Stage 3 is determined using p-values. Test statistics from Stage 3 are utilized to assess the accuracy of drift detection by establishing their statistical bounds. Stage 4 is crucial for calculating the drift confidence interval, which indicates the likelihood that the observed change is attributable to concept drift rather than noise or random sample selection bias (Lu et al., 2014).

Concept Drift in Healthcare

In a multi-label hospital discharge dataset (Stiglic and Kokol, 2011) containing diagnosis information, a proposed method utilizes relative risk and phi-correlation. Monthly discharge statistics and motion charts are employed for visualization to detect concept drift. Static and dynamic ensemble classifiers are then utilized to assess accuracy and recommend the optimal classifier to use in response to concept drift.

TABLE 1 Concept drift by the probabilistic source of change (Bayram et al., 2022).

Primary term	Sudden drift	Gradual drift	Recurring drift	Incremental drift
Alternative terms	Abrupt drift	Evolutionary drift	Recurring contexts	Stepwise drift
	Concept shift			
	Revolutionary drift		Replacing drift	Development drif
	Immediate drift			

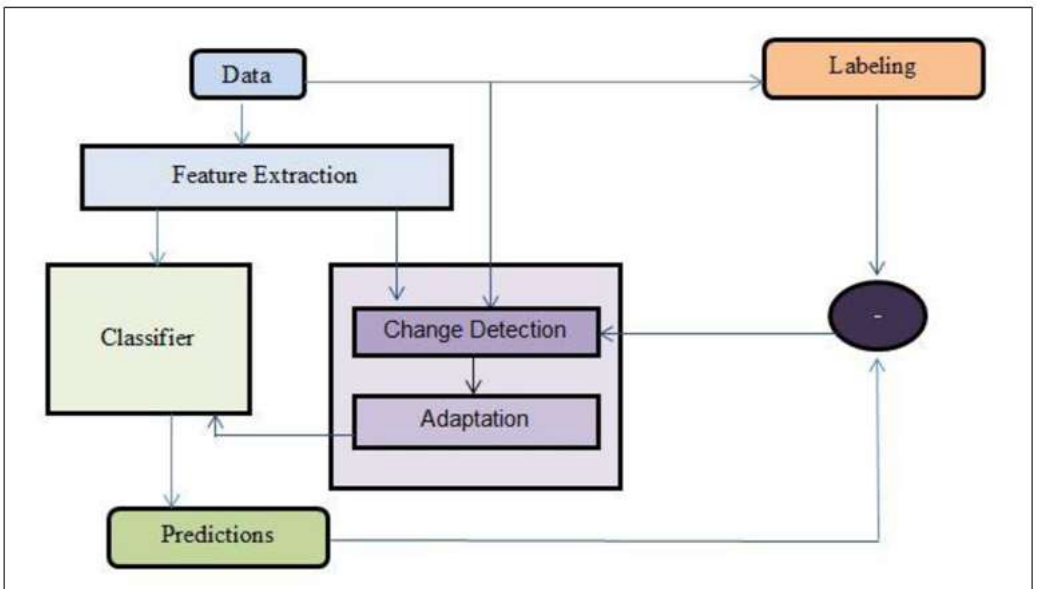

FIGURE 3

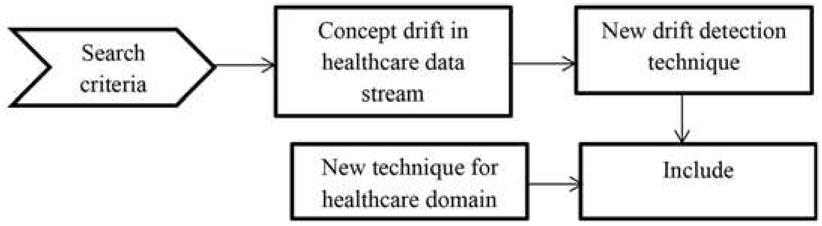

FIGURE 4
Systematic review process of concept drift.

Medical sensors designed for general healthcare or rehabilitation purposes (Toor et al., 2020) can be repurposed for ICU emergency procedures when needed. However, detecting concept drifts becomes challenging when dealing with skewed class distributions, a common occurrence in e-health data collected by medical sensors. The Reactive Drift Detection Method (RDDM) is efficient in identifying prolonged concepts but is susceptible to errors and unable to address class imbalances. In contrast, the Enhanced Reactive Drift Detection Method (ERDDM) effectively resolves concept drift in class-imbalanced data streams.

