International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

Int. J. Sci. Res. Arch.
IJSRA
International Journal of Science and Research Archive
Research Journal Archive, INDIA

(Review Article)

Tackling data and model drift in AI: Strategies for maintaining accuracy during ML model inference

Surya Gangadhar Patchipala *

Director, Consulting Expert - Data, AI, ML Engineering

International Journal of Science and Research Archive, 2023, 10(02), 1198–1209

Publication history: Received on 19 September 2023; revised on 26 November 2023; accepted on 30 November 2023

Article DOI: https://doi.org/10.30574/ijsra.2023.10.2.0855

Abstract

In machine learning (ML) and artificial intelligence (AI), model accuracy over time is very important, particularly in dynamic environments where data and relationships change. Data and model drift pose challenging issues that this paper seeks to explore: shifts in input data distributions or underlying model structures that continuously degrade predictive performance. It analyzes different drift types in-depth, including covariate, prior probability, concept drift for dasta, parameters, hyperparameter, and algorithmic model drift. Key causes, ranging from environmental changes to evolving data sources and overfitting, contribute to decreased model reliability.

The article also discusses practical strategies for detecting and mitigating Drift, such as regular monitoring, statistical tests, and performance tracking, alongside solutions like automated recalibration, ensemble methods, and online learning models to enhance adaptability. Furthermore, the importance of feedback loops and computerized systems in handling Drift is emphasized, with real-world case studies illustrating drift impacts in financial and healthcare applications. Finally, future AI system drift management will be highlighted from emerging directions such as AI-based drift prediction, transfer learning, and robust model design.

Keywords: Data Drift; Model Drift; AI Model Accuracy; Machine Learning Drift Detection; Concept Drift

Graphical Abstract

[*] Corresponding author: Surya Gangadhar Patchipala

1. Introduction

Even in the fast-changing world of artificial intelligence (AI) and machine learning (ML), artificially intelligent (AI) models can be perturbed by data, and the model drifts to the point of dramatically decreasing its accuracy. When the environment changes (i.e., there is a change in the data patterns or relationships between variables), a drift accompanies this, and the model performance deteriorates. Picture your machine learning model as an instrument that's so perfectly tuned when it's deployed, but then the world around it begins to shift, and it loses its ability to perform exactly as you want it to.

Data and model drift happen for various reasons. User behavior, market trends, and even how data is collected could be these factors. Eventually, if allowed to continue growing out of control, Drift can corrupt the accuracy of your model. Despite this, Drift can be detected and corrected early with proactive strategies, and the model will continue to be accurate during inference.

1.1. Understanding Data Drift

Data drift occurs when the statistical properties of input data change over time. So, in simpler terms, the data your model sees in real-world use is different than the data it was trained on. For example, if you had built a model to predict customer preferences given data from 2019 and are now, in 2023, still using that same model without updates, you probably wouldn't see much benefit as user behavior, preferences, and external factors would have changed. The model needs to improve its effectiveness in making predictions.

There are many reasons data drifts. However, these can include seasonal trends like spending more during holidays or temperature trends that change user's behavior. In addition, the patterns in the data will be dramatically changed by external events like economic recessions or political changes. Additionally, as your user base grows and changes, the demographic makeup of the data might evolve, making it less reflective of the training set.

1.2. Understanding Model Drift

Model drift refers to the model no longer properly accomplishing its task. A model's ability to accurately predict is decreased over time, especially when relationships in variables the model relies on change. For instance, a model trained from past financial data may be hard to predict in the stock market since financial policies may change, technological advances may occur, and investors may change their behavior.

Second, model drift can also happen due to overfitting when the model is too finely tuned to the historical data and needs to generalize better to new data. In such a situation, the model may do well at first, but its accuracy will degrade as it is exposed to data that doesn't fit its learned patterns.

1.3. Why Drift is a Big Deal

Data and model drift can have real-world consequences. Unreliable predictions created by a drop in model performance can either mean inefficient processes or critical failures, depending on the use case. On the one hand, in high-stakes fields like healthcare or finance, decisions are made based on AI predictions, so this slight drop in accuracy can mean the wrong diagnosis, financial losses, or safety risks.

Making it crucial to stay forward, recognize, and manage Drift. Identifying Drift early allows you to take corrective actions before it causes significant issues.

2. Types of Data Drift

In this context, we refer to data drift as statistical properties of input data change affecting model performance. Data drift has different effects on different forms of machine learning models.

2.1. Covariate Drift

Covariate drift happens when input features go through a change of distribution. Suppose you have an e-commerce recommendation system trained on customer preferences in one season. If new products are introduced in another season, the patterns in customer preferences may shift, and your model may need help to make accurate recommendations. Covariate drift doesn't necessarily change the relationship between inputs and outputs, but the inputs evolve, making it harder for the model to predict outcomes reliably.

