Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

Comparative Analysis of Classification based Techniques using Change Detection and Ensemble models

Leena A Deshpande[1*], M.R. Narasingarao[2]
[1] Research scholar, Department of Computer Science and Engineering,
KLEF Deemed to be University,
Vijayawada ,522502, India
[2]Department of Computer Science and Engineering,
KLEF Deemed to be University,
Vijayawada ,522502, India
*leena.deshpande@viit.ac.in

ABSTRACT

In the data streaming environment, more and more data gets generated due to various applications like network intrusion detection, weather forecast, finance, sensor based data acquisition systems and so on. Due to rapid advancement in technology, the arriving pattern is expanding and hence data arrives continuously which need to be monitored rather than storage. Some of the application requires a quick analysis of arriving data. In such situations, storing such continuous and huge data and then analyzing the performance will not be useful for better decision. Since the probability distribution of such data changes over the period of time it is necessary to timely analyze such patterns in the data such change is occurred due to unstable data distribution. This raises a need of continuous monitoring of such changes in in streaming environment the researchers have given major attention to concept change which is referred as concept drift. In this paper we have surveyed the various classification-based methods for handling change in the distribution, in introduction section we have discussed various types of concept drift and their application, followed by the methodologies used to handle the diversity and further, discussed the experimental work performed on Covid-19 dataset. The last section concludes how the concept drift learners are extensively used in covid-19 data to detect the change in the distribution and how an adaptive learners and evaluators are used for solving real life applications.

Keywords
Classification, ;Concept drift; Ensemble model; Accuracy

Introduction

A *data stream* is an ordered sequence of instances that arrive continuously with time-varying intensity[ ] , which is often impossible to store instances permanently or process them more than once. Because of its rapid pattern of data arriving, the data distribution does not remain stationary. Traditional knowledge of data mining algorithms lower down the accuracy due to such change in data distribution. It may discover some patterns as frequent while other patterns tend to disappear and gets wrongly predicted. Thus a traditional data mining process may not be effective to analyses hidden pattern in such data environment and there is a need of online stream mining**.**Traditional classification algorithm works on the data which does not changes its distribution and hence the training data and testing data is assumed to be of same distribution. In real life situation like twitter data analysis [2], Spam filtering [3] , sensor based acquisition system, Chemical reactor plants [4] ,the distribution of data changes as it is arrived . In such scenarios, the change in the distribution may degrade the performance. Such degradation of classifier performance needs to be monitored carefully by identifying the class labels of the model. The change in the class label leads to concept drift where target class in introduced newly after certain period or it may remain for a particular period of time. Thus, the concept drift in data may occur due to 3 different conditions i.e prior probability, posterior probability and class conditional probability distribution. [5]

Types of Concept Drift

Concept drift is divided into 5 basic types [6]

a) **Sudden drift:** Sudden drift occurs when there is sudden peak or fall in the data which causes a change in the target concept. E.g sudden change in the share market after declaration of demonetization of Indian currency notes by Indian government. A sale of particular mobile after declaration of sale and discount by online shopping

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

websites. In both the situations the weight assigned to the attribute gets changed and hence prediction accuracy may get affected. A sudden concept drift occurs when sudden change in the distribution of data occurred in the arriving pattern. E.g in a chemical reactor plant a sudden change in the sensor reading may cause change in temperature or pressure level . Such changes need to be monitored and handled at run time only as a machine learning model, such activity or process needs to be monitored. A sudden drift is also observed in finance application share market or stock exchange program. A quick change in regulation or policies may impact the stock market . Such change in the stock market has a huge influence on individual as well as companies. For handling such change, an efficient predictive model needs to be constructed where a sudden change in the concept may be identified using adaptive windowing techniques [7]

b)Gradual drift: gradual drift is a slow and gradual change in the data distribution which needs to be handled by the system

E.g. people's opinion on social media about something or someone after certain period of time may get changed gradually. Such gradual change must be monitored by the model for better labelled sentiment data. Application like traffic monitoring, customer purchase patterns are likely to adapt to change after certain interval of time Many times this change is referred as a seasonal change. A concept drift monitoring system should be able to adapt to this smooth change. If we observe current Covid-19 situation, due to covid effect ,a customer buying pattern , web usage patterns are changing . Such change is a gradual change which needs to be identified as a concept drift and learning algorithm must be deployed to tune with the new pattern and forget the old one.