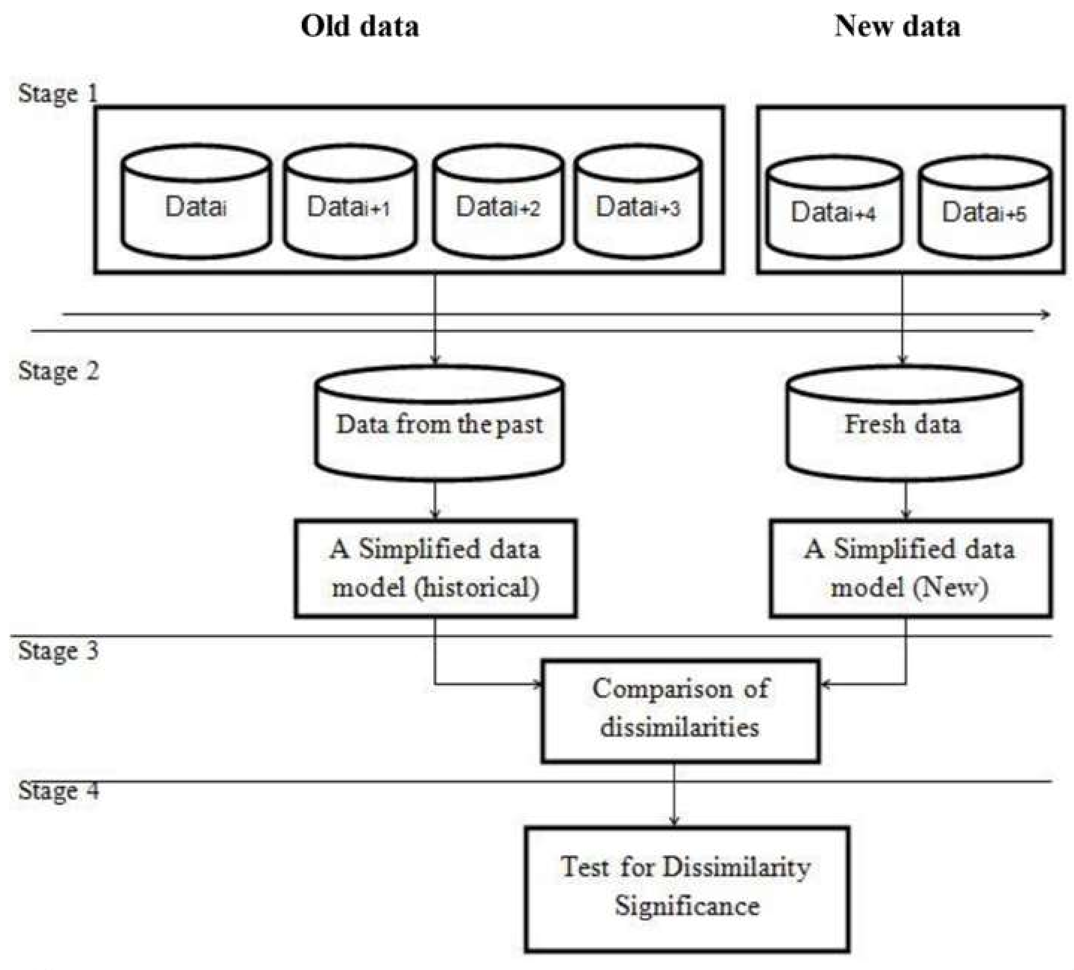

FIGURE 5
An overall framework for concept drift detection (Lu et al., 2019).

We conducted a comparison between ERDDM and three recent techniques across various metrics including prediction error, drift detection delay, latency, and data imbalance.

Clinicians utilize referral documents to triage patients referred to medical facilities (Huggard et al., 2020), which typically contain both free-text and structured data. By training a model to predict triage decisions from these documents, we aim to partially automate the triage process, enabling more efficient and systematic decision-making. However, this task requires resilience against changes in triage priorities due to factors such as policy adjustments, budget constraints, staffing variations, and other considerations. Concept drift occurs when features within the referral documents and triage labels undergo changes, necessitating model retraining to adapt to these alterations. In this domain, a unique calibrated drift detection method (CDDM) is utilized, which has demonstrated superior performance compared to state-of-the-art detectors on benchmark and simulated medical triage datasets, with fewer false positives induced by feature drift.

Calibration drift detection serves to alert users to model performance degradation (Davis et al., 2020). Our detector incorporates a robust calibration measure utilizing dynamic calibration curves, a novel approach for monitoring model performance as it evolves. An adaptive windowing strategy known as Adaptive Windowing (Adwin) monitors this calibration parameter for signs of drift as data accumulates (Wang and Abraham, 2015).

In this study, a surgical prediction model for hip replacement datasets was developed (Davis et al., 2020). Concept drift is identified by shifts in data distribution that lead to an increase in error rates and classifier performance degradation. To address concept drift in surgical prediction, a trigger-based ensemble method has been implemented, which processes each sample and rapidly adapts the model to changes in data distribution.