2.2. Prior Probability Drift

This form of Drift refers to shifts in the overall probability distribution of target classes. In a fraud detection model, for example, if the incidence rate of fraudulent transactions suddenly increases or decreases, this would cause prior probability drift. Your model, which may have been trained on a stable fraud rate, would have a harder time making accurate predictions due to the shifting likelihood of fraud in the population.

2.3. Concept Drift

Concept drift happens when the relationship between input and output variables changes. Unlike covariate drift, which affects only the inputs, concept drift alters the underlying logic or patterns that your model has learned. For example, that's common in predicting the stock market or weather because external factors make relationships between variables change over time. For example, the same set of economic indicators may generate different market outcomes due to the various ways investors behave or, alternately, the influence of external events.

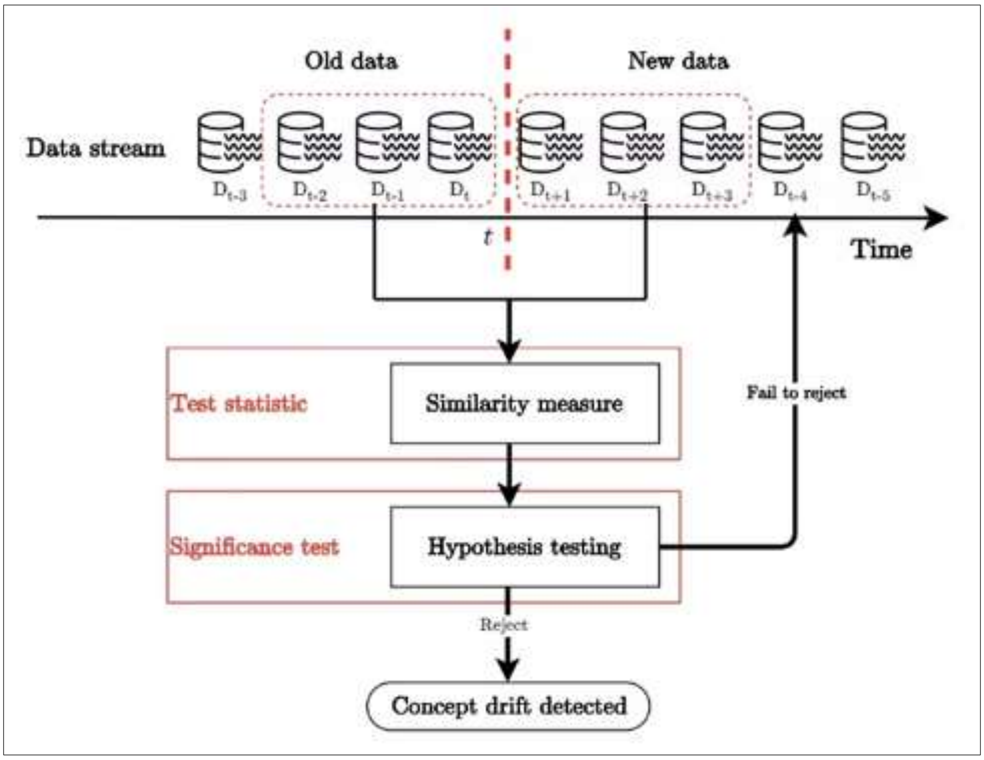

Figure 1 Concept Drift Detection Framework

3. Types of Model Drift

Model drift, on the other hand, is specific to changes in the model itself—either in its performance or in its internal configuration. Several types of model drift can affect the accuracy and reliability of machine learning systems.

3.1. Parameter Drift

Parameter drift occurs when the parameters of a model tuned out during training start to stop working as intended. For starters, these parameters were set to individually optimize the model's performance based on the training data, but as the data changes, these values no longer adequately predict. They keep model accuracy as time elapses and more "unseen" or evolving data get processed and decay.

3.2. Hyperparameter Drift

Unlike parameters, which the model learns, hyperparameters are set before training and control how the model learns. Hyperparameter drift happens when the originally selected hyperparameters no longer produce the best results because the data environment has changed. This type of Drift is often subtle and may require regular hyperparameter tuning to restore performance, especially in models deployed in dynamic settings where the nature of data changes rapidly.

3.3. Algorithmic Drift

Algorithmic Drift is a broader concept that occurs when the machine learning algorithm becomes outdated or less effective. This often happens because AI methodology advances, but it usually occurs because the data architecture underlying the data has changed in a way the algorithm can't efficiently deal with. Even when an algorithm used to detect anomalies in data used to work well, as data complexity grows, it may become inefficient and need to be switched to a more complex algorithm or architecture.

4. Causes of Data Drift

Understanding why data drift happens is key to developing viable safeguards against the effect of data drifts in machine learning models. The ambient environment can cause data Drift, a change in human behavior, or the data pipeline.

