

A Critical Review on Adverse Effects of Concept Drift over Machine Learning Classification Models

Syed Muslim Jameel¹

Department of Computer and Information Sciences
Universiti Teknologi PETRONAS (UTP), Perak
Malaysia

Mobashar Rehman⁴

Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman
Perak, Malaysia

Manzoor Ahmed Hashmani², Hitham Alhussain³

Department of Computer and Information Sciences,
Center for Research in Data Science (CERDAS), UTP
High-Performance Cloud Computing Centre (HPC3), UTP
Universiti Teknologi PETRONAS (UTP), Perak, Malaysia

Arif Budiman⁵

Faculty of Computer Science
University of Indonesia
Depok, West Java
Indonesia

Abstract—Big Data (BD) is participating in the current computing revolution in a big way. Industries and organizations are utilizing their insights for Business Intelligence using Machine Learning Models (ML-Models). Deep Learning Models (DL-Models) have been proven to be a better selection than Shallow Learning Models (SL-Models). However, the dynamic characteristics of BD introduce many critical issues for DL-Models, Concept Drift (CD) is one of them. CD issue frequently appears in Online Supervised Learning environments in which data trends change over time. The problem may even worsen in the BD environment due to veracity and variability factors. Due to the CD issue, the accuracy of classification results degrades in ML-Models, which may make ML-Models not applicable. Therefore, ML-Models need to adapt quickly to changes to maintain the accuracy level of the results. In current solutions, a substantial improvement in accuracy and adaptability is needed to make ML-Models robust in a non-stationary environment. In the existing literature, the consolidated information on this issue is not available. Therefore, in this study, we have carried out a systematic critical literature review to discuss the Concept Drift taxonomy and identify the adverse effects and existing approaches to mitigate CD.

Keywords—Big data classification; machine learning; online supervised learning; concept drift; Adaptive Convolutional Neural Network Extreme Learning Machine (ACNNELM); Meta-Cognitive Online Sequential Extreme Learning Machine (MOSELM); Online Sequential Extreme Learning Machine (OSELM); Real Drift (RD); Virtual Drift (VD); Hybrid Drift (HD); Deep Learning (DL); Shallow Learning (SL); Concept Drift (CD)

I. INTRODUCTION

State of the art Big Data (BD) and Machine Learning (ML) is one of the fundamental pillars of the 4th Industrial Revolution (IR 4.0). BD generates from a variety of sources, including scientific research, finance, government, internet search, sensors, documents, image, audio and video, and others. The nature of BD is very complex due to its non-stationary characteristics (volume, velocity, variety, veracity and variability). ML approaches (specifically Deep Learning) are considered the main drivers to utilize BD for intelligence

for offline learning scenarios only. However, these approaches failed to maintain performance accuracy during online learning scenarios. One of the online learning scenarios where ML models degrade their performance accuracy due to the non-stationary environment is Concept Drift (CD)[1].

A. Taxonomy of Concept Drift

Dynamic assumptions of data (features of data changes over time) called Concept Drift [2]. The Concept Drift term in Machine Learning (ML) is being recognized as the most critical problem since many decades for traditional data and big data. Many assumptions in ML is by using static data [3]. However, this issue frequently occurs in an Online Machine Learning scenario where these dynamic conditions change frequently. Therefore, due to the addition of new features in data, ML models degrade their performance accuracy or could fail to classify or predict to correct output.

Notably, in Supervised Online ML, the model is learned through the input and output features from data of one-time span and will be likely to predict or classify the output (class category) from another time. The change in features (among both time) is due to various conditions. It could be due to the data format (variety), distribution (variability), or sources (complexity), which change over time. Another term for Concept Drift refers to the classification boundary or clustering centers that continuously change with time elapsing [04]. These conditions will adversely affect the classification performance of the model. In studies the term CD is modeled based on Bayesian decision theory for class output 'c' and input data 'X' as shown in eq (1);

$$P\left(\frac{c}{X}\right) = P(c) P\left(\frac{X}{c}\right) / P(X) \quad (1)$$

Where $P(c/X)$, $P(c)$, $P(X/c)$, and $P(X)$ are posterior, prior, conditional, and feature-based probabilities respectively [3]. The possible conditions of Concept Drift arise $P(c/X)$ undergo changes and causes the shift in the class boundary or conditional probabilities (the number of classes increase), this type of Concept Drift is referred to as Real Drift [5].