Concept Drift Detection Categories

Concept drift detection can be categorized into two main types:

1. Supervised
2. Unsupervised

Supervised Methods of Concept Drift Detection Categories

Performance-based Methods

Performance-based concept drift detection techniques focus on monitoring the performance of a model and identifying dips in performance as indicators of concept drift. In the context of healthcare, these techniques typically monitor vital parameter values in patients' records and trigger alerts in response to changes. The following categories of techniques are commonly employed to monitor the performance of healthcare data:

Statistical Process Control/Error Rate-based Methods

Statistical process control methods assess a model's error rates, which is particularly crucial in production environments where performance may fluctuate over time. These methods aim to issue alerts when the model's error rate surpasses predefined thresholds. Specific techniques within this category include:

1. CUSUM Test:
The CUSUM test calculates the deviation of observed values from the mean and signals concept drift if it exceeds a user-defined threshold. Concept drift is indicated when the error mean significantly deviates from zero, triggering an alert based on the log-likelihood ratio of probabilities before and after the change.

2. Drift Detection Method (DDM):
DDM models errors as a binomial random variable and monitors the error rate over time. It compares the ratio of log-likelihood between two probabilities to detect changes in error rates.

Window-based Methods

Window-based methods analyze data within sliding windows to detect changes in performance over time. These methods monitor the statistics of data windows and issue alerts when significant deviations are observed.

Sequential Analysis

Sequential analysis techniques continuously monitor data streams and analyze them sequentially to detect concept drift. These methods adaptively adjust detection thresholds based on sequential observations and signal drift when predefined criteria are met.

Ensemble Methods

Ensemble methods combine multiple models or detectors to improve detection accuracy and robustness against concept drift. By aggregating predictions or drift signals from multiple sources, ensemble methods enhance the reliability of concept drift detection in dynamic environments.

References	Drift detection methods	Health care datasets	Hypothesis test	Pros	Cons
Stiglic and Kokol(2011)	Motion charts for drift detection in the detection of sudden concept drift	NHDS data	Relative risk andphi-correlation	Visualization helps select abnormally dynamic features	Detects only one type of drift. i.e., (sudden drift)
Toor et al. (2020)	Enhanced Reactive DriftDetection Method (ERDDM)	Medical Sensors measuring for general healthcare or rehabilitation.	–	To address the class imbalance, SMOTE wasused.	Handles abrupt and gradual drifts.
Huggard et al. (2020)	calibrated drift detection method (CDDM)	benchmark and synthetic medical triage datasets	Nemenyi *post-hoc* test to compare the detection methods.	CDDM is h less prone to false positives.	A single system that can handle all changes in triage priorities.
Davis et al. (2020)	Adaptive windowing (Adwin)	Department of Veterans Affairs (VA)	–	Calibration curves show model performance over anticipated probability ranges. Loess smoothing or logistic regression is used to produce these curves. Addresses all types of drift	Accuracy in the detection of different types of drift can be improved.
Beyene et al. (2015)	Trigger-based Ensemble (TBE)	hip-replacement dataset	Nemenyi *post-hoc* test	Automate the prediction task of surgery.	The ensemble size does not become overly large

Page–Hinckley Test:

The Page–Hinckley Test detects abrupt changes in the average of a Gaussian signal, and the detection process involves running two tests in parallel, testing between the hypotheses of no change (H0: $r \leqslant n$) and change (H1: $r > n$). To detect an increase in average, the test calculates the probability range (R) and confidence (δ), where R is typically 1 for probability and log c for information gain (where c is the number of classes). For a decrease in average, the calculation involves a similar process but with different parameters. Variants in the HDDM family include:

1. Hoeffding Drift Detection Method (HDDM): HDDM enhances DDM by utilizing Hoeffding's inequality to identify significant alterations in the performance estimate's moving average. An A_test considers the difference between moving averages and estimates the error given a predefined significant level.

2. Hoeffding Drift Detection Method with Weighted Moving Averages (HDDMW): This method employs a broader statistical test for weighted moving averages, giving greater weight to current real values over older ones for quicker

detection.

3. Fast Hoeffding Drift Detection Method (FHDDM): FHDDM utilizes a sliding window and Hoeffding's inequality to compare the highest probability of correct predictions with the most recent probability to detect drift.

4. Stacking Fast Hoeffding Drift Detection Method (FHDDMS): FHDDMS expands upon FHDDM by maintaining windows of various sizes, combining short and long sliding windows to reduce false negatives and detection delays.

5. Additive FHDDMS (FHDDMSadd): FHDDMSadd detects abrupt concept drifts with shorter delays and reduces false negatives for gradual drifts.

6. Exponentially Weighted Moving Average (EWMA): EWMA adapts EWMA charts to detect changes in classifier error rates by computing the dynamic standard deviation.