4.1. External Environmental Changes

The external environment can undergo unexpected shifts, which can majorly impact data patterns and increasingly reduce the predictability of data on which future predictions are based. For example, those businesses would likely see big swings in how consumers behave, spend, or make lifestyle decisions. Such changes can write these new data patterns so far from what training models have seen that the models can't generalize well to the new realities. Additionally, governmental policy changes, regulatory shifts, or economic downturns can introduce abrupt variations in data trends, furthering the likelihood of Drift. For example, an economic recession may alter spending or investment behaviors in financial sectors, which can mislead algorithms if these new patterns aren't detected and accounted for.

4.2. Human Behavior Shifts

Human behavior in the facing industry is dynamic and unpredictable, a large source of data drift. Historical trends in data form only part of a social trend and shifting consumer preference, which can result in new data patterns that are disconnected from the norm. Suppose an increased awareness of sustainable and eco-friendly products would change what people buy. On the other hand, if a recommendation engine in an e-commerce setting is trained on prior consumer preferences without these new sustainable tendencies, such that it fails to recommend a relevant item, inaccuracies, and poor model performance are the result. Similar shifts are common in social media, where trends evolve rapidly, and data patterns from even a few months ago may become irrelevant to current behaviors.

4.3. Changes in Data Distribution

Data drift can also occur when the process of collecting and obtaining the data changes, when the sources of the data change, or when the structure and structure of the data change. The underlying distribution of data may vary in response to modifications in the data pipeline, such as a switch in suppliers, changes to how the collection methodology worked, or data processing in new ways. A slightly different way data was collected and on which the model was trained versus the data the model sees during production can result in performance degradation. For example, label criteria for data can change the data structure, which leads to unintentional Drift. In healthcare, changing the types of health metrics collected for patient analysis may lead to shifts that impact the accuracy of a previously trained predictive model.

5. Causes of Model Drift

Model drift refers to a gradual decline in a model's performance due to various factors, from shifting variable relationships to issues within the model itself.

5.1. Changing Relationships Between Variables

As real-world relationships between variables evolve, they can impact how effectively the model makes predictions. For example, if age and income once had a strong predictive relationship in a certain model, economic or societal changes might weaken or alter that relationship. When these changes go unnoticed, models make less accurate predictions, necessitating retraining or adjustment.

5.2. Overfitting to Past Data

When a model is overly specialized to historical data (overfitting), it cannot generalize (and adapt) to new data. If you've overfitted your model, it may do well initially, but it must be more accurate as data patterns change. Overfitting can make the model rigid and less able to capture emerging trends or changes in behavior.

5.3. Degradation in Prediction Performance

Over time, all models are subject to some level of performance degradation. A model's predictive accuracy may gradually decline due to ongoing shifts in the input data or relationships between variables. This degradation can manifest slowly, often unnoticed, until the model's predictions no longer meet acceptable standards. Consistent monitoring and recalibration are needed to address and counteract this Drift.

6. The Impact of Drift on Machine Learning Models

Unchecked data or model drift can have a very negative effect on the machine-learning model. Drift can be subtle and sometimes significant, resulting in a gradual decrease in accuracy, unreliable predictions, or the model simply not working as intended.

6.1. Reduced Accuracy

One of the most immediate effects of Drift is a drop in accuracy. The model struggles to make correct predictions if data patterns change (switch) when the model-learned relationships fail to follow new inputs (generalize themselves, generalize in the wrong way). In critical applications such as fraud detection or medical diagnosis, such loss in accuracy is particularly dangerous, as predictive accuracy is vital.

6.2. Increased Prediction Errors

As drift increases, so does the frequency of prediction errors. When the model starts to behave inappropriately, errors can pile up as the years pass, eventually leading to bad decision-making. For instance, in a financial forecasting model, this might represent very large, costly misjudgments regarding economic loss and strategic errors.

6.3. Failure to Generalize to New Data

Generalizing new data well is one of the core purposes of a machine learning model. However, when there is a drift, the model breaks down in its ability to work with previously unseen data. If the input distributions change and diverge from the original training data, the model struggles to recognize new patterns and relationships, ultimately rendering it ineffective or obsolete.

7. How to Detect Data and Model Drift

It is important to detect Drift early since it keeps a model accurate and reliable. There are many ways to locate Drift before it creates problems.

7.1. Monitoring Data Distributions

Regularly monitoring the distribution of incoming data can reveal early signs of Drift. Data scientists can check for discrepancies in the distribution of incoming data relative to original training data, which might point to covariate and prior probability drift. That's especially useful for spotting those changes that can otherwise go unnoticed.

