DRIFT-DRIVEN REGRESSION FOR PREDICTING THE EVOLUTION OF PANDEMICS

Khaled Jouini and Ouajdi Korbaa
University of Sousse, MARS Research Lab, LR17ES05, ISITCom
4011 H. Sousse, Tunisia

ABSTRACT

Pandemics will never cease to emerge and threaten public health and the global economy. Predicting the evolution of a pandemic is of paramount interest as it enables policymakers understand the potential spread of a virus and make informed decisions to mitigate its impact. *Concept drifts* refer to situations where the relationship between the input features (*model input*) and the learning targets (*model output*) changes or evolves over time. Concept drifts are common in pandemic curves, not only because new variants appear over time, but also due to factors such as seasonality, policy responses to the pandemic, and changes in the way the disease is treated. Without proper intervention, the accuracy of conventional (*batch*) machine learning models will deteriorate after drift occurs, since they were trained on outdated data. *Incremental learning* is an alternative to batch learning, where the training phase never ends, and the model is incrementally updated as new data becomes available. While incremental models are in general less accurate than batch models, they adapt better to concept drifts as they are continuously refined using the most recent data. Batch and incremental learning are often considered as two distinct and mutually exclusive approaches (Montiel et al., 2018a). In this work we propose CDR (*Collaborative Drift- Driven Regression*), a novel collaborative regression strategy where incremental and batch regressors work together to complement each other's strengths, ergo the overall predictive performance. Experiments conducted on COVID-19 pandemic data, show that CDR is an efficient collaborative learning strategy that yields better results than the underlying batch and incremental models used separately.

KEYWORDS

Concept Drift, Incremental Learning, Collaborative Learning, Pandemic Forecasting

1. INTRODUCTION

The risk of extreme pandemics like COVID-19 is increasing due to climate change, ease of global travel, and increasing rates of disease emergence from animal reservoirs (Marani et al., 2021). The accurate prediction of a pandemic's evolution is crucial for improved preparedness and prompt responses. It is however a particularly challenging task due to factors such as, changes in the way the virus spreads or mutates, changes in testing or reporting practices, and changes in the way the disease is treated. In machine learning and related fields, changes over time in the relationship between input data and the learning target are known as *Concept Drifts* (*a.k.a. data non-stationarity*) (Bifet & Gavaldà, 2007).

Most state-of-the-art machine learning algorithms, referred to as *batch learners* in the sequel, operate under the premise of data stationary, and assume that all training data is available prior to the learning process. Without proper intervention, the predictive performance of batch learners inevitably deteriorates as drifts occur, since they were trained on historical data and no longer accurately reflect the current relationship between input and target variables. A common intervention to handle drifts is to periodically *retrain* the batch model to take into account recent observations. Besides the computational burden, model retraining raises two significant challenges: (*i*) determining when a model is no longer valid and requires retraining (*i.e.* stability-plasticity dilemma); and (*ii*) deciding how much new data to collect, given that collecting more data enhances the chances of generating an accurate model, but also delays the replacement of the old poorly performing model.

Incremental learning (*a.k.a. online learning* or *lifelong learning*) is an alternative to batch learning, wherein the model is trained on small amounts of data at a time, rather than in a single batch. In incremental learning, the training phase is perpetual, and the model undergoes continuous and incremental refinement as new data becomes available. This continual updating enables incremental models to exhibit superior adaptability to drifts when compared to batch models. Another desirable feature of incremental learning is its *"anytime property"* which refers to the ability of a model to make predictions at any point in time during the learning process. By being responsive to changes, easy to maintain (models do not need to be retrained), and able to start making predictions after the first few training instances, incremental learning seems to be well-suited to the requirements of modeling the evolution of a pandemic. The downside of incremental learning, however, is that it builds models by making assumptions on upcoming data (an incremental model is in fact an approximation of the corresponding batch model) (Bifet & Gavaldà, 2007). In contrast to incremental algorithms, batch algorithms need time to collect enough data before building a model, but once the model is built, it is often more accurate than the corresponding incremental model.