Furthermore, if the $P(X)$ (feature-wise distribution of data changes) due to insufficient or partial feature representation of existing data distribution (new additional feature adds or some feature updates) called as Virtual Drift [5]. Also, a study introduces Hybrid Drift as a condition $P(c/X)$, and $P(X)$ occurred consequently [2], as shown in Fig. 1. However, few studies discuss possible configuration pattern based on the frequency of drift, gradual drift (when the variety of concepts changes gradually), consecutive drift (when previous concepts reoccur) and sudden drift pattern (when a concept changes/substitutes abruptly) [6], [7], as shown in Fig. 2.

ML models train to classify according to input and output features with a predefined number of classes. If a feature or class-wise distribution changes over time, then ML models will face a substantial degradation in their performance (because ML models do not have prior knowledge of these changes). However, if these ML models retrain according to newly-arrived data, then they are unable to keep knowledge of the recurrent context (previous training knowledge). As shown in Fig. 3.

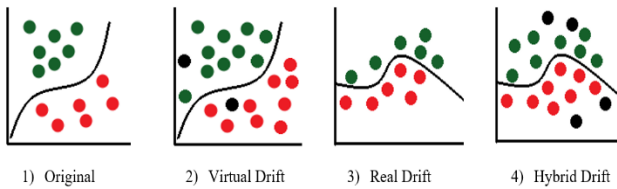


Fig. 1. Types of Concept Drift.

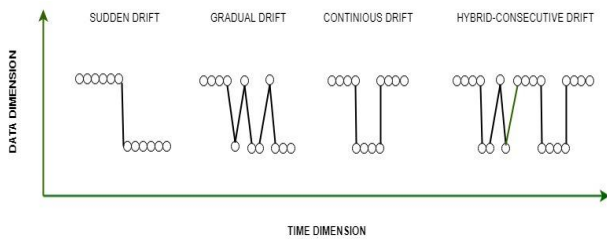


Fig. 2. Configuration Pattern of Concept Drift.

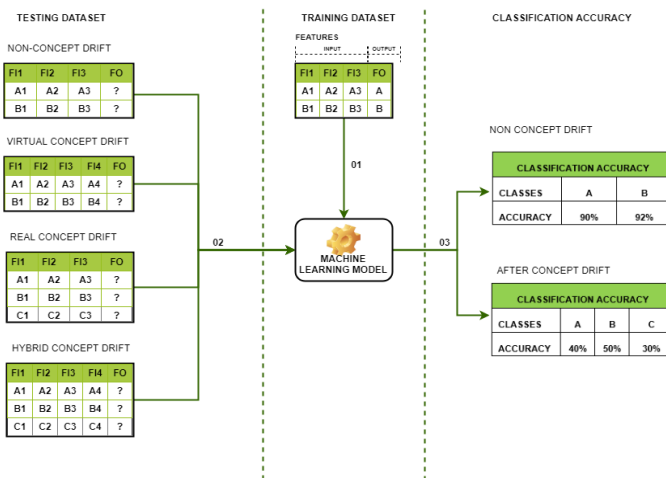


Fig. 3. Performance Degradation in the ML Model Due to Concept Drift.

B. Causes and Mitigation of Concept Drift in the Classification Problem

Zilobite I. [8], the classification problem determines through the prior probabilities $P(c_i)$ and class conditional probabilities $P(x/c_i)$ without considering the Concept Drift scenario. Zilobite I. defines the fixed set of prior probabilities of class and class-conditional as follow;

$$S = \{(P(c_1), P(\frac{x}{c_1}), P(c_2), P(\frac{x}{c_2}), \dots, (P(c_k), P(\frac{x}{c_k}), \dots))\} \quad (2)$$

Where “S” represents the data source at a given time.