7.2. Using Statistical Tests

Differences between training and real-world data can be identified through statistical tests, such as the Kolomogorov-Smirnov test for continuous variables or the Chi-squared test for categorical variables. These tests give a quantitative sense of how much data entered differs from the data the model was trained on, allowing us to catch data drift early.

7.3. Model Performance Tracking

By tracking model performance metrics, such as accuracy, precision, recall, and F1 score, over time, it is possible to detect drift at a model level. And when they're all going down, these metrics are typically a red flag that the model is drifting and needs to be checked. Setting performance thresholds and monitoring for deviations can enable automated alerts, helping teams address issues before they escalate.

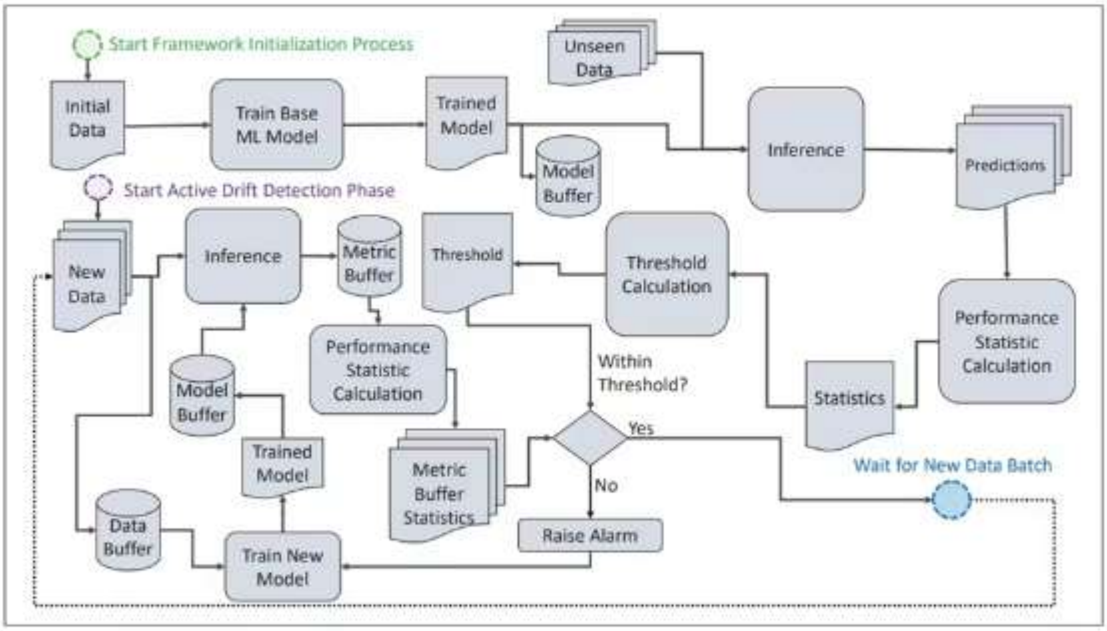

Figure 2 Drift detection and adaptation framework

8. Strategies for Tackling Model Drift

Proactive measures that maintain the accuracy and reliability of machine learning models need to be used to perform well on model drift. By implementing these solutions, you can mitigate the effects of Drift and keep your model performing well in dynamic environments.

8.1. Regular Model Recalibration

One straightforward approach to managing model drift is regular recalibration. By fine-tuning your model periodically, you can ensure it adapts to new data conditions and changes between variables. Whether adjusting parameters, retraining the model with fresh data, or updating hyperparameters, recalibration helps maintain accuracy as the environment evolves.

8.2. Using Ensemble Methods

Ensemble methods, which combine the predictions of multiple models, are another effective strategy for combating Drift. The diversity of models in an ensemble helps balance the effects of Drift on individual models. But if one model starts to fail, the other models can fill in, keeping the impact on performance at a minimum. Ensembles are especially useful in environments like the one described above, which are subject to Drift because techniques such as bagging, boosting, and stacking are used to tune predictions to be as robust as possible.

8.3. Online Learning Approaches

Online learning models are designed to update themselves as new data becomes available continuously. Online learning models, however, are different from traditional models that can be used only after periodic retraining, as they can incorporate new data on the fly and are thus very resilient to drift. This real-time learning approach helps address one of the most critical vulnerabilities in machine learning systems: drift.