c) **Incremental**: Incremental drift occurs when the change in the variable assignment is incremental over the period of time. , system must be retrained with the incremental pattern.

d) Recurring drift: Recurring drift are also called as local drift [8] . Certain changes in the distribution occurs at certain period of time and again it appears with the original distribution. E.g Due to Covid -19 pandemic situation, the usage of internet broadband service is changed as more and more online teaching platforms are used . In such situations some features are getting deviated and old features are replaced with new features. Such drift is not periodical as you do not predict as when the data gets its original distribution. In traffic monitoring system , drift detection must react to change in the traffic patterns which may arise due to change in rules or routes . This may be a seasonal change which adapts to a transition and predicts the new routes. But such change may occur for some period of time. It may occur due to seasons, festivals or some common regular social gathering of people, due to which the traffic routes may require changes . Such change must be predicted by learning the seasonal impact on traffic monitoring system.

e)Blip: Blip is also referred as outliers. Such outliers are detected as it does not meet the properties of the class label assigned and hence detected as outlier by the model. E.g in an intrusion detection system, when model is unable to detect the presence of newly introduced intruder, it labels it as an outlier. Such outliers need to be monitored separately to identify new patterns in the system. In text mining, the change in people's opinion may shift due to various reasons and such change should be identified with correct feature selection techniques. Many political applications like exit poll, popularity index of celebrity or the opinion by common people on various government policies Such information is captured using various social media platforms. Such data comes with mixed opinion along with a noice as some opinions are not about the actual topic. It is deviated from the current topic and hence need to be treated as an outlier. So identifying a new context or meaning in a text is a Blip detection monitoring where a Blip distribution will be eliminated and a new feature selection needs to be revised using essential learning algorithms [9]

Selection strategies for sampling:

To mine the streaming data with concept drift identification, the essential function block is the selection strategy. since a stream of data is generated continuously data may not be required to store as a historical data and hence to build the model summary of data can be computed and model accordingly. Also, the historical data may no longer in existence or important. If we consider such data to the concept drift, it may give the false classification outcome. Thus, a good sampling strategy is required for the designing a concept drift model. Researchers have extensively discovered various sampling algorithms; the random selection is a commonly used sampling techniques where other sample selection techniques are decision tree.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