Batch and incremental learning are commonly considered as distinct and mutually exclusive approaches (Montiel et al., 2018a). In this paper, we propose CDR (*Collaborative Drift-Driven Regression*), a novel collaborative regression strategy in which incremental and batch regressors work together to complement their strengths, with the ultimate goal of accurately predicting the evolution of pandemics. CDR continuously refines the underlying incremental model as data arrives, retrains a new batch model on recent observations whenever it detects a drift and dynamically selects the best performing model to produce predictions. Experiments conducted on COVID-19 data show that CDR yields better results compared to using the underlying incremental and batch regressors separately.

The remainder of this paper is organized as follows. Section 2 briefly reviews the main concepts related to incremental learning and provides an overview of related work. Section 3 details our CDR approach. Section 4 outlines the experimental evaluation and examines the main findings. Conclusions and future work are discussed in Section 5.

2. PRELIMINARIES AND RELATED WORK

In sensitive areas like pandemics forecasting, interpretability matters as much as accurate predictions, and white-box models are often preferred to black-box models (Salah, I. et al., 2023). Decision Trees are among the most popular white-box models. This section first introduces the key concepts related to incremental trees and then briefly reviews the main approaches utilizing them for pandemics forecasting.

2.1 Incremental Trees

A decision tree is learned top-down by recursively replacing leaves by test nodes. The recursion is completed when a node is deemed homogeneous enough, or when splitting no longer improves predictions. Batch trees scan the entire dataset to discover the attribute leading to the highest homogeneity. The aforementioned process cannot be adopted directly in contexts like pandemics forecasting where only a small fraction of data is accessible during learning. The *Hoeffding Tree* (HT) (Domingos & Hulten, 2000) is the de-facto standard in stream mining and has inspired many state-of-the-art incremental algorithms. The main idea behind HT is that a small fraction of data can often be enough to choose the best splitting attribute (*i.e.* the attribute leading to the highest homogeneity). This idea is supported by the *Hoeffding Bound* which states that, with probability ($1 - \delta$), the true mean of a random variable of range R will not differ from the estimated mean after n independent observations by more than (Domingos & Hulten, 2000):

$$\epsilon = \sqrt{\frac{R^2 \ln(1/\delta)}{2n}}$$

For the purpose of deciding which attribute to split on, the random variable being estimated is the difference in information gain between the best and second-best attributes, *resp.* referred to in the sequel as X_a and X_b. If the computed difference of information gains between X_a and X_b is higher than ϵ, the algorithm asserts with confidence $1-\delta$, that X_a will always remain a better split option than X_b. The *Fast Incremental Model Tree with Drift Detection* (FIMT-DD) (Ikonomovska et al., 2011) is among the state-of-the-art incremental regression trees. Like the *Hoeffding Adaptive Tree* (HAT) (Domingos & Hulten, 2000), FIMT-DD is an adaptive incremental tree that encompasses local change detectors. The salient features of FIMT-DD are synthesized in the following points.

Splitting Criterion. FIMT-DD uses the *standard deviation reduction* (SDR) as splitting criterion. Similarly, to HT and HAT, FIMT-DD uses the Hoeffding Bound to control the risk that, as data arrives, the merit of splitting on X_b exceeds the merit of splitting on X_a. In practice, before splitting a node, FIMT-DD waits until the following condition is met.

$$\frac{SDR(X_b)}{SDR(X_a)} < 1 - \epsilon$$

Linear model at the leaves. FIMT-DD trains a perceptron at each leaf of the tree. The weights of the perceptrons are continuously updated as new data arrives using the incremental stochastic gradient descent and with the objective of minimizing the mean squared error. In addition to their proven effectiveness, perceptrons have the crucial advantage of naturally adapting to drifts (Ikonomovska et al., 2011).

Drift handling. FIMT-DD uses the Page-Hinkley (PH) change detection test (Mouss et al., 2004) at inner nodes to detect changes in the error rate. When a change is detected in a node *inner*, an alternate tree rooted at *inner* is grown with new incoming instances: every new instance that reaches *inner* is used for growing both subtrees. The new subtree replaces the original one when (and if) it performs better.