Sequential Analysis-Based Methods

Sequential analysis methods inspect data instances one after another to detect changes in data distribution beyond predefined thresholds, signaling drift when observed. These methods monitor the contingency table's four rates (deviation Zt and error rate pˆ) to signal concept drift. FP-ELM and OS-ELM are notable techniques within this category, providing incremental and online learning capabilities.

Window-Based Methods

Window-based methods group incoming data into batches or windows, either fixed or adjustable in size, to analyze changes over time. These methods adapt window sizes based on drift conditions, shrinking windows in the presence of drift and widening them otherwise. From a healthcare perspective, windows can represent recordings of patient details over time, facilitating the monitoring of changes in their condition.

Adaptive Windowing (ADWIN and ADWIN2):

Adaptive windowing, proposed by Bifet and Gavalda (2007), compares the distribution between all partitions of two windows to detect concept drift. Each partition's mean error rate is calculated, and if the absolute difference exceeds a threshold based on the Hoeffding bound, the last element in the window is dropped.

Wilcoxon Rank Sum Test Drift Detector (WSTD):

WSTD, introduced by de Barros et al. (2018), utilizes a two-window approach for concept drift detection, employing the Wilcoxon rank sum statistical test instead of traditional methods. The test equation for WSTD involves calculating the difference in means and comparing it to a predefined threshold.

Detection Method Using Statistical Testing (STEPD):

STEPD relies on statistical testing to detect concept drift, assuming that recent accuracy significantly lower than overall accuracy indicates a changing concept. Nishida and Yamauchi (2007) developed a detection approach using a statistical test of equal proportions.

Fishers Exact Test:

Three approaches—Fisher Test Drift Detector (FTDD), Fisher Square Drift Detector (FSDD), and Fisher Proportions Drift Detector (FPDD)—employ Fisher's exact test to detect concept drift (de Lima Cabral and de Barros, 2018).

Cosine Similarity Drift Detection (CSDD):

CSDD uses cosine similarity to detect drift, computing confusion matrices using Positive Predictive Value (PPV) and False Discovery Rate (FDR) rates instead of traditional metrics (Hidalgo et al., 2019).

McDiarmid Drift Detection Method (MDDM):

MDDM utilizes McDiarmid's inequality to detect concept drift, weighting recent entries more heavily to highlight their importance (Pesaranghader et al., 2017).

Margin Density Drift Detection (MD3):

MD3 signals drift based on the number of samples mapped to a classifier's uncertainty zone (Sethi and Kantardzic, 2015).

Kolmogorov–Smirnov Test (KS Test):

KSWIN employs the Kolmogorov–Smirnov (KS) statistical test to detect concept drift without assumptions about data distribution (unlike traditional methods). KS test is run on identically sized windows, comparing the empirical cumulative distribution distance.

Identifying Concept Drift

When R and W are drawn from the same distribution, a notable disparity in empirical data distributions between them may signal concept drift.

Raza et al. (2015) introduce innovative covariate shift-detection techniques employing a two-stage structure for both univariate and multivariate time-series analysis. The initial stage employs an exponentially weighted moving average (EWMA) model-based control chart in online mode to detect covariate shift-points in non-stationary time-series. The subsequent step validates the shift detection using the Kolmogorov-Smirnov (K-S) statistical hypothesis test for univariate time-series and the Hotelling T-Squared test for multivariate time-series.

CIDD-ADODNN Deep Learning Framework

The CIDD-ADODNN model, as proposed by Priya and Uthra (2021), efficiently classifies extremely imbalanced streaming data. It leverages the adaptive synthetic (ADASYN) methodology to weigh minority class examples based on learning difficulties, effectively handling class imbalanced data. Concept drift detection is facilitated through an adaptive sliding window mechanism (ADWIN).

Concept Drift Adaptation using Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are employed by Saurav et al. (2018) to detect anomalies in time series data. With the gradual addition of new data, RNNs can adapt to changes in data distribution. Anomalies and change points are identified using RNN predictions of the time series, with significant prediction errors indicating deviant behavior.

Ensemble Methods

The architecture of the ensemble method is illustrated in Figure 7. This architecture consists of n models, combining the predictions of all models to produce the output. In the healthcare domain,

Streaming Ensemble Algorithm (SEA)

SEA (Street and Kim, 2001) automates drift management by incorporating a new learner for each incoming data chunk until reaching the maximum number of learners. Learners are iteratively enhanced based on their performance with predictions.

Accuracy-Weighted Ensemble (AWE)

AWE (Wang et al., 2003) comprises a group of classification models, with each model meticulously weighted according to its projected accuracy in classifying test data within a dynamic environment. This ensemble strategy ensures effectiveness and resilience to concept-drifting streams.

Accuracy Updated Ensemble (AUE)

AUE (Brzezin´ski and Stefanowski, 2011), an advancement of AWE, employs online component classifiers updated in line with the current distribution, thereby enhancing accuracy. Additional adjustments in weighting functions mitigate unintended classifier discarding, a concern observed in AWE.