Concept Drift based on Bayesian decision theory, as shown below;

$$P(c/X) = P(c) \left\{ \frac{P(\frac{X}{c})}{P(X)} \right\} \quad (3)$$

$$P(X) = \sum_c P(c)P(X/c) \quad (4)$$

The fundamental causes of the possible change in source data (S) due to $P(c/X)$ are presented by [3][9], which are mentioned below;

$P(c)$ change in class priors (Class Drift)

$P(X/c)$ class distribution permute (Virtual Drift)

$P(c/X)$ posterior distribution of class members changes (Real drift)

$P(X/c) + P(c)$, class and posterior distribution changes (Hybrid drift)

However, the mitigation strategies are not identical to each type and frequency pattern of Concept Drift (Class, Virtual, Real, continuous, gradual, and sudden or abrupt). For example, we probably like to reuse the past trained classifier if changes reappear (continuous drift pattern), or we may want to suddenly stop classifier and retrain classifier from newly detected changes (abrupt). Thus, to provide a simple approach to handle various types of CD is critical.

In recent studies, researchers propose the term “Adaptability” to avoid performance degradation due to Concept Drift in ML models. The adaptability refers to the feature of ML models (capability) to dynamically adjust itself as per the data changes. This approach allows the ML models to tune or self-regulate for new concept adjustment. Furthermore, this approach possesses the potential to eliminate performance degradation through its dynamic capabilities. However, due to the recurrent context adjustment, the practical implementation for adaptability arises many critical and fundamental questions for researchers. For example, Machine Learning can be categories as context replacement and recurrent context. Context replacement means, how ML models train new concept and forget the previous one. This is simple and can be easily incorporated through the basic adaptability features. Recurrent context refers to how ML models learn a new concept by keeping the previous one. The recurrent context is very challenging. For example, how good accuracy of the ML model can achieve for a new concept and how good it can retain with minimum re-training old data is one of the challenges. In the literature, the adaptability factor can be categorized as semi-adaptive (fundamental dynamic

changes in a classifier level) and fully-adaptive (self-regulatory and more autonomous approaches to make classifiers self-regulatory), which are defined in detail in section 02. The fact is that there are few research studies, which provide the practical implementation of recurrent context through the adaptability feature. However, these studies are particular to the type or frequency pattern of Concept Drift or type of data stream. Whereas, the studies define the framework of Concept Drift adaptation in machine learning models present the generalized framework for classifiers. This generalized framework encompasses to make the proper future assumption of data sources, detect all the possible change pattern, tune the classifier parameters or select the appropriate strategy (training, testing or feature manipulation) for specific type of Concept Drift, and optimal model selection (more appropriate model towards the target function) with minimum error rate. Through a dynamic mechanism, in the defined framework, a classifier could be able to regularly evolve and maintain its performance after any Concept Drift., as shown in Fig. 4.

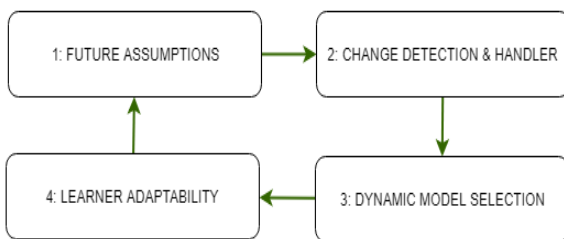


Fig. 4. The Framework for Concept Drift Learners.

II. APPROACHES TO MITIGATE THE CONCEPT DRIFT IN ML CLASSIFICATION MODELS

Multidimensional properties of Big Data (BD) and non-stationary nature of Online Supervised Learning (OSL), upfront several new challenges to handle Concept Drift (CD) issue [10]. However, the contribution from the research community to mitigate the adverse effects of CD on Big Data Classification models are rarely reported [2]. Furthermore, most existing approaches are based on Shallow Learning models (ELM, SVM, and others), hence not capable of reasonably handle the non-stationary feature of BD in the OSL scenario [1]. Moreover, several studies urged to adopt these dynamic changes (in classifier) through self-regulatory mechanisms [11] [12] [13].