2.2 Related Work

The volume of literature on predicting the evolution of the COVID-19 pandemic is monumental and covers a wide range of techniques (Miralles-Pechuán et al., 2023). Despite the abundance of studies, little work has been dedicated to incremental learning (Miralles-Pechuán et al., 2023). Existing incremental approaches can be broadly classified into two families: *compartmental models* and *machine learning models*. Machine learning models rely on historical data to forecast future outcomes. Compartmental models divide the population into different groups and use differential equations to model the transition of individuals between these groups. The approach of (Camargo et al., 2022) is representative of compartmental methods and involves a two-component architecture, where the first component is a feature engineering process that selects the predictor variables used by the predictive models of the second component. Specifically, the first component analyzes the temporal dependencies between the SEIRD variables (*Susceptible, Exposed, Infected, Recovered, and Dead*) and identify, for each target variable, the best subset of predictors. The second component of the architecture employs an ensemble learning approach, where various models are trained using different batch and incremental algorithms. The final output is chosen based on the predictions of the top-performing model. When the predictions of the top-performing model are not deemed accurate enough, the approach builds new predictive models. The approach of (Camargo et al., 2022) lacks a drift detection mechanism for automatically detecting drops in performance. The authors do not specify how and when the predictions of the top-performing model are no longer considered as good enough and, hence, when the employed models need to be retrained. Furthermore, (Camargo et al., 2022) retrains its models on the entire available dataset, which implies that over time, these models are trained on a diminishing proportion of recent data and are consequently less and less sensitive to changes.

The study of (Miralles-Pechuán et al., 2023) is representative of machine learning approaches and com- pares batch and incremental algorithms using COVID-19 data from 50 countries. Results showed that incremental methods are more effective in adapting to changes and have lower computational cost compared to techniques such as LSTMs. A salient feature of (Miralles-Pechuán et al., 2023) is that it tests three different approaches. The first approach (*resp.* the second approach), involves training each model using data from a single country (*resp.* using data from the 50 countries). In the third approach, clustering is first applied to identify the most similar countries to the one being predicted and then each model is trained using data from those countries. Results showed that the third approach outperforms the two others. While the study of (Miralles-Pechuán et al., 2023) includes HT and HAT, surprisingly, it overlooks FIMT-DD. Similar to (Camargo et al., 2022), the work of (Miralles-Pechuán et al., 2023) does not include a drift detector for an automated management of the life cycle of machine learning models.

3. COLLABORATIVE DRIFT-DRIVEN REGRESSION (CDR)

The goal of a regression task is to learn a model M that predicts a real value, and not one of a discrete set of values as in classification. Formally, let S be a continuous stream of data: $S = \{(\overrightarrow{x^t}, y^t)\}$, where $\overrightarrow{x^t}$ is a feature vector, $y^t \in \mathbb{R}$ is the target variable and t the arrival timestamp. The goal is to incrementally learn $M: \vec{x} \rightarrow y$ as new data becomes available (Gomes et al., 2018). The predicted value of M is denoted as $\hat{y}$. When the actual value y gets revealed, the performance P is measured according to a loss function l: $P(M) = l(y, \hat{y})$. Performance of incremental models is typically measured using *prequential evaluation* (*a.k.a. test-then-train* evaluation), where each instance is used to test the model before it is used for training (Gomes et al., 2018). For the i^{th} instance: $\widehat{y^i} = M^{i-1}(\overrightarrow{x^i})$.

Our drift-driven collaborative regression approach CDR draws inspiration from the work of (Montiel et al., 2018a) on fast and slow classifiers. CDR combines incremental learning and batch learning to harness the strengths of both: (*i*) the accuracy of batch regressors, and (*ii*) the anytime property and adaptability to drifts of incremental regressors. Illustrated in Figure 1, the learning process of CDR involves the utilization of incremental learning to continuously and incrementally train and refine a regressor I, as new samples become available. Concurrently, batch learning is used to train a sequence of (batch) regressors $\{B_1, B_2, \dots, B_n\}$. As illustrated in Figure 1, whenever a drift is detected, the current batch regressor B_i is invalidated and is subsequently replaced by a new regressor B_{i+1}. While I is trained on single samples as they arrive, a batch regressor B_i is trained on a (micro-) batch M_i containing the k most recent samples. This implies that the training process of B_{i+1} is deferred until k samples are gathered.