Dynamic Weighted Majority (DWM)

DWM (Kolter and Maloof, 2007) employs four strategies to combat concept drift: development of online learners, weighting based on performance, disposal of underperforming learners, and addition of new experts based on overall ensemble performance.

Learn++.NSE

In Learn++.NSE (Polikar et al., 2001), a group of learners is trained using examples from data chunks. Training instance weighting depends on ensemble error for each case. Instances are assigned a weight of 1 if correctly classified by the ensemble, otherwise penalized to $wi = 1/e$. Learner ensembles are weighted using a sigmoid function based on mistakes in previous and current chunks.

Adaptive Random Forest (ARF)

ARF (Gomes et al., 2017) employs efficient resampling and adaptive operators to address concept drifts without the need for dataset optimization.

DDD (Diversity-Driven Detection)

DDD regulates learner diversity by incorporating low and high-diversity ensembles, selecting the appropriate ensemble after drift detection.

DDE (Detector-Directed Ensemble)

DDE (Bruno Maciel et al., 2015) orchestrates a small ensemble to manage signals from three drift detectors at both warning and drift levels. Sensitivity settings determine the number of detectors required to confirm an alarm or drift level, with each setting having a corresponding default detector setup.

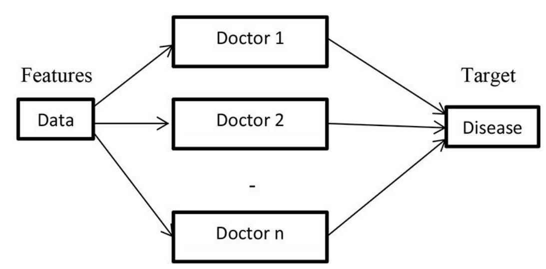

FIGURE 7
Ensemble methods.

Data Distribution Methods

This approach to drift detection assesses contextual shifts by comparing recent examples with previous data instances. Typically, these methods analyze statistical significance and are often employed alongside window-based approaches. Determining the location of drift involves computing changes in data distribution, which may incur computational costs, as depicted in Figure 8.

From a healthcare standpoint, the distribution of current vital health parameters is compared with those from previous days, weeks, or hours, both pre-diagnosis and post-diagnosis. Health parameters are meticulously observed and studied, allowing doctors to adjust medication or decide on patient discharge based on the distributions of these parameters.

SyncStream

SyncStream (Shao et al., 2014) is a prototype-based categorization model designed for evolving data streams. It dynamically models shifting concepts and provides local predictions by maintaining a collection of prototypes in a new data structure called the P-tree. SyncStream captures evolving notions through limited clustering and error-driven representativeness learning, with heuristics based on PCA and statistics for detecting abrupt concept drift.

PCA-Based Change Detection Framework

The methodology proposed by Qahtan et al. (2015) is based on estimating data for a subset of key components. Densities in reference and test windows are estimated and compared for each projection. Subsequently, divergence measures determine change-score values, with the largest change score from multiple principal components considered as the final change score, assigning equal weight to all selected principal components.

Hellinger Distance Threshold Technique

Ditzler and Polikar (2011) introduced a technique that utilizes an adjustable threshold to compute the Hellinger distance, serving as a measure to determine drift between two batches of training data.

Least Squares Density-Difference Estimation Method

Bu et al. (2018) devised a novel test for detecting changes without relying on the probability density function. This method, based on the least squares density-difference estimation method, works online with multidimensional inputs and employs a reservoir sampling mechanism. It automatically derives the required thresholds for change detection once the false positive rate is established by the application designer.

Local Drift Degree (LDD)

Liu et al. (2017a) introduced the concept of local drift degree (LDD), which measures alterations in local density over time. LDD synchronizes regional density disparities after a drift, rather than suspending all historical data.

Reactive Robust Soft Learning Vector Quantization (RRSLVQ)

RRSLVQ (Raab et al., 2020) combines the Kolmogorov–Smirnov (KS) test with the Robust Soft Learning Vector Quantization (RSLVQ) to detect concept drift.

Healthcare Datasets

Concept drift detection in healthcare often relies on diverse datasets, including:

• National Hospital Discharge Survey (NHDS) data: This dataset comprises hospital discharge records from approximately 1% of hospitals across the United States.

• MIMIC-III (Medical Information Mart for Intensive Care) (Johnson et al., 2016): MIMIC-III is a comprehensive critical care database accessible to the public. It contains extensive information on patients admitted to critical care units at a large tertiary care hospital, encompassing vital signs, medications, laboratory measurements, care providers' notes, fluid balance, procedure codes, diagnostic codes, imaging reports, hospital length of stay, survival data, and more.

• Veterans Health Administration (VHA): As one of the three administrations within the Department of Veterans Affairs (VA), VHA represents the largest integrated health system in the United States.