Shallow Learning approaches (for example, Extreme Learning Machine (ELM), Support Vector Machine (SMV), Multi-Layer Perception Neural Network (MLP NN), Hidden Markov Model, and others.) handle classification and regression problems efficiently in structured data [14][15]. These approaches not perform well for complex unstructured data (Big Data) [1]. However, Deep Learning algorithms found a better selection to handle Big Data stream and extract value with more accuracy over conventional approaches [3]. Besides, some studies ascertained through comprehensive experiments that Deep Learning approaches are appropriate to learn from

BD and urged the researcher to explore further new means to handle CD issue due to OSL. A study argued that to find new means to handle Concept Drift in the context of Big Data and OSL is an essential task for the future of Machine Learning [10].

In literature, the issue of CD is mostly handled through the different configurations of the Extreme Learning Machine (ELM). These configurations are either based on a single classifier or ensemble classifier [8][11] [16] [17]. Ensemble classifier considers effective solution than single classifier to improve the classification performance (in terms of accuracy) after CD. Nevertheless, the ensemble approach does not adapt to the numerous drift cases [18] [19], such kind of drift may be handled through the adaptive nature of classifiers.

Few recent studies concentrated towards adaptive learning techniques using ELM based single classifier [2] [15] [19] and ensemble classifier for CD mitigation [20] [21] [22]. However, all these solutions lie in this semi-adaptive category (does not implements the fully autonomous learning behavior). For example, Incremental Data Stream ELM used an incremental approach to train the classifier. In this approach, the number of neuron in hidden layers and selection of the activation layer is dynamic, which enhance the performance of the model. Whereas, this approach handles stream data for gradual drift scenario only [22].

A Dynamic-ELM model uses ELM as a first classifier, whereas the online learning approach is adopted to train the double hidden layer structure of ELM. The improvement in the generalized characteristics of the classifier is incorporated by adding more hidden layers. This approach is capable of mitigating the CD in a short time; however, the performance of this model suffers due to the fast processing speed [21].

Meta-Cognition Online Sequential Extreme Learning Model (MOSELM) proposed for improving class imbalance (binary and multiclass) and Concept Drift for online data classification. This model is first to use Meta-Cognition principles and Online Sequential Extreme Learning Machine (OSELM) but only handle Real Drift [23]. A new adaptive windowing approach is proposed to improve adaptability in Real Drift only [15]. Online Pseudo Inverse Method (OPIUM) is based on Gravel methods, the incremental solutions to computing pseudo-inverse of a matrix. OPIUM tackles the real Concept Drift with the discriminant function boundary shift in streaming data only [19].

A recent study proposed an adaptive ML model (AOSELM) using a single classifier approach based on Online Sequential Extreme Learning Machine (OSELM) [23], and Constructive Sequential Extreme Learning Machine (COSELM) [24] to handle the Concept Drift issue for classification and regression problem. AOSELM is the simple solution used matrix adjustment. Results were satisfactory for handling Real Drift but not satisfactory to handle virtual and Hybrid Drift and did not yield better output on real data. Single classifier results may not exceed the adoptable ensemble or full batch approach due to its shared weight changes [2]. Table I represents the notable contributions (concept drift adaptation models) and highlights its pitfalls.