CDR uses *ADaptive WINdowing* (ADWIN) (Bifet & Gavaldà, 2007) for drift detection. The basic idea behind ADWIN is to maintain a variable-length sliding window W which increases in size as long as no drift is detected. To detect a drift, ADWIN repeatedly partitions W into two adjacent sub-windows W_0 and W_1 and compares their average to decide whether they are likely to originate from the same distribution. If W_0 and W_1 exhibit distinct enough averages and are of sufficient size, then a drift is detected and W is shrunk by dropping W_0 items from the window. In practice, ADWIN tests if the difference between the averages of W_0 and W_1 is larger than a variable value ϵ_{cut} computed as (Bifet & Gavaldà, 2007):

$$\epsilon_{cut} = \sqrt{\frac{1}{2m} ln \frac{4|W|}{\delta}}$$

, where m is the harmonic mean of W_0 and W_1. Unlike other existing drift detectors, ADWIN is assumption-free. Its only parameter is a confidence bound $\delta \in [0, 1]$, which enables adjusting the sensitivity to drifts.

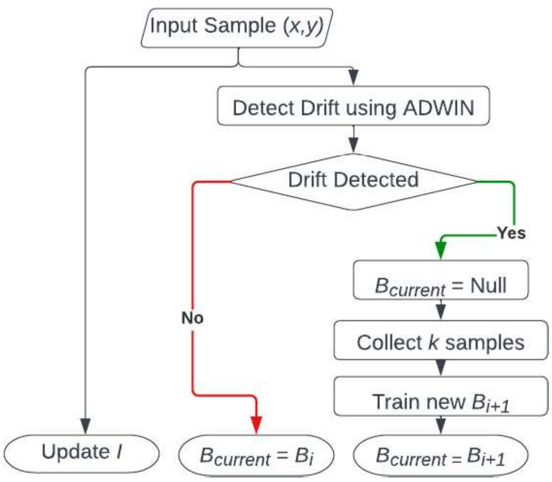


Figure 1. CDR - Learning Process

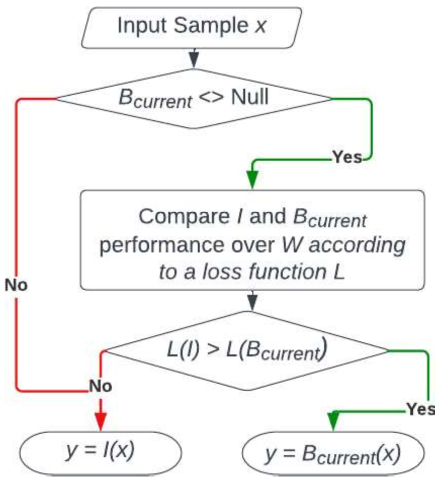


Figure 2. CDR - Inference Process

The inference process of CDR is depicted in the flowchart of Figure 2. As illustrated in Figure 2, when a batch regressor B_i is invalidated and until k samples are collected to train a new batch regressor B_{i+1}, only the incremental regressor I is used for inference. When a new batch regressor $B_{current}$ becomes available, CDR tracks the predictive performance of I and $B_{current}$ over a sliding window W containing the most recent observations. The top-performing model over W is then selected for inference. The aforementioned process is repeated for each new incoming instance. It is important to note that CDR is independent of the underlying learners and is compatible with any combination of incremental and batch methods.

4. EXPERIMENTAL EVALUATION

4.1 Tools and Datasets

We considered the dataset *Coronavirus Pandemic (COVID-19)* (Mathieu, E. et al., 2023), provided by Our World in Data (OWID). The original dataset includes daily information about the pandemic in 219 countries. Our target variables are the daily new confirmed cases and deaths per million people. We model the evolution of the daily new confirmed cases and deaths as function of the previously reported daily new confirmed cases per million people. For each country *C* and each record of *C* with a timestamp *t*, we consider the number of cases per million reported at 8 time points[1]: t minus 1 week, t minus 2 weeks,.., t minus 8 weeks. The obtained dataset contains nine input variables, ≈177k samples and covers the period starting from March 28, 2020 to November 30, 2022. When restricted to Tunisian data, the dataset contains ≈910 samples.