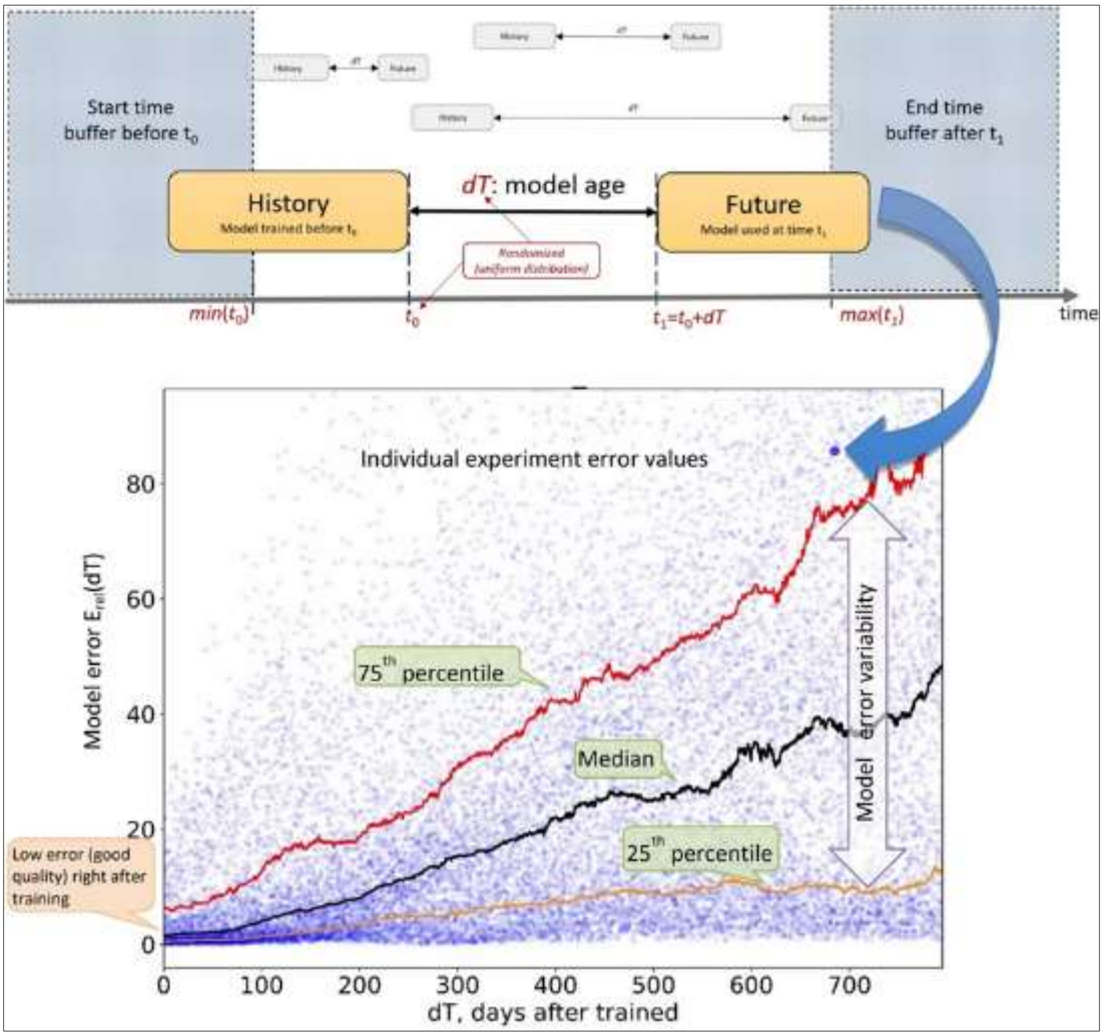

Figure 3 A Model Ageing Chart Based on a Model Degradation Experiment

9. Using Automation in Drift Detection and Correction

Automation is essential in handling Drift in large-scale deployments where drift monitoring and correction is cumbersome. It makes it faster to respond and more reliable.

9.1. Automated Monitoring Systems

Automated systems that monitor data and model drift can provide real-time insights into when and where Drift occurs. These systems track incoming data distributions and model performance metrics, alerting data scientists or triggering automatic adjustments when Drift is detected. Automated monitoring is especially helpful for early detection, allowing teams to address Drift before it leads to significant performance degradation.

9.2. Self-Healing Models

Self-healing models take automation a step further by retraining when Drift is detected. These models take on the ability to find Drift in their data or performance and trigger retraining without human intervention. Self-healing models prevent the need for manual oversight and thus minimize downtime. They also address performance issues resulting from Drift.

10. Automation in the Drift Detection and Correction

Automation is critical for managing data and model drift in large-scale machine-learning systems. Organizations can detect and resolve Drift efficiently, eliminating the need for manual oversight and speeding up response time by leveraging automated monitoring combined with self-healing models.

10.1. Automated Monitoring Systems

Automated monitoring systems provide continuous oversight of data and model drift in real time. These systems analyze incoming data distributions, comparing them with the original training data to spot deviations. When Drift is detected, they can trigger alerts, allowing data scientists to respond quickly. Advanced automated systems go further, making immediate corrections or initiating retraining processes to recalibrate the model without human intervention. This capability helps models stay accurate and responsive, addressing Drift as it arises and maintaining its performance over time.