TABLE. I. LATEST CONCEPT DRIFT HANDLING APPROACHES

Year, Author	Method	Advantages	Disadvantages	Ref
1. Dynamic integration of classifiers for handling Concept Drift				
2008, Tsymbal, A	Ensemble Integration	Handle Virtual Drift	No-Adaptability in model Synthetic Dataset used	[16]
2. Reacting to different types of Concept Drift: The accuracy updated ensemble algorithm				
2014, Brzezinski, D	AUE	Better classification accuracy Less memory consumption	Used Synthetic data set Further, need improvement inaccuracy	[25]
3. Reacting to different types of Concept Drift with adaptive and incremental one-class classifiers				
2015, Krawczyk, Bartosz,	AW-SVM classifier	No need to drift detector Provide classification accuracy for stream data.	Further improvement needed in accuracy and adaptability.	[26]
4. Classification of uncertain data streams based on extreme learning machine				
2015, Cao, Keyan, et al	WEC-ELM	Improved efficiency and accuracy during the drift. Deal with real-time uncertain data streams.	Handle only gradual drift Need to work on high dimensional data Not viable for Big Data.	[27]
5. A fast-incremental extreme learning machine algorithm for data streams classification called IDS-ELM				
2016, Xu, Shuliang,	IDS-ELM	Improved performance Suitable for real-time data due to fast processing	Handle gradual drift in streaming Need further improvement	[22]
6. Dynamic extreme learning machine for data stream classification				
2017, Xu, Shuliang	Dynamic-ELM	Improved accuracy Fast processing speed	Accuracy suffers due to speed I need further improvement.	[21]
7. Self-Adaptive Windowing Approach for Handling Complex Concept Drift				
2015, Khamassi, Imen, et al	EDIST2	Concept Drift Detection and handling data streams. Reasonable accuracy rate at Synthetic and Real data Used Self-Adaptive Windowing approach.	Need to handle another type of drift i.e. Virtual.	[28]
8. Meta-Cognitive online sequential extreme learning machine for imbalanced and concept-drifting data classification				
2016, Mirza, Bilal	OS-ELM	Imbalanced and Concept Drift learning model Metacognition is used as self-regulatory	It only works RD drift	[17]
9. Adaptive Online Sequential ELM for Concept Drift Tackling				
2016, Budiman, Arif	AOS-ELM	Handle Real Drift, Virtual Drift, and Hybrid Drift Work well on sudden drift and recurrent Concept Drift type. The public data set is used for regression and classification problem. Good adaptability and retaining recognition	Improvement is needed in VD and HD. Work on specific dataset Need further adaptability factor	[2]
10. Adaptive Convolutional ELM for Concept Drift Handling in Online Stream Data				
2016, Budiman, Arif	CNN+ELM	A hybrid model ACNNELM proposed the integration approach of CNN with ELM. Enhanced CNN features used with ELM. Simple solution good accuracy results and work.	Results non-image data, i.e. voice, action, logging or 1D or more than 4D data. Further improvement needed to increase accuracy. Handling complex variable needed. Need to enhance adaptability factor.	[3]

III. RESULT ANALYSIS AND DEDUCTION

Through the comprehensive literature analysis, we can safely state that the performance degradation in Big Data Classification models (in terms of accuracy) due to Concept Drift is still a critical problem. The existing solutions can be categorized into as follow;

- 1) Non-adaptive and semi-adaptive (single classifier based) SL approaches.
- 2) Non-adaptive and semi-adaptive (ensemble classifier based) SL approaches.

3) Non-adaptive and semi-adaptive (single classifier based) DL approaches.

4) Non-adaptive and semi-adaptive (ensemble classifier based) DL approaches.

5) Semi-adaptive Hybrid (DL and SL) approaches.

The existing solutions are either limited to a specific type of CD, or their results are biased towards specific CD conditions or dataset. In addition to that, the classification degradation does not reasonably retain after CD handling for complex datasets (CIFAR 10), as shown in Table II. Table II demonstrates the simulations on the most prominent Big Data classification models under certain CD conditions. The experiments carried out to validate the problem formulation of

the Concept Drift issue. In this experiment, we used MNIST [29], Not-MNIST, and CIFAR 10 [30] dataset. The MNIST dataset is recognized as the benchmark dataset for the classification problem, whereas Not-MNIST is an extension of the MNIST dataset, contains some foolish images for providing some challenging data environment. CIFAR 10 is the dataset for color images. MATLAB model R2018a, using Deep Learning toolbox [31] using NVIDIA GeForce GTX 950, 768 GPU cores with 2 GB RAM. The cross-validation and holdout method is used for evaluation and the testing accuracy is measured after a specific type of CD. Interestingly, through the results, we can determine that ACNNELM is better for handling CD for MNIST and Not-MNIST dataset, whereas we found the promising testing accuracy of CNN in CIFAR10 dataset (color images). Also, in our previous study [32], we have performed several experiments to validate the Concept Drift issue.