Challenges of streaming data

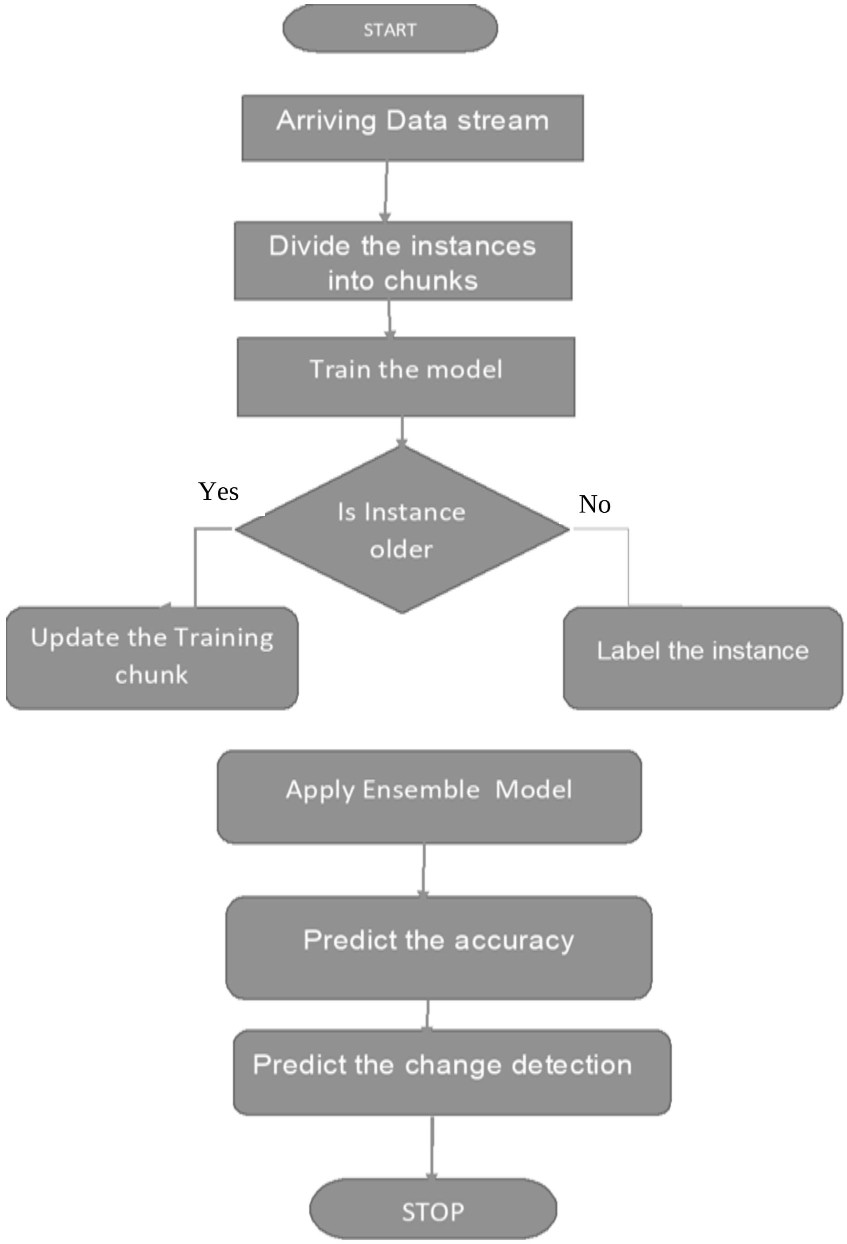
When large amount of data is generated, it becomes practically impossible to store such of data offline. Sometimes in the continuous arriving data pattern, all data need not to be processed, due to which current or active data only needs to be stored and processed. In such situations, online learners also called as incremental approach [10] is used for designing a model which evaluate the arriving pattern as and when it arrives. The advantage of incremental learning is that the instances or data streams are labelled instantly and thus if any further change is detected in the arriving pattern, the classifier updates the label and thus able to tune the change in the training data. Another advantage of incremental learning is you train the model in one pass thus reads the block a at a time and hence require less time and memory. Another way of detecting the change in the concept is by detecting the impact on classifier accuracy [11] This impact may degrades the performance of classifier which is usually occurred due to flexible training sample after certain period of stability. Thus change in the distribution causes a classifier to rebuild the model by forgetting the old parameters and values associated with it by adopting various windowing techniques [7][12] . An adaptive windowing techniques help to adapt a new class by forgetting the old instances and train the model with changing window size . Researchers have also extensively studied on parameters affecting the drift and their trade off with accuracy. As mentioned in the types of drift, each kind of drift is a kind of drift detector which detect the sudden, gradual drift and thus reacts to the change. Another challenge in handling concept drift is class imbalance problem. A class imbalance problem occurs due to skewed distribution of positive and negative classes.[13][14] Such imbalance causes a machine learning model to degrade its performance .However a single classification model may not perform well for accurate change detection and thus requires ensemble models which are widely used to boost the accuracy with change in distribution and flexibility. In some applications a class imbalance problem is found in e.g. fraud detection, network intrusion detection, disease detection etc. IN fraud detection system, in a transactional history a regular customer may found with exceptional transaction and transactions may be lower in numbers. Although it is minor in number it is a major class to predict such exception to detect whether it is a fraudulent or non-fraudulent transaction. A concept drift detector must be able to detect this imbalance [15] proposed a decision trees for handling streaming data, named as Very Fast Decision Tree (VFDT) . Thus, one of the main research contribution is to develop an efficient change detector system for different sets of problem.

Change in data distribution:

Variety of literature is been reviewed for monitoring the performance of change concept and regular stream Street and Kim [16] proposed an ensemble method called Streaming Ensemble Algorithm (SEA). A similar way of restructuring an ensemble was proposed by Wang et al. [7][17]. In their algorithm, called Accuracy Weighted Ensemble (AWE), they train a new classifier on each incoming data chunk and use that chunk to evaluate all the existing ensemble members to select the best component classifiers. Wang et al. stated and proved that an ensemble (Ek) is built from the k most recent data chunks and a single classifier which uses an internal change detector parameter which optimizes the accuracy and resources. Warse [18] has given in depth literature review on various statistical approaches for handling concept drift and the metric used for evaluating drift detectors.. In his review, he has also stated the limitation of benchmark datasets available for experimentation. A more acceptable data set with a real change in concept needs to be available for experimentation.

Following flow chart represents the overall change detection model of our proposed system . The model is trained in incremental way where older instances are ignored and newer instances gains higher weight to retrain the model.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.



