Journal of Engineering Sciences
Vol 10, Issue 11, Nov /2019
ISSN NO:0377-9254

POLYNOMIAL KERNELIZED FEATURE SELECTION AND TRIMMED BOOTSTRAP AGGREGATING CLASSIFIER FOR TIME SERIES DATA ANALYSIS

A.SELVAKUMAR [1], Dr.S.PRASATH [2]

1 Ph.D Research Scholar Part-time, Department of Computer Science, Nandha Arts and Science College, Erode, Tamilnadu, India.email:arshariniselva@gmail.com

2 Assistant Professor, Department of Computer Science, Nandha Arts and Science College, Erode, Tamilnadu, India. email:softprasaths@gmail.com

***Abstract*— With increasing trend of time series data accessibility, accurate prediction is a challenging one. In order to improve the prediction accuracy and minimize the complexity, a Polynomial Kernel Feature Selection based Trimmed Bootstrap Aggregating Data Classification (PKFS-TBADC) technique is introduced. The PKFS-TBADC technique includes two processes, namely Polynomial Kernel-based Feature Selection and Trimmed Bootstrap Aggregating Data Classification for predicting future outcomes. Polynomial Kernel Feature Selection process is carried out to select relevant features for performing data classification. Trimmed Bootstrap Aggregating Data Classification is carried out for performing the predictive analysis with time series data by constructing weak learner (i.e. C4.5). The weak learner results are combined to provide accurate classification with minimum out of sample error. This helps to improve prediction accuracy and reduce false-positive rate. Experimental evaluation is carried out on prediction accuracy, false-positive rate, time complexity, and space complexity with respect to number of time series data. From the results PKFS-TBADC technique achieves higher prediction accuracy with minimum complexity and false-positive rate than state-of-the-art methods.**

Keywords— Time Series Data, Feature Selection, Polynomial Kernel, data classification, Trimmed Bootstrap Aggregating algorithm, weak learner

1.INTRODUCTION

Time series data classification is an important research area in machine learning to predict future events with lesser complexity. A multivariate convolutional neural network (MVCNN) was developed in [1] for classifying time series data. But, accurate classification was not performed by MVCNN technique. Multivariate and Multi-Output Weighted Nearest Neighbor's (MV-kWNN) classification algorithm was designed in [2] for time-series data prediction. The algorithm failed to minimize prediction time and space complexity.

Cycle Deep Belief Network method was developed in [3] to categorize the time series data with minimum error rate. The designed method failed to consider the performance of complexity involved in data classification. A Long Short Term Memory and Fully Convolutional Network (LSTM-FCN) was introduced in [4] for categorizing the multivariate time series data. But the performance of false-positive rate remained unaddressed. A multiobjective model-metric (MOMM) learning was developed in [5] for time series

prediction through the classification. The model has high computational complexity and the performance of the classification was not improved. An information geometry framework was introduced in [6] to classify time series data in tangent space. The performance of framework was not enhanced due to high time complexity.

An effective online gradient learning model was designed in [7] for predicting time series data with minimum error. But, prediction time was not minimized. Relative Position Matrix and Convolutional Neural Network (RPMCNN) was developed in [8] for performing time series data classification task. But, designed method failed