At the current state of our work, we implemented CDR using MOA (*Massive On line Analysis*) (Bifet et al., 2018), *Scikit-Multiflow* (Montiel et al., 2018b) and *Scikit-Learn* (Pedregosa et al., 2011). We configured CDR using FIMT-DD for incremental learning, DT (Pedregosa et al., 2011) for batch learning, and ADWIN for drift detection. All methods were run using their default settings. We compared the performance of batch and incremental models using a sliding window of one week and selected the best performing model to make predictions for the current input. Batch models were trained on windows of three weeks, with two weeks before and one week after a detected drift. The models' performance was evaluated using the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). In general, RMSE is more sensitive to large errors and outliers, while MAE is more interpretable.

4.2 Results and Discussion

A large number of experiments have been performed to demonstrate the effectiveness of CDR. Due to the lack of space, only few results are presented herein. In the first set of experiments, we compare the performance of CDR and established incremental methods using a prequential evaluation scheme. The obtained results are summarized in Table 1, and partially illustrated in Figures 3 and 4. As shown in Table 1, when compared to FIMT-DD, CDR achieves an improvement of 2.99%, 9.78%, 5.72% and 5.51% (resp. 44.59%, 63.97%, 15.75% and 16.41%) with regards to $RMSE_{Deaths}$, MAE_{Deaths}, $RMSE_{Cases}$ and MAE_{Cases} on world data (resp. single-country data). In contrast to HAT and FIMT-DD, HT lacks specialized mechanisms for handling drifts. When compared to HT, CDR achieves an improvement of 20.36%, 31.52%, 33.62% and 29.54% (resp. 57.32%, 79.41%, 26.12% and 30.7%) with regards to $RMSE_{Deaths}$, MAE_{Deaths}, $RMSE_{Cases}$ and MAE_{Cases} on world data (resp. single-country data). As we can observe in figures 3 and 4, CDR benefits from the continuous selection of predictions from both the incremental and the batch regressors. Near drift points, CDR opts in most cases for the batch model, which is only trained on recent data. However, the performance of the batch learner deteriorates quickly as we move away from drift points, primarily due to overfitting. Consequently, CDR transitions to the incremental learner. It's important to emphasize that the model selection near drift points is solely based on the performance (Mean Absolute Error) of both regressors, evaluated within a sliding window of one week.

[1] We do not consider the cases reported less than a week before *t*. Although including such observations could lead to more accurate models, they are not practical for timely policy responses.

ISBN: 978-989-8704-53-5 © 2023

Our second set of experiments compares the performance of a conventional batch regression tree against CDR. The prequential evaluation approach, commonly used in incremental learning, can be applied to batch learning by repeatedly retraining and reevaluating the model. For each training/testing round, the dataset is split into a training set and a test set in an order-preserving fashion. Instances used for testing the i^{th} batch model are appended to the training set of the $(i+1)^{th}$ model. The evaluation of the $(i+1)^{th}$ batch model is then performed on instances that arrived after its training and before the training of a new model. The process is repeated for multiple rounds until all the data has been used for both training and testing (except the first batch of data, which is only used for training, and the last batch, which is only used for testing). In (Miralles-Pechuán et al., 2023), training/evaluation rounds are referred to as *milestones*. In this study, we followed the aforementioned process (used also in (Miralles-Pechuán et al., 2023)) and adopted a realistic scenario where a new batch model is trained from scratch every ≈3 months, resulting in a set of 9 milestones (and, hence, 9 batch models). Tables 3 and 4 report the performance of the considered batch models and of CDR over the 9 milestones. As illustrated in Tables 3 and 4 CDR by far outperforms the corresponding batch model and respectively achieves an average improvement of 111.87%, 78.14%, 59.02% and 20.64% (resp. 124.83%,