TABLE. II. LATEST ML MODELS AND THEIR TESTING ACCURACY IN DIFFERENT IMAGE DATASETS

No	Category	ML model	Dataset	Testing Accuracy %
1	Shallow Learning (Non-adaptive)	Support Vector Machine (SVM)	MNIST	83.47
			Not-MNIST	72.65
			CIFAR 10	22.98
2	Deep Learning Model (Non-adaptive)	Convolutional Neural Network (CNN)	MNIST	84.34
			Not-MNIST	72.98
			CIFAR 10	45.67
3	Shallow Learning (Semi-adaptive)	Online Sequential ELM (OSELM)	MNIST	82.34
			Not-MNIST	73.98
			CIFAR10	38.68
4	Hybrid Deep Learning (Semi-adaptive)	Adaptive Convolutional ELM (ACNNELM)	MNIST	90.45
			Not-MNIST	78.24
			CIFAR 10	37.56

IV. CONCLUSION

Concept Drift issue can be handled by improving the ML model accuracy and enhancing the adaptability factor. The adaptability feature talks about how an ML model capable of retaining its previous training data knowledge. The ultimate goal of improvement in the ML model and adaptability is handling CD issues, whereas adaptability in the ML model can reduce the computational processing and training time too. According to literature, adaptability can be classified into two types; semi-adaptive (less adaptability) and self-regulatory (a more general aspect of autonomous learning). Current solutions to handle Concept Drift either handle image data or stream data. However, these data classification only provides non-adaptive or semi adaptive solutions (which restrict to utilizing the complete essence of adaptability factor and to handling Concept Drift in Big Data environment). Some possible research directions to overcome CD are; to investigate and formulate the relationship between Concept Drift (Big Data) and exiting Machine Learning models. Quantification and characterization of Concept Drift for Big Data streams. Propose a framework for fully adaptive models for Big Data streams.

Current solutions improved classification accuracy by working on the fewest parameters. For example. The best ML model ACNNELM for Big Data stream classification worked on six parameters which are; Training data composition.

- 1) Number of kernels
- 2) Number of layers
- 3) Type of activation function
- 4) Number of iteration
- 5) Variable learning rate

However, to identify the latest critical parameters (for example, hyperspectral features) of advance ML models (Deep Learning) and model the matrix to measure the adaptability factor are potential research directions.

ACKNOWLEDGMENT

This research study is conducted in Universiti Teknologi PETRONAS (UTP), Malaysia as a part of research project “Correlation between Concept Drift Parameters and Performance of Deep Learning Models: Towards Fully Adaptive Deep Learning Models” under Fundamental Research Grant Scheme (FRGS) Ministry of Education (MoE) Malaysia (Grant Reference: FRGS/1/2018/ICT02/UTP/02/2).

REFERENCES

- [1] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, “Deep learning applications and challenges in big data analytics,” *J. Big Data*, 2015.
- [2] A. Budiman, M. I. Fanany, and C. Basaruddin, “Adaptive Online Sequential ELM for Concept Drift Tackling,” *Comput. Intell. Neurosci.*, vol. 2016, 2016.
- [3] A. Budiman, M. I. Fanany, and C. Basaruddin, “Adaptive Convolutional ELM For Concept Drift Handling in Online Stream Data,” Oct. 2016.
- [4] W. Zang, P. Zhang, C. Zhou, and L. Guo, “Comparative study between incremental and ensemble learning on data streams: Case study,” *J. Big Data*, vol. 1, no. 1, pp. 1–16, 2014.
- [5] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, “A survey on Concept Drift adaptation,” *ACM Comput. Surv.*, vol. 46, no. 4, pp. 1–37, 2014.
- [6] L. Kuncheva, Classifier ensembles for changing environments, in: *Multiple Classifier Systems*, Vol.3077 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2004, pp. 1–15.
- [7] L.I.Kuncheva,Classifierensemblesfordetectingconceptchangeinstreaming data: Overview and perspectives, in: *2nd Workshop SUEMA 2008 (ECAI 2008)*, 2008, pp. 5–10.
- [8] Žliobaitė, Indrė. "Learning under concept drift: an overview." arXiv preprint arXiv:1010.4784 (2010).
- [9] Kelly, M. G., Hand, D. J., & Adams, N. M. (1999, August). The impact of changing populations on classifier performance. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 367-371). ACM.
- [10] M. Rouse., “Predictive Analytics Definition,” 2009.
- [11] R. Elwell and R. Polikar, “Incremental learning of Concept Drift in nonstationary environments,” *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1517–1531, 2011.
- [12] G. Ditzler and R. Polikar, “Incremental learning of Concept Drift from streaming imbalanced data,” *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 10, pp. 2283–2301, 2013.
- [13] I. Zliobaite et al., “Next challenges for adaptive learning systems,” *ACM SIGKDD Explor. Newsl.*, vol. 14, no. 1, p. 48, 2012.
- [14] G.-B. Huang et al., “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.