• Hip-replacement dataset: This dataset, collected from the orthopedics department of Blekinge Hospital, contains information related to patients undergoing hip replacement procedures.

• Il Paese Ritrovato Dataset: This dataset originates from a healthcare facility in Monza designed for the residential care of individuals affected by Alzheimer's disease.

Future Research Prospects

There are several promising directions for future research in healthcare and concept drift detection:

Drug Manufacturing Process: Changes in drug formulations or manufacturing processes can have significant economic implications for pharmaceutical companies. Early detection of these changes could inform decision-making regarding production adjustments.

Monitoring Health Parameters During Surgery: Real-time monitoring of patients' health parameters during surgical procedures could provide valuable insights for timely diagnosis and intervention, potentially improving patient outcomes.

Pandemic Disease Information: Recent events like the COVID-19 pandemic have highlighted the importance of early detection and response to infectious diseases. Research focusing on early detection and monitoring of pandemic diseases could help mitigate their impact on public health systems.

Deep Learning Techniques for Concept Drift: Exploring new deep learning techniques tailored to address concept drift in healthcare data could lead to more effective and robust models. These techniques could enhance the accuracy and reliability of predictive models in dynamic healthcare environments.

Conclusion

This study delved into drift-handling algorithms tailored for healthcare applications, focusing on both supervised and unsupervised learning tasks. Supervised learning tasks involve mapping feature instances to target variables, while unsupervised tasks lack class labels. The article examined machine learning algorithms designed to manage concept drift in medical domain problems, exploring various types of concept drift and strategies for addressing them through implicit and explicit approaches. Future research can further incorporate different techniques for handling concept drift in the medical sector, offering promising avenues for advancement.

References List

[1]. Shuford, J. (2023). Contribution of Artificial Intelligence in Improving Accessibility for Individuals with Disabilities. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(2), 421-433.
DOI: https://doi.org/10.60087/jklst.vol2.n2.p433

[2]. Chentha, A. K., Sreeja, T. M., Hanno, R., Purushotham, S. M. A., & Gandrapu, B. B. (2013). A Review of the Association between Obesity and Depression. Int J Biol Med Res, 4(3), 3520-3522.

[3]. Gadde, S. S., & Kalli, V. D. R. (2020). Descriptive analysis of machine learning and its application in healthcare. *Int J Comp Sci Trends Technol*, *8*(2), 189-196.

[4]. Atacho, C. N. P. (2023). A Community-Based Approach to Flood Vulnerability Assessment: The Case of El Cardón Sector. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 434-482.
DOI:https://doi.org/10.60087/jklst.vol2.n2.p482

[5]. jimmy, fnu. (2023). Understanding Ransomware Attacks: Trends and Prevention Strategies. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(1), 180-210. https://doi.org/10.60087/jklst.vol2.n1.p214

[6]. Bayani, S. V., Prakash, S., & Malaiyappan, J. N. A. (2023). Unifying Assurance A Framework for Ensuring Cloud Compliance in AIML Deployment. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(3), 457-472. DOI: https://doi.org/10.60087/jklst.vol2.n3.p472

[7]. Bayani, S. V., Prakash, S., & Shanmugam, L. (2023). Data Guardianship: Safeguarding Compliance in AI/ML Cloud Ecosystems. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(3), 436-456.
DOI: https://doi.org/10.60087/jklst.vol2.n3.p456

[8]. Karamthulla, M. J., Malaiyappan, J. N. A., & Prakash, S. (2023). AI-powered Self-healing Systems for Fault Tolerant Platform Engineering: Case Studies and Challenges. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 327-338. DOI:

https://doi.org/10.60087/jklst.vol2.n2.p338

[9]. Prakash, S., Venkatasubbu, S., & Konidena, B. K. (2023). Unlocking Insights: AI/ML Applications in Regulatory Reporting for US Banks. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *1*(1), 177-184. DOI: https://doi.org/10.60087/jklst.vol1.n1.p184

[10]. Prakash, S., Venkatasubbu, S., & Konidena, B. K. (2023). From Burden to Advantage: Leveraging AI/ML for Regulatory Reporting in US Banking. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *1*(1), 167-176. DOI: https://doi.org/10.60087/jklst.vol1.n1.p176

[11]. Prakash, S., Venkatasubbu, S., & Konidena, B. K. (2022). Streamlining Regulatory Reporting in US Banking: A Deep Dive into AI/ML Solutions. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *1*(1), 148-166. DOI: https://doi.org/10.60087/jklst.vol1.n1.p166

[12]. Tomar, M., & Jeyaraman, J. (2023). Reference Data Management: A Cornerstone of Financial Data Integrity. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(1), 137-144. DOI: https://doi.org/10.60087/jklst.vol2.n1.p144