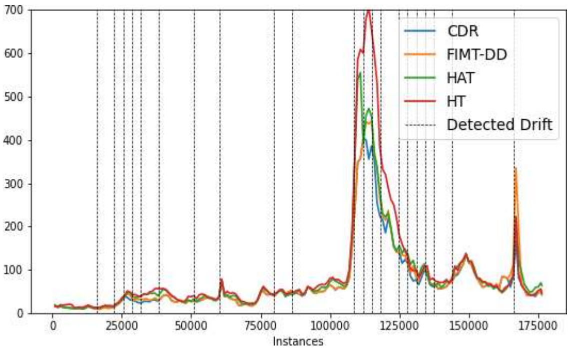


Figure 3. World - Daily New Confirmed Cases MAE achieved by CDR and incremental learners

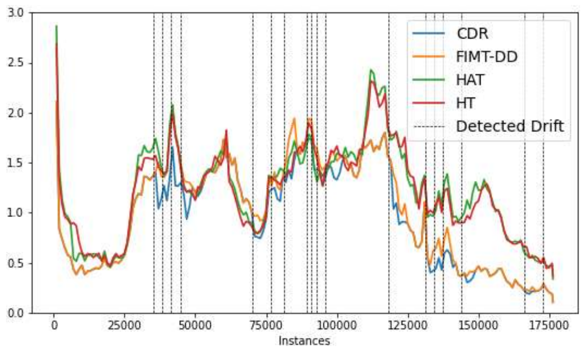


Figure 4. World - Daily New Confirmed Deaths MAE achieved by CDR and incremental learners

90.29%, 109.17% and 65.13%) with regards to resp. $RMSE_{Deaths}$, MAE_{Deaths}, $RMSE_{Cases}$ and MAE_{Cases} on world data (resp. single-country data).

Table 1. MAE and RMSE achieved by CDR and Incremental Learners

		Daily New Confirmed Deaths		Daily New Confirmed Cases	
		$RMSE_{Deaths}$	MAE_{Deaths}	$RMSE_{Cases}$	MAE_{Cases}
World	HT	2.01	1.21	174.1	91.51
	HAT	2.03	1.23	150.12	80.67
	FIMT-DD	1.72	1.01	137.75	74.53
	CDR	**1.67**	**0.92**	**130.29**	**70.64**
Tunisia	HT	2.47	2.44	61.56	58.71
	HAT	2.46	2.43	59.35	56.68
	FIMT-DD	2.27	2.23	56.50	52.29
	CDR	**1.57**	**1.36**	**48.81**	**44.92**

Besides confirming that CDR yields better results than those attained by each of the contributing models separately, the aforementioned experiments allow to draw the following important conclusions.

Retraining on recent data vs. retraining on the entire dataset. While (re-)training a batch model on the whole available dataset is beneficial in some cases to learn stable concepts or recurrent drifts, it is not the most suitable approach when it comes to predicting a pandemic's evolution. This can be observed in Tables 2 and 3, where in most cases, CDR becomes increasingly more efficient than the batch model over time. The reason behind this is that the batch model is (re-)trained on datasets with a diminishing proportion of recent data, causing it to become less and less sensitive to changes.

Learning using data from a single country vs. Learning using data from multiple countries. As it can observed in Tables 1 and 2 and 3, CDR exhibits higher improvements over incremental and batch models when training is performed using data from a single country. This is mainly due to the occurrence of drifts at different time points across countries, and to the fact that using data from multiple countries results in models that do not accurately represent any specific country. Overall, even though using data from multiple countries provides more training data, using data from a single country yields better results, as detecting (and, hence, adapting to) drifts is easier. It should be noted that this confirms the results found in (Miralles-Pechuán et al., 2023).