Ensemble methodologies

Ensemble algorithms uses multiple classifiers for prediction. A final prediction is based on combined outcome of different classifiers. These algorithms are effectively used on streaming problem .[19] A combined majority voting algorithm is used to vote the classifier possible outcome. However, if a classifier is failed to predict the correct outcome, its weight is reduced and a new classifier thus replaced a weak classifier. An individual classifiers performance decided the weight of the classifier. Thus, the complexity of machine learning algorithms is extensively handled by ensemble models. Researchers have extensively used ensemble for various reasons. We have listed popular features of ensemble in the following section.

1. **Generalized performance**: Ensemble learning methods are extensively used for various reasons. The Ensemble classifier gives you the generalized performance over the single classifier of the use of multiple

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

classifiers. The use of multiple classifiers replaces the old or weak classifier and hence allows us to make more generalized selection of training samples. Another reason for ensemble is its efficiency of handling huge data.[20][21][22]

2. **Subset of data**: Todays machine learning model needs a special learning technique to select the subsample out of the vast amount of training samples that are created. Ensemble model allows you to use different samples for different models and they combine the different predictions to make a final decision. Thus, a huge stream of data can be efficiently trained by ensemble model. From such large stream of data, new features along with the old one may generate. e.g. social media data ,video surveillance, biomedical data [23][24] etc. , appropriate selection of sub features are needed to make the model more efficient by deploying the optimum feature selection strategy for such data .

3. **Bagging and Boosting:** Due to diversity in the data, the decision boundaries become complex and reduces the possibility if isolated boundary class. This diversity can be improved by resampling method. A bagging strategy is applied on all training sample [25]. In bagging, bootstrap samples are created , [20][26]and then given to different classifiers for training. Different classifiers vote the outcome and based on majority voting a final decision is taken. Because of its diversity in the samples, different decision boundaries must be obtained by reducing the repetition of data in various subsets. This can be achieved with decision trees where training parameters changes randomly. This avoids the repetition of similar samples in different subsets. However, in boosting, the weal learners are retrained and boosted with assigned weights so that most informative data would be selected for training the next classifier This is achieved by assigning weights to the distribution of training samples.

4. **Heterogeneous feature Sets:** another powerful feature of ensemble is its ability to combine heterogeneous feature sets under one family distribution. The typical example of such feature is medical diagnostic data, where different features are gathered for predicting the disease like MRI scan, The blood reports, Patients history, the X- ray report etc.[27]] Such heterogeneous source of data generates various features , such data can be combined to form a feature set , here even if some of the features are missing , ensemble would handle it to generate a new feature set for heterogeneous data.

Literature Review

Ensemble model solution

This section explains the multiple categories of ensemble model, approaches used and, limitation of each category. Researchers have proposed a possible solution for each category.

Table 1: Ensemble models

S No	Ensemble Models	Proposed by	Strategy
1	Weighted Majority Voting [1]	Kolter and Maloof	Incremental learning approach with weights assigned to each classifier
2	Accuracy weighted Ensemble	Wang et al.	New classifiers are built based on chunk size and MSE (mean square error) is used as a measure to replace the old chunk
3	Streaming ensemble algorithm	Street and Kim	Block based algorithms. Sequential blocks are built for classification
4	Bagging and Boosting (Ada Boost)	Breiman, Schapire	Generate subset of data by random selection and boost the performance of weak learner to strong level

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

5	Learn ++ algorithm	Polikar et al	Series of classifiers are trained for single batch of data and prediction error is used for weighted distribution of all instances

. Several Algorithms and technique shave been proposed for streaming data , in view of its performance capability and compatibility . WE tries to analyses these algorithms to highlight the merits of each one and limitations if any with respect new challenges in the arriving pattern that may arise in streaming data. Table 1 illustrates the ensemble models and strategies used by each of the model. Table 2 summarizes the concept change detection techniques based on single classifier model, ensemble model and drift detection model. Year wise invention is described in the table with the methodologies adopted and limitations of each of the methodology Based on these comparative analysis we may say that the attempts made by earlier researchers for solving the drift is limited to fixed size window or adaptive window size, but an adaptive windowing technique is computationally high in case of large data size . So most of the literature is found on use of various efficient models to improve further and very less work haven been given on sampling strategies and its impact on building the model. Following table summarizes the concept change detection techniques based on ensemble model and its limitation. Year wise study is considered from the date of invention