10.2. Self-Healing Models

Self-healing models represent an advanced level of automation, where the model can autonomously retrain itself upon detecting Drift. These models monitor their performance, and when accuracy declines, they initiate corrective actions independently. By adapting to new data conditions without requiring manual recalibration, self-healing models minimize downtime and ensure ongoing accuracy and relevance, enhancing the efficiency and reliability of machine learning systems at scale.

In this immensely growing field of learning, it is important to be able to manage Drift due to its imperfections. Automated systems and self-healing models provide the essential tools that can be used at minimum cost to ensure that the models perform at peak and allow organizations to take as few actions as possible but with maximum performance and longevity.

11. Managing Drift through the Role of Feedback Loops

Effective drift management requires feedback loops to provide continuous updates using real-world data. The model has a robust feedback loop to hold it in alignment with its dynamic environment so that it can change quickly and continue to do well.

11.1. Continuous Feedback from Real-world Data

Drift is detected early by incorporating continuous feedback from the model's operating environment. By regenerating the model with real-time data, we can instantly uncover any discrepancies before they become serious. This feedback mechanism is indispensable in systems where rapid adaptation and noise removal propels the perception layer into a different state, such as recommendation engines and autonomous systems that must maintain accuracy consistently.

11.2. Human-in-the-Loop Systems

Human-in-the-loop systems add an extra layer of quality control by involving expert intervention when needed. Automated systems may detect Drift and take corrective actions, but complex cases or critical applications might still benefit from human judgment. This approach combines automation with human expertise, allowing data scientists to oversee model updates and ensure that any adjustments made by automated systems meet accuracy and reliability standards.

12. Real-World Examples of Data and Model Drift

Data and model drift are problems encountered in practical applications across various fields that necessitate immediate correction to keep model accuracy and relevance. Since machine learning (ML) models can be very sensitive to changing conditions, some industries, such as finance and healthcare, are particularly prone to Drift. Below are two case studies showing the two fields that Drift affects and the corrective action often taken to correct it.

12.1. Case Study: Drift in Financial Models

The market conditions are constantly shifting, as is the case in the finance industry, and hence, the reliability of the ML models used for trading, risk assessment, credit scoring, etc, is directly under the influence. If, for example, a trading algorithm trained on data from a bull market, defined as increasing stock prices and high trading volume, falls short in a bear market when the prices go down and liquidity dries up. The model is built to suggest trades under certain market conditions, and the prediction accuracy and the suggested trades can deteriorate in case of any significant change in market conditions. This describes a concept drift situation in which the relationship between inputs (market indicators) and outputs (buy/sell) varies with time.

Financial models are used to combat Drift and need to frequently retrain to incorporate data from various market conditions, whether a bull or bear market. Organizations sometimes apply ensemble methods or online learning approaches to allow the models to adapt in real time as conditions evolve. By regularly updating models and including new data, financial institutions can help ensure these models remain effective across varied economic landscapes.

12.2. Case Study: Drift in Healthcare Prediction Systems

Similarly, healthcare prediction systems are equally vulnerable to Drift, as they rely on patient data that changes over time as medical treatments, practices, and demographics change. For instance, a predictive model to predict which patients are at risk of having particular conditions would be trained on historical data that helped define old medical guidelines and practices. However, when new treatments or protocols are tried, such as updated guidance in treating diabetes and some cardiovascular diseases, the model may not portray patient care practice.

Moreover, changes in demographics—such as transitions in age distributions or fluctuations in health trends—can result in data drift, and poor performance is possible. Healthcare models need continuous updates with the most recent patient data and medical knowledge to maintain accuracy. These updates often involve adding new data or adjusting the model structure to accommodate more current medical practices. Organizations may also deploy feedback loops from real-world patient data to detect early signs of Drift, prompting retraining when significant changes in patient demographics or treatment protocols are detected.

13. Challenges in Managing Drift

Despite effective strategies to handle Drift in machine learning models, several challenges exist. The first step towards solving this problem is in recognizing these hurdles.

13.1. Lack of Real-Time Monitoring

Managing Drift is challenging because real-time monitoring systems that we could use don't exist. Detecting Drift early becomes easier with these systems and suffers from unnoticed performance degradation. Many organizations may rely on periodic checks or manual audits, often insufficient for identifying subtle changes in data distributions or model performance. As a result, Drift can go undetected until it significantly impacts accuracy and decision-making.

13.2. Limited Resources for Retraining

Model retraining is often time-consuming, consuming resources such as time, computing power, or human expertise. Computational resources, or even personnel, can constrain organizations' ability to retrain models regularly, making it difficult for organizations to consider a more flexible schedule. This can lead to outdated models that need to reflect current data conditions, increasing the risk of drift-related performance issues.