Table 2. Daily New Confirmed Deaths - MAE and RMSE achieved by CDR and the batch Decision Tree

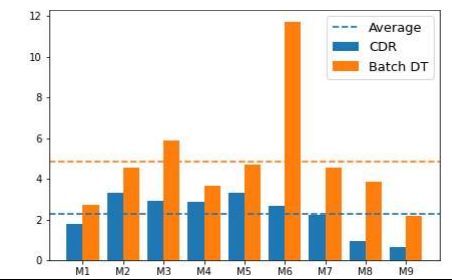
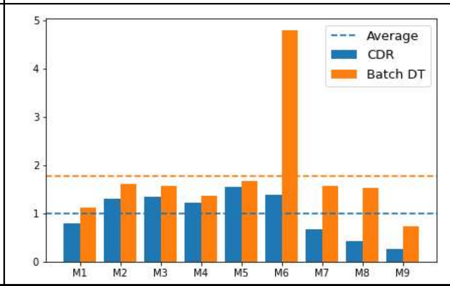
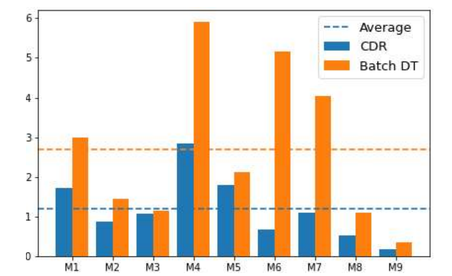
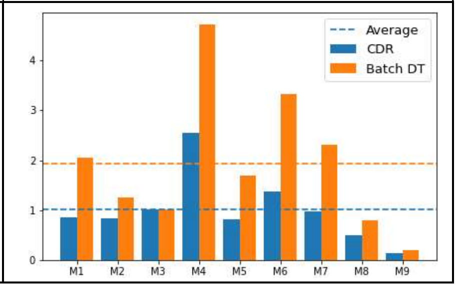
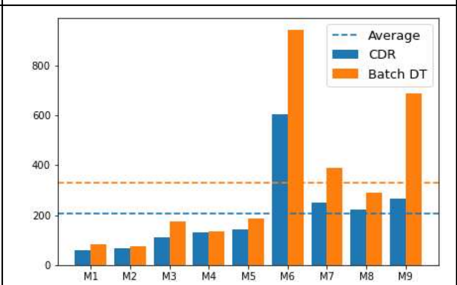
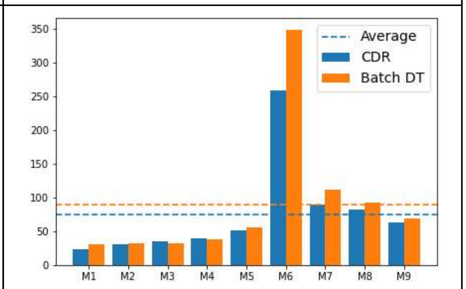
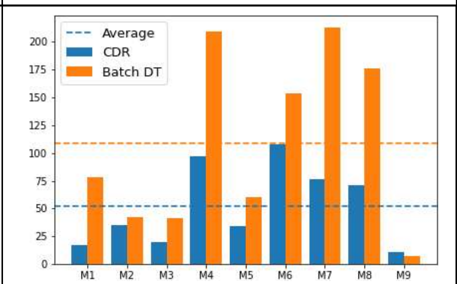
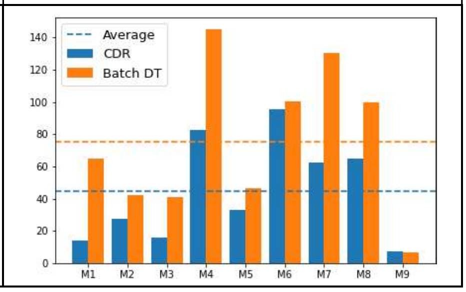
<table>
<tr><th></th><th colspan="2">Daily New Confirmed Deaths (per 1M)</th></tr>
<tr><th></th><th>$RMSE_{Deaths}$</th><th>MAE_{Deaths}</th></tr>
<tr><td>World</td><td></td><td></td></tr>
<tr><td>Tunisia</td><td></td><td></td></tr>
</table>

Table 3. Daily New Confirmed Cases - MAE and RMSE achieved by CDR and the batch Decision Tree

<table>
<tr><th></th><th colspan="2">Daily New Confirmed Cases (per 1M)</th></tr>
<tr><th></th><th>$RMSE_{Cases}$</th><th>MAE_{Cases}</th></tr>
<tr><td>World</td><td></td><td></td></tr>
<tr><td>Tunisia</td><td></td><td></td></tr>
</table>