Table 2. Comparative analysis of classification-based technique

S No	Category	Approaches	Algorithms and Year of invented	Limitations /Findings	Memory and computational efficiency
1	Single classifier 1997	Statistical, Non parametric approach	Naïve Bayes Neural Network [1997[	Inefficient in case of large data . Only static concepts are detected	Clustering approach to identify summarized samples to train large data sample
			Decision Tree: C 4.5 [2000]	Rebuilding of classifier in case of change is detected, VFDT requires more example to train	
			CVDFT, SEA[2001]	Expensive in case of large rules , not suitable for high speed data	
			Similarity based classification [2004]	Suitable for text data and for capturing quick	

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

				changes	
2	Windowing Techniques 2000	Non parametric approach	Sliding window [2002]	Accuracy is affected by the window size	Efficiency is increased by multimode parallelization
			Weighted window [2003]	Overhead of different decaying function for different window	
			ADWIN [2004]	Computationally expensive as more memory consumption in case of large sub windows.	
			Parallel ADWIN [2008-2009]	Single node and multinode parallelization	
3	Drift Detection 2001	Incremental approach	DDM [2008]	Works better for sudden drift but poor in case of no drift or gradual drift	Ensemble models are integrated with drift detection techniques
			EDDM [2008][2009]	Works better for gradual drift but sensitive to noise	
			Hoeffding Tree [2011]	Less memory and time required However inefficient when large no of rules are formed	
			Modified weighted one class SVM [2014]	One-class classifier with incremental learning and forgetting	

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

			Online Stream classifier with incremental semi-supervised learning [2015]	Can work with unlabelled data , uses semi supervised approach	
			Adaptive micro cluster [2017]	Parallel data stream mining, choice of ensemble is more effective than drift detector [36]	Memory and Cpu cost optimization due to parallelization efficient
			Prequential error rate using Hoefdding inequality [2020]	Inconsistency of error rate is measured to detect concept drift . [37][38]	Desired performance is more reliable than efficiency .
			Ensemble algorithms with Nb and HDDM as base learners	Suitable for abrupt drift along with additional overhead of auxiliary drift detection algorithm	Additional Memory requirement
			Several Ensemble algorithms and a regressor is combined to detect drift	Computationally costly	Faster throughput
			Efficient optimization algorithms using	Suitable for Video surveillance data	Efficient in space and time complexity

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

			decision trees as base learner [2020]		
			Prequential error rate using Hoefdding inequality [2020]	Inconsistency of error rate is measured to detect concept drift	

Results

We have used corona virus data set which is available on covid19india.org. Corona virus has infected all over the globe in large extend and thus needs special attention and contribution along with doctors and medical science research group, equal contribution of researches specially in machine learning is needed for predicting the impact of virus on human being. However, we have collected data set of India and its various states. The statistics tells about no of deaths, confirmed cases, no of cured cases in each state, symptoms if any along with date, time, gender. In Indian context we observed that, the corona virus is speeded in diversified way. Some states like Kerala, Maharashtra pickup up the cases after May and suddenly some states since march and it the growth is exponential till date. We have applied traditional machine learning and ensemble learning model for predicting the future no of cases. We have categorized the data state wise and month wise for analysis. We observed that data is coming with varied pattern as in the beginning, the no o cases were very low. but suddenly few of the states has shown the variability in the data wherein some states the no of cases are increased with minimum variance. Such diversified patterns are well captured by ensemble models and traditional regression models. However, the data set could be modelled with more no of features like patient's medical history, lockdown period, type of treatment etc. Since this experimental work only surveys how block-based ensemble model proves better over the other methods, we have experimented it on basic 6 features with 5754 instances

Table 3 shows the spread of the data for each attribute using statically measures followed by a plot (Figure 1) which shows the state wise no of cured, confirm and deaths record.

Measures	Cured	Death	Confirmed
Mean	220.3421	28.28947	879.1579

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

Standard Error	60.99688	12.59213	298.0092
Median	32.5	1	66.5
Mode	0	0	0
Standard Deviation	376.01	77.6231	1837.052
Sample Variance	141383.5	6025.346	3374760
Kurtosis	5.075574	20.81646	15.85465
Skewness	2.279225	4.339628	3.622577
Range	1593	432	9915
Minimum	0	0	0
Maximum	1593	432	9915
Sum	8373	1075	33408
Count	38	38	38

Table 3. Statistical distribution of Covid data

The statistical analysis of hypothesis testing is shown in the following table. The F-test statistics is a measure of how different the means are within each group.