13.3. Complexities in Understanding Drift

It can be challenging to spot Drift, and identifying and correcting Drift in complex datasets with complex models can get harder. In particular, it can be challenging to tell when Drift has occurred and to what extent when data change is minimal, as it can greatly affect model performance in high dimensions. Additionally, the relationships between variables can change unexpectedly, complicating efforts to maintain model accuracy.

14. Future Directions for Drift Management

Consequently, researchers study novel means of addressing the crisis of drift management in the field of AI.

14.1. AI Used to Predict Drift

Predicting the likelihood of Drift is one promising development area for AI systems. These systems analyze historical data and detect trends that precede drift events so that organizations can take preventive actions before huge performance issues occur. Predictive models are a useful early warning system to improve the responsiveness of current drift management strategies.

14.2. Creating More Robust Models

Future research also focuses on creating more robust models capable of handling various data variations. Researchers wish to reduce Drift's impact by designing models that are less sensitive to particular data distributions. Regularization, dealing with adversarial training, and diverse training datasets could help us develop a more resilient model to environmental changes.

14.3. Mitigating Drift Using Transfer Learning.

Another approach to overcome this is data and model drift, which is transfer learning. Specifically, it allows a model trained on one domain to transfer its knowledge to another where the data is scarce. For instance, a model trained on data from one of the demographics may be fine-tuned to work well for another demographic to close the gap created through Drift in the training set.

15. Conclusion

In the fast and ever-evolving field of AI, it is critical to remain alert against Drift. The drift problem is inevitable, but adopting a strategic and proactive stance is essential to minimize the resulting consequences. To keep machine learning models accurate and reliable over time, they can include robust monitoring systems, employ predictive AI tools to detect early Drift, and build flexible models.

The ultimate goal is to create AI systems that are adaptable from the get-go, not just performing well when they first go live but able to thrive in the face of changing constraints of the real world. With new techniques and tools emerging, drift management research and innovation will enable AI models to remain responsive, resilient, and relevant in an ever-changing world.

Suppose you want AI systems to succeed in the unpredictable world. In that case, managing for Drift is about the immediate accuracy of your decision-making and the long-term success and robustness of your AI systems.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Analytics Vidhya. (n.d.). drift Archives. https://www.analyticsvidhya.com/blog/tag/drift/

[2] Lee, Yongsoo & Lee, Yeeun & Lee, Eungyu & Lee, Taejin. (2023). Explainable Artificial Intelligence-Based Model Drift Detection Applicable to Unsupervised Environments. Computers, Materials & Continua. 76. 1701-1719. 10.32604/cmc.2023.040235.

[3] Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Ali, A. B. M. S., & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artificial Intelligence Review, 56(11), 13521–13617. https://doi.org/10.1007/s10462-023-10466-8

[4] Manias, D.M., Chouman, A., & Shami, A. (2023). Model Drift in Dynamic Networks. IEEE Communications Magazine, 61, 78-84.

[5] Jiaoyan Chen, Freddy Lécué, Jeff Z. Pan, Shumin Deng, Huajun Chen, Knowledge graph embeddings for dealing with concept drift in machine learning, Journal of Web Semantics, Volume 67, 2021, 100625, ISSN 1570-8268, https://doi.org/10.1016/j.websem.2020.100625.

[6] S, A. R. M., R, N. C., R, S. B., Lahza, H., & Lahza, H. F. M. (2023). A survey on detecting healthcare concept drift in AI/ML models from a finance perspective. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.955314

[7] Aldoseri A, Al-Khalifa KN, Hamouda AM. Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges. Applied Sciences. 2023; 13(12):7082. https://doi.org/10.3390/app13127082

[8] Firas Bayram, Bestoun S. Ahmed, Andreas Kassler, From concept drift to model degradation: An overview on performance-aware drift detectors, Knowledge-Based Systems, Volume 245, 2022, 108632, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2022.108632.

[9] Dwork, C.; Roth, A.; Naor, M. Differential Privacy: A Survey of Results. In Theory and Applications of Models of Computation; Springer: New York, NY, USA, 2018; pp. 1–19.

[10] Truex, S.; Xu, C.; Calandrino, J.; Boneh, D. The Limitations of Differential Privacy in Practice. In Proceedings of the 28th USENIX Security Symposium, Santa Clara, CA, USA, 14–16 August 2019; pp. 1045–1062.

[11] Goodfellow, I.; Shlens, J.; Szegedy, C. Explaining and Harnessing Adversarial Examples. Commun. ACM 2022, 65, 56–65.

[12] Sina_Mobile. (2023, September 19). ChatGPT 笨了，还是老了？. 新浪移动_手机新浪网. https://finance.sina.cn/blockchain/2023-09-19/detail-imzncnqe5076884.d.html?oid=3848037900393170&vt=4&wm=4007&cid=76601&node_id=76601

[13] Akhtar, N.; Mian, A. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. IEEE Access 2018, 6, 14410–14430.