[13]. Tomar, M., & Periyasamy, V. (2023). The Role of Reference Data in Financial Data Analysis: Challenges and Opportunities. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 1(1), 90-99.
DOI: https://doi.org/10.60087/jklst.vol1.n1.p99

[14]. Tomar, M., & Periyasamy, V. (2023). Leveraging Advanced Analytics for Reference Data Analysis in Finance. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(1), 128-136.
DOI: https://doi.org/10.60087/jklst.vol2.n1.p136

[15]. Sharma, K. K., Tomar, M., & Tadimarri, A. (2023). Unlocking Sales Potential: How AI Revolutionizes Marketing Strategies. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 231-250.
DOI: https://doi.org/10.60087/jklst.vol2.n2.p250

[16]. Sharma, K. K., Tomar, M., & Tadimarri, A. (2023). Optimizing Sales Funnel Efficiency: Deep Learning Techniques for Lead Scoring. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 261-274.
DOI: https://doi.org/10.60087/jklst.vol2.n2.p274

[17]. Shanmugam, L., Tillu, R., & Tomar, M. (2023). Federated Learning Architecture: Design, Implementation, and Challenges in Distributed AI Systems. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 371-384.
DOI: https://doi.org/10.60087/jklst.vol2.n2.p384

[18]. Sharma, K. K., Tomar, M., & Tadimarri, A. (2023). AI-driven Marketing: Transforming Sales Processes for Success in the Digital Age. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *2*(2), 250-260.
DOI: https://doi.org/10.60087/jklst.vol2.n2.p260

[19]. Gadde, S. S., & Kalli, V. D. (2021). The Resemblance of Library and Information Science with Medical Science. International Journal for Research in Applied Science & Engineering Technology, 11(9), 323-327.

[20]. Gadde, S. S., & Kalli, V. D. R. (2020). Technology Engineering for Medical Devices-A Lean Manufacturing Plant Viewpoint. Technology, 9(4).

[21]. Gadde, S. S., & Kalli, V. D. R. (2020). Medical Device Qualification Use. International Journal of Advanced Research in Computer and Communication Engineering, 9(4), 50-55.

[22]. Gadde, S. S., & Kalli, V. D. R. (2020). Artificial Intelligence To Detect Heart Rate Variability. *International Journal of Engineering Trends and Applications*, *7*(3), 6-10.

[23]. Chentha, A. K., Sreeja, T. M., Hanno, R., Purushotham, S. M. A., & Gandrapu, B. B. (2013). A Review of the Association between Obesity and Depression. *Int J Biol Med Res*, *4*(3), 3520-3522.

[24]. Tao, Y. (2022). Algorithm-architecture co-design for domain-specific accelerators in communication and artificial intelligence (Doctoral dissertation).
https://deepblue.lib.umich.edu/handle/2027.42/172593

[25]. Tao, Y., Cho, S. G., & Zhang, Z. (2020). A configurable successive-cancellation list polar decoder using split-tree architecture. IEEE Journal of Solid-State Circuits, 56(2), 612-623.
DOI: https://doi.org/10.1109/JSSC.2020.3005763

[26]. Tao, Y., & Choi, C. (2022, May). High-Throughput Split-Tree Architecture for Nonbinary SCL Polar Decoder. In 2022 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 2057-2061). IEEE.
DOI: https://doi.org/10.1109/ISCAS48785.2022.9937445

[27]. Tao, Y. (2022). Algorithm-architecture co-design for domain-specific accelerators in communication and artificial intelligence (Doctoral dissertation).
https://deepblue.lib.umich.edu/handle/2027.42/172593

[28]. Mahalingam, H., Velupillai Meikandan, P., Thenmozhi, K., Moria, K. M., Lakshmi, C., Chidambaram, N., & Amirtharajan, R. (2023). Neural attractor-based adaptive key generator with DNA-coded security and privacy framework for multimedia data in cloud environments. Mathematics, 11(8), 1769.
https://doi.org/10.3390/math11081769

[29]. Padmapriya, V. M., Thenmozhi, K., Praveenkumar, P., & Amirtharajan, R. (2020). ECC joins first time with SC-FDMA for Mission "security". Multimedia Tools and Applications, 79(25), 17945-17967.
DOI https://doi.org/10.1007/s11042-020-08610-5

[30]. Padmapriya, V. M. (2018). Image transmission in 4g lte using dwt based sc-fdma system. Biomedical & Pharmacology Journal, 11(3), 1633.
DOI : https://dx.doi.org/10.13005/bpj/1531

[31]. Padmapriya, V. M., Priyanka, M., Shruthy, K. S., Shanmukh, S., Thenmozhi, K., & Amirtharajan, R. (2019, March). Chaos aided audio secure communication over SC-FDMA system. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) (pp. 1-5). IEEE.
https://doi.org/10.1109/ViTECoN.2019.8899413