Table 4. Statistical Analysis using Annova

Source of Variation	SS	df	MS	F	P-value	F critical
Between Groups	15135394	2	7567697	6.44577	0.002247	3.078057
Within Groups	1.3E+08	111	1174056	-	-	-
Total	1.45E+08	113		-	-	-

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

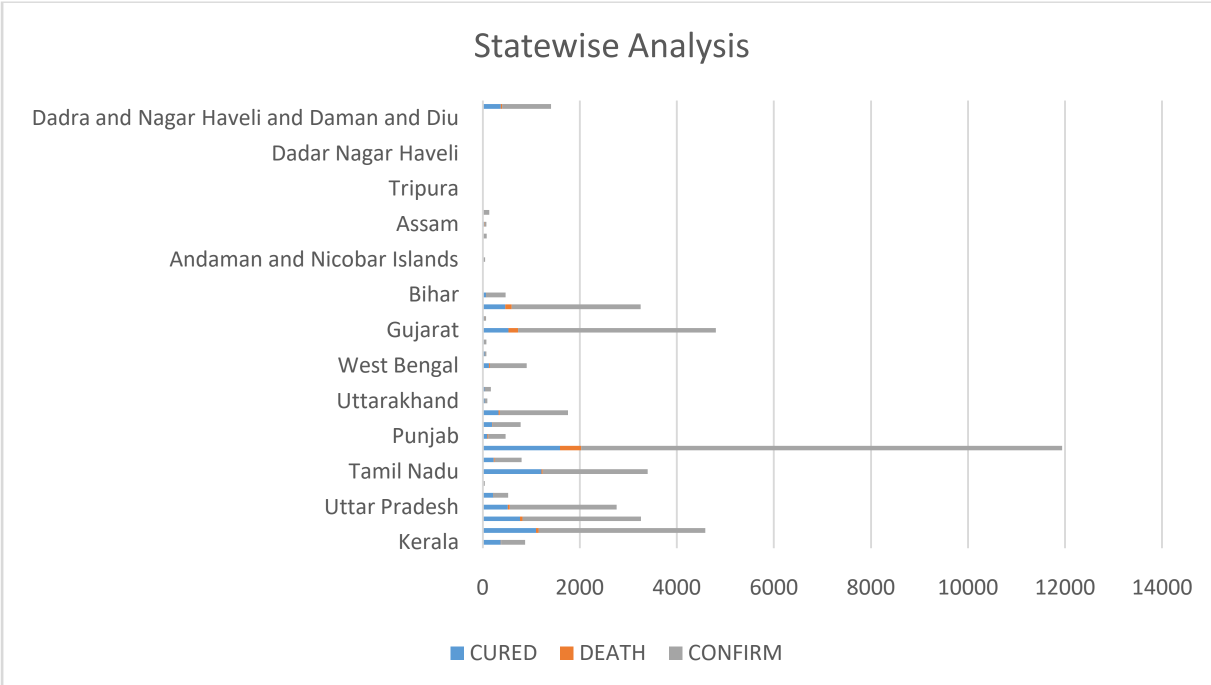


Figure 1. State-wise analysis of no of deaths, no of confirmed and no of cured patients.

The analysis shows a bivariate relationship exist between the confirm cases, cured cases and the death. In state wise analysis, Maharashtra state has most no of confirm cases after the surge of August month. Using our regression model, total cases predicted in Maharashtra would be 1.4 million by September 30, 2020. The aim of this analysis is to predict the suitable model for the data which exponentially rises after certain period. For such exponential rise the use of concept drift detection algorithms and evaluators are experimented. In Table 5, the conceit drift techniques are evaluated against the data set. The 3 main standards are used to evaluate the model, Accuracy and Kappa Statistics.

Table 5.Comparritive analysis based on performance evaluator and learner category on COVID-19 data

Algorithm	Category	Accuracy (%)	Kappa Statistics
DDM with Naïve Bayes	Learner	73.66	47.69
DDM with Naïve Bayes multinomial	Learner	61.42	44.30
Single classifier	Learner	73.52	44.39
Concept Drift performance evaluator	Evaluator	2	1.23
Basic regression performance evaluator	Evaluator	0.33	1.07
EWMA classic evaluation performance evaluator	Evaluator	1.07	505
Window classification performance model	Evaluator	0.58	0.76
Window Regression performance model	Evaluator	0.58	0.76