[14] Steinhardt, J.; Koh, P.W.; Liang, P. Certified Defenses against Adversarial Examples. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.

[15] Zhu, M.; Yin, H.; Yang, X. A Comprehensive Survey of Poisoning Attacks in Federated Learning. IEEE Access 2021, 9, 57427–57447.

[16] Sun, Y.; Zhang, T.; Wang, J.; Wang, X. A Survey of Deep Neural Network Backdoor Attacks and Defenses. IEEE Trans. Neural Netw. Learn. Syst. 2020, 31, 4150–4169.

[17] Gu, T.; Dolan-Gavitt, B.; Garg, S. BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain. In Proceedings of the 28th USENIX Security Symposium, Santa Clara, CA, USA, 14–16 August 2019; pp. 1965–1980.

[18] Liu, Y.; Ma, X.; Ateniese, G.; Hsu, W.L. Trojaning Attack on Neural Networks. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, Toronto, ON, Canada, 15–19 October 2018; pp. 27–41.

[19] Gao, Y.; Sun, X.; Zhang, Y.; Liu, J. Trojan Attacks on Federated Learning Systems: An Overview. IEEE Netw. 2021, 35, 144–150.

[20] Agrahari S., Singh A. K. (2021). Concept drift detection in data stream mining: a literature review. J. King Saud Univ. Comput. Inform. Sci. 34, 9523–9540. 10.1016/j.jksuci.2021.11.006

[21] Alippi C., Roveri M. (2008). Just-in-time adaptive classifiers-part I: detecting nonstationary changes. IEEE Trans. Neural Netw. 19, 1145–1153. 10.1109/TNN.2008.2000082

[22] Baena-Garcia M., del Campo-Avila J., Fidalgo R., Bifet A., Gavalda R., Morales-Bueno R. (2006). “Early drift detection method,” in Proc. 4th Int. Workshop Knowledge Discovery from Data Streams.

[23] Bayram F., Ahmed B. S., Kassler A. (2022). From concept drift to model degradation: an overview on performance-aware drift detectors. Knowledge Based Syst. 245, 108632. 10.1016/j.knosys.2022.108632

[24] Beyene A. A., Welemariam T., Persson M., Lavesson N. (2015). Improved concept drift handling in surgery prediction and other applications. Knowledge Inform. Syst. 44, 177–196. 10.1007/s10115-014-0756-9

[25] Bifet A., Gavalda R. (2007). “Learning from time-changing data with adaptive windowing,” in Proc. 2007 SIAM Int. Conf. Data Mining, SIAM 2007 (Minneapolis, MN:). 10.1137/1.9781611972771.42

[26] Bruno Maciel I. F., Silas Santos G. T. C., Barros R. S. M. (2015). “A lightweight concept drift detection ensemble,” in IEEE 27th International Conference on Tools with Artificial Intelligence (Vietri sul Mare:). 10.1109/ICTAI.2015.151

[27] Brzeziński D., Stefanowski J. (2011). “Accuracy updated ensemble for data streams with concept drift,” in Hybrid Artificial Intelligent Systems, HAIS, 2011, eds E. Corchado, M. Kurzyński, and M. Wozniak (Berlin; Heidelberg: Springer;). 10.1007/978-3-642-21222-2_19

[28] Krishna, K. (2022). Optimizing query performance in distributed NoSQL databases through adaptive indexing and data partitioning techniques. International Journal of Creative Research Thoughts (IJCRT). https://ijcrt. org/viewfulltext. php.

[29] Krishna, K., & Thakur, D. (2021). Automated Machine Learning (AutoML) for Real-Time Data Streams: Challenges and Innovations in Online Learning Algorithms. Journal of Emerging Technologies and Innovative Research (JETIR), 8(12).

[30] Murthy, P., & Thakur, D. (2022). Cross-Layer Optimization Techniques for Enhancing Consistency and Performance in Distributed NoSQL Database. International Journal of Enhanced Research in Management & Computer Applications, 35.

[31] Murthy, P., & Mehra, A. (2021). Exploring Neuromorphic Computing for Ultra-Low Latency Transaction Processing in Edge Database Architectures. Journal of Emerging Technologies and Innovative Research, 8(1), 25-26.

[32] Mehra, A. (2024). HYBRID AI MODELS: INTEGRATING SYMBOLIC REASONING WITH DEEP LEARNING FOR COMPLEX DECISION-MAKING. Journal of Emerging Technologies and Innovative Research (JETIR), Journal of Emerging Technologies and Innovative Research (JETIR), 11(8), f693-f695.

[33] Thakur, D. (2021). Federated Learning and Privacy-Preserving AI: Challenges and Solutions in Distributed Machine Learning. International Journal of All Research Education and Scientific Methods (IJARESM), 9(6), 3763-3764.