5. CONCLUSION

Pandemics will never cease to emerge and threaten both public health and the global economy. Therefore, it is crucial to learn from past pandemics to develop effective tools for preventing and controlling future outbreaks. This involves not only researching treatments and therapies, but also implementing efficient epidemiological surveillance systems. In this paper we proposed CDR, an innovative collaborative regression strategy for predicting the evolution of pandemics. CDR continuously refines the underlying incremental model as data arrives, uses ADWIN as drift detector and retrains a new batch model on recent observations whenever it detects a drift. At the current state of our work, we implemented CDR using FIMT-DD as incremental regressor and the Decision Tree as batch regressor. Experiments on COVID-19 data, showed that CDR is an effective collaboration strategy that yields better results than those attained by the incremental and the batch model separately.

In this paper, we mainly focused on connecting the dots between batch regressors and incremental regressors and on the relevance of using them in conjunction to predict the evolution of a pandemic. The results obtained are highly encouraging for exploring other forms of collaboration between batch learning and incremental learning. a part of our future work, we intend to investigate the use of incremental and batch ensemble models instead of single models to further alleviate the effects of concept drifts. On the other hand, our experimental study revealed that even though using data from multiple countries provides more training data, using data from a single country yields better results. This is mainly due to the occurrence of drifts at different time points across countries. In our future work, we intend to identify subgroups of countries exhibiting similar patterns and apply CDR to each subgroup independently. We believe that doing so will allow to leverage more training data while maintaining the ability to efficiently detect and adapt to drifts.

REFERENCES

Bifet, A. et al., 2018. *Machine learning for data streams: with practical examples in MOA*. MIT Press, Massachusetts, USA.

Bifet, A. and Gavaldà, R., 2007. Learning from time-changing data with adaptive windowing. *Proceedings of the 2007 SIAM International Conference on Data Mining (SDM)*. pp. 443–448.

Camargo, E. et al., 2022. An incremental learning approach to prediction models of seird variables in the context of the covid-19 pandemic. *In Health and Technology*, Vol. 12, No. 4, pp. 2190–7196.

Domingos, P. and Hulten, G., 2000. Mining high-speed data streams. *Proceedings of the sixth ACM SIGKDD international conference on Knowledge Discovery and Data mining*. Boston, USA, pp. 71–80.

Gomes, H. M. et al., 2018. Adaptive random forests for data stream regression. *26th European Symposium on Artificial Neural Networks (ESANN 2018)*, Bruges, Belgium, pp. 267–272

Ikonomovska, E. et al., 2011. Learning model trees from evolving data streams. *Data mining and knowledge discovery*, Vol. 23, No. 1, pp. 128–168.

Marani, M. et al., 2021. Intensity and frequency of extreme novel epidemics. *Proceedings of the National Academy of Sciences*, Vol. 118, No. 35, e2105482118.

Mathieu, E. et al., 2023. OurWorldInData.org. *Coronavirus Pandemic (COVID-19)*. Retrieved from: “https://ourworldindata.org/coronavirus” [Online Resource].

Miralles-Pechuán, L. et al., 2023. Forecasting covid-19 cases using dynamic time warping and incremental machine learning methods. *Expert Systems*, Vol. 40, No. 6, e13237.

Montiel, J. et al., 2018a. Learning fast and slow: A unified batch/stream framework. *2018 IEEE International Conference on Big Data*, Seattle, WA, USA, pp. 1065–1072.

Montiel, J. et al., 2018b. Scikit-multiflow: A multi-output streaming framework. *Journal of Machine Learning Research*, Vol. 19, No. 72, pp. 1–5.

Mouss, H. et al., 2004. Test of page-hinckley, an approach for fault detection in an agro-alimentary production system. *2004 5th Asian Control Conference (IEEE Cat. No. 04EX904)*, Melbourne, VIC, Australia, pp. 815–818.

Pedregosa, F. et al., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830.

Salah, I. et al., 2023. On the use of data augmentation for stance and fake news detection. *Journal of Information and Telecommunication*, Vol. 7, No. 3, pp. 359–375.