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

Conclusion

Handling uncertain and diversified data is a challenging problem in various commercial applications. The arriving pattern of such application falls under data stream mining category. Mining such Data stream always poses a new challenge that arises either with change in concept or a change in distribution of data. Traditional single classifier may not handle the change in distribution of data as a model needs to be retrained entirely to detect the changes. These multiple classifiers are built to train the diversified data so as to take better decision in terms of accuracy. The literature in this paper concludes that there are many approaches experimented by researchers using ensemble algorithms, where each weak classifier is replaced with new one by associating weights to those training samples. Various types of drift that occur in the application are handled using drift detection techniques like SEA , DDM, CVFDT,ADWIN algorithm. However more optimum classification models need to be designed with relevant statistical measures to handle sudden, gradual, recurrent drift in the arriving pattern of data samples.

Limitations and Future Studies

From the literature survey, it may be concluded as an emphasis of present work is on detecting the change data stream in arriving pattern using both parametric and non-parametric approach. However a parametric approach is applied on known data distribution where expected mean and standard deviation is known for the data sample. But in real life data, such data distribution does not come with known parameters. The work done so far using non parametric approaches are executed on synthetic data as getting real data set with a concept drift is difficult to get as it is not necessary that data distribution always comes with a sudden, gradual, recurring drift. Some data sets may come along with a noise which should not be treated as a change detection in data stream further. Hence more experiments to be applied on real time data sets also where a generalized performance must be achieved through ensemble and drift detection models. A better model needs to be designed by incorporating parametric and non-parametric approach. It is observed that in traditional data mining approach the fixed pair of parameter may be responsible for overall performance degradation where as a better online change detection method is needed for pre and post change detection ina sequential change detection model. Also a massive growth in data poses a challenge in the research community for making an optimal model performance [28]. Also the change in concept remains for a very short period of time [28] [29] , finding such small shifts in big data stream is still A concept drift model for big data stream is a challenging research area for the academicians and researchers.

References

[1] D. Brzezinski and J. Stefanowski, “Reacting to Different Types of Concept Drift :,” vol. 25, no. 1, pp. 81–94, 2014.

[2] R. Gull, U. Shoaib, S. Rasheed, W. Abid, and B. Zahoor, “Pre Processing of Twitter ’ s Data for Opinion Mining in Political Context,” *Procedia - Procedia Comput. Sci.*, vol. 96, no. September, pp. 1560–1570, 2016, doi: 10.1016/j.procs.2016.08.203.

[3] S. J. Delany, P. Cunningham, A. Tsymbal, and L. Coyle, “A case-based technique for tracking concept drift in spam filtering,” *Knowledge-Based Syst.*, vol. 18, no. 4–5, pp. 187–195, Aug. 2005, doi: 10.1016/j.knosys.2004.10.002.

[4] B. Zhang and Y. Chen, “Research on detection and integration classification based on concept drift of data stream,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2019, no. 1, Dec. 2019, doi: 10.1186/s13638-019-1408-2.

[5] J. Sarnelle, A. Sanchez, R. Capo, J. Haas, and R. Polikar, “Quantifying the limited and gradual concept drift assumption,” in *Proceedings of the International Joint Conference on Neural Networks*, 2015, vol. 2015-Septe, doi: 10.1109/IJCNN.2015.7280850.

[6] S. Wang, S. Schlobach, and M. Klein, “Concept drift and how to identify it,” *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 9, no. 3, pp. 247–265, Sep. 2011, doi: 10.1016/j.websem.2011.05.003.

[7] Y. Sun, Z. Wang, H. Liu, C. Du, and J. Yuan, “Online Ensemble Using Adaptive Windowing for Data Streams with Concept Drift,” vol. 2016, 2016, doi: 10.1155/2016/4218973.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.

[8] R. Elwell and R. Polikar, "Incremental learning in nonstationary environments with controlled forgetting," 2009, doi: 10.1109/IJCNN.2009.5178779.

[9] L. A. Deshpande and M. R. Narasingarao, "Addressing social popularity in twitter data using drift detection technique," *J. Eng. Sci. Technol.*, vol. 14, no. 2, 2019.

[10] M. Wozniak, P. Ksieniewicz, B. Cyganek, A. Kasprzak, and K. Walkowiak, "Active learning classification of drifted streaming data," *Procedia Comput. Sci.*, vol. 80, pp. 1724–1733, 2016, doi: 10.1016/j.procs.2016.05.514.