- [15] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst. Man, Cybern. Part B, Cybern.*, vol. 42, pp. 513–29.
- [16] A. Tsymbal, M. Pechenizkiy, P. Cunningham, and S. Puuronen, "Dynamic integration of classifiers for handling Concept Drift," *Inf. Fusion*, vol. 9, no. 1, pp. 56–68, 2008.
- [17] B. Mirza and Z. Lin, "Meta-cognitive online sequential extreme learning machine for imbalanced and concept-drifting data classification," *Neural Networks*, vol. 80, pp. 79–94, 2016.
- [18] N. Liu and H. Wang, "Ensemble based extreme learning machine," *IEEE Signal Process. Lett.*, vol. 17, no. 8, pp. 754–757, 2010.
- [19] A. van Schaik and J. Tapson, "Online and adaptive pseudoinverse solutions for ELM weights," *Neurocomputing*, no. Part A, pp. 233–238, 2015.
- [20] J. Zhai, J. Wang, and X. Wang, "Ensemble online sequential extreme learning machine for large dataset classification," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2014, pp. 2250–2255.
- [21] S. Xu and J. Wang, "Dynamic extreme learning machine for data stream classification," *Neurocomputing*, vol. 238, pp. 433–449, 2017.
- [22] S. Xu and J. Wang, "A fast incremental extreme learning machine algorithm for data streams classification," *Expert Systems with Applications*, vol. 65, pp. 332–344, 2016.
- [23] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A Fast and Accurate Online Sequential Learning Algorithm for Feedforward Networks," *IEEE Trans. Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [24] Y. Lan, Y. C. Soh, and G. Bin Huang, "A Constructive Enhancement for Online Sequential Extreme Learning Machine," in *Proceedings of the International Joint Conference on Neural Networks*, 2009, pp. 1708–1713.
- [25] D. Brzezinski and J. Stefanowski, "Reacting to different types of Concept Drift: The accuracy updated ensemble algorithm," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 25, no. 1, pp. 81–94, 2014. [38].
- [26] B. Krawczyk and M. Wozniak, "Reacting to different types of Concept Drift with adaptive and incremental one-class classifiers," in *Proceedings - 2015 IEEE 2nd International Conference on Cybernetics, CYBCONF 2015*, 2015, pp. 30–35.
- [27] K. Cao, G. Wang, D. Han, J. Ning, and X. Zhang, "Classification of Uncertain Data Streams Based on Extreme Learning Machine," *Cognit. Comput.*, vol. 7, no. 1, pp. 150–160, Feb. 2015.
- [28] I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghādira, "Self-Adaptive Windowing Approach for Handling Complex Concept Drift," *Cognit. Comput.*, vol. 7, no. 6, pp. 772–790, 2015.
- [29] <http://yann.lecun.com/exdb/mnist/>.
- [30] Krizhevsky, Alex, Vinod Nair, and Geoffrey Hinton. "The CIFAR-10 dataset." online: <http://www.cs.toronto.edu/kriz/cifar.html> (2014).
- [31] <https://www.mathworks.com/help/nnet/ug/deep-learning-in-matlab.html>.
- [32] Manzoor Ahmed Hashmani, Syed Muslim Jameel, Hitham Al-Hussain, Mobashar Rehman and Arif Budiman, "Accuracy Performance Degradation in Image Classification Models due to Concept Drift" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 10(5), 2019. <http://dx.doi.org/10.14569/IJACSA.2019.0100552>.