[31]. Padmapriya, V. M., Thenmozhi, K., Praveenkumar, P., & Amirtharajan, R. (2022). Misconstrued voice on SC-FDMA for secured comprehension-a cooperative influence of DWT and ECC. Multimedia Tools and Applications, 81(5), 7201-7217.
DOI https://doi.org/10.1007/s11042-022-11996-z

[32]. Padmapriya, V. M., Sowmya, B., Sumanjali, M., & Jayapalan, A. (2019, March). Chaotic Encryption based secure Transmission. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) (pp. 1-5). IEEE.
DOI https://doi.org/10.1109/ViTECoN.2019.8899588

[33]. Sowmya, B., Padmapriya, V. M., Sivaraman, R., Rengarajan, A., Rajagopalan, S., & Upadhyay, H. N. (2021). Design and Implementation of Chao-Cryptic Architecture on FPGA for Secure Audio Communication. In Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020, Volume 3 (pp. 135-144). Springer Singapore
https://link.springer.com/chapter/10.1007/978-981-15-9774-9_13

[34]. Padmapriya, V. M., Thenmozhi, K., Avila, J., Amirtharajan, R., & Praveenkumar, P. (2020). Real Time Authenticated Spectrum Access and Encrypted Image Transmission via Cloud Enabled Fusion centre. Wireless Personal Communications, 115, 2127-2148.
DOI https://doi.org/10.1007/s11277-020-07674-8

[35]. Kommaraju, V., Gunasekaran, K., Li, K., Bansal, T., McCallum, A., Williams, I., & Istrate, A. M. (2020). Unsupervised pre-training for biomedical question answering. arXiv preprint arXiv:2009.12952.

[36]. Bansal, T., Gunasekaran, K., Wang, T., Munkhdalai, T., & McCallum, A. (2021). Diverse distributions of self-supervised tasks for meta-learning in NLP. arXiv preprint arXiv:2111.01322.

[37]. Gunasekaran, K., Tiwari, K., & Acharya, R. (2023, June). Utilizing deep learning for automated tuning of database management systems. In 2023 International Conference on Communications, Computing and Artificial Intelligence (CCCAI) (pp. 75-81). IEEE.

[38]. Gunasekaran, K. P. (2023, May). Ultra sharp: Study of single image super resolution using residual dense network. In 2023 IEEE 3rd International Conference on Computer Communication and Artificial Intelligence (CCAI) (pp. 261-266). IEEE.

[39]. Gillespie, A., Yirsaw, A., Gunasekaran, K. P., Smith, T. P., Bickhart, D. M., Turley, M., ... & Baldwin, C. L. (2021). Characterization of the domestic goat γδ T cell receptor gene loci and gene usage. Immunogenetics, 73, 187-201.

[40]. Yirsaw, A. W., Gillespie, A., Zhang, F., Smith, T. P., Bickhart, D. M., Gunasekaran, K. P., ... & Baldwin, C. L. (2022). Defining the caprine γδ T cell WC1 multigenic array and evaluation of its expressed sequences and gene structure conservation among goat breeds and relative to cattle. Immunogenetics, 74(3), 347-365.

[41]. Gunasekaran, K. P., Babrich, B. C., Shirodkar, S., & Hwang, H. (2023, August). Text2Time: Transformer-based Article Time Period Prediction. In 2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI) (pp. 449-455). IEEE.

[42]. Gunasekaran, K., & Jaiman, N. (2023, August). Now you see me: Robust approach to partial occlusions. In 2023 IEEE 4th International Conference on Pattern Recognition and Machine Learning (PRML) (pp. 168-175). IEEE.

[43]. Gillespie, A., Yirsaw, A., Kim, S., Wilson, K., McLaughlin, J., Madigan, M., ... & Baldwin, C. L. (2021). Gene characterization and expression of the γδ T cell co-receptor WC1 in sheep. Developmental & Comparative Immunology, 116, 103911.

[44]. Gunasekaran, K. P. (2023). Leveraging object detection for the identification of lung cancer. *arXiv preprint arXiv:2305.15813*.

[45]. Gunasekaran, K. P. (2023). Exploring sentiment analysis techniques in natural language processing: A Comprehensive Review. arXiv preprint arXiv:2305.14842.

[46]. Lee, S., Weerakoon, M., Choi, J., Zhang, M., Wang, D., & Jeon, M. (2022, July). CarM: Hierarchical episodic memory for continual learning. In Proceedings of the 59th ACM/IEEE Design Automation Conference (pp. 1147-1152).

[47]. Lee, S., Weerakoon, M., Choi, J., Zhang, M., Wang, D., & Jeon, M. (2021). Carousel Memory: Rethinking the Design of Episodic Memory for Continual Learning. arXiv preprint arXiv:2110.07276.