[11] X. Li, Y. Zhou, Z. Jin, P. Yu, and S. Zhou, "A Classification and Novel Class Detection Algorithm for Concept Drift Data Stream Based on the Cohesiveness and Separation Index of Mahalanobis Distance," *J. Electr. Comput. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/4027423.

[12] I. Khamassi and M. S. Moez, "Self-Adaptive Windowing Approach for Handling Complex Concept Drift," 2015, doi: 10.1007/s12559-015-9341-0.

[13] T. R. Hoens, R. Polikar, and N. V Chawla, "Learning from streaming data with concept drift and imbalance: An overview," *Prog. Artif. Intell.*, vol. 1, no. 1, pp. 89–101, 2012, doi: 10.1007/s13748-011-0008-0.

[14] J. Błaszczyński, M. Deckert, J. Stefanowski, and S. Wilk, "Integrating selective pre-processing of imbalanced data with Ivotes ensemble," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6086 LNAI, pp. 148–157, doi: 10.1007/978-3-642-13529-3_17.

[15] R. H. Rad and M. A. Haeri, "Hybrid Forest: A Concept Drift Aware Data Stream Mining Algorithm," pp. 1–25, 2019, [Online]. Available: http://arxiv.org/abs/1902.03609.

[16] G. Ditzler, G. Rosen, and R. Polikar, "Domain adaptation bounds for multiple expert systems under concept drift," in *Proceedings of the International Joint Conference on Neural Networks*, 2014, pp. 595–601, doi: 10.1109/IJCNN.2014.6889909.

[17] Y. Sun, H. Shao, and S. Wang, "E ffi cient Ensemble Classification for Multi-Label Data Streams with Concept Drift," 2019.

[18] S. Wares, J. Isaacs, and E. Elyan, "Data stream mining: methods and challenges for handling concept drift," *SN Appl. Sci.*, vol. 1, no. 11, Nov. 2019, doi: 10.1007/s42452-019-1433-0.

[19] D. M. Farid *et al.*, "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Syst. Appl.*, vol. 40, no. 15, pp. 5895–5906, Nov. 2013, doi: 10.1016/j.eswa.2013.05.001.

[20] A. Moubayed, E. Aqeeli, and A. Shami, "Ensemble-based Feature Selection and Classification Model for DNS Typo-squatting Detection," Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.09272.

[21] S. H. Pradeepini, G., Pradeepa, G., Tejanagasri, B., & Gorrepati, "Data classification and personal care management system by machine learning approach," *Int. J. Eng. Technol.*, vol. 7, no. 32, pp. 219–223, 2018.

[22] S. S. P., Sreedevi, M., & Amiripalli, "An ensemble feature selection framework of sonar targets using symmetrical uncertainty and multi-layer perceptron," *Adv. Intell. Syst. Comput.*, p. Potharaju, S., 2019.

[23] M. R. Sajana, T., & Narasingarao, "Majority voting algorithm for diagnosing of imbalanced malaria disease."

[24] M. R. Sajana, T., & Narasingarao, "Machine learning techniques for malaria disease diagnosis - A review," *J. Adv. Res. Dyn. Control Syst.*, vol. 9, pp. 349–369, 2017.

[25] E. Cesario, C. Mastroianni, and D. Talia, "Distributed volunteer computing for solving ensemble learning problems," *Futur. Gener. Comput. Syst.*, vol. 54, pp. 68–78, 2016, doi: 10.1016/j.future.2015.07.010.

[26] G. Ditzler, J. Labarck, J. Ritchie, G. Rosen, and R. Polikar, "Extensions to Online Feature Selection Using Bagging and Boosting," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 29, no. 9, pp. 4504–4509, 2018, doi: 10.1109/TNNLS.2017.2746107.

[27] M. R. (2019) Sajana, T., & Narasingarao, "Ensemble framework for classification of imbalanced malaria disease data.," *J. Adv. Res. Dyn. Control Syst.*, 2017.

[28] Janardan, Mehta S. Concept drift in Streaming Data Classification: Algorithms, Platforms and Issues. Procedia Comput Sci [Internet]. 2017;122:804–11. Available from: https://doi.org/10.1016/j.procs.2017.11.440

[29] Tennant M, Stahl F, Rana O, Bártolo J. Scalable real-time classification of data streams with concept drift. Futur Gener Comput Syst [Internet]. 2017;75:187–99. Available from: http://dx.doi.org/10.1016/j.future.2017.03.026

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4787 - 4800
Received 25 April 2021; Accepted 08 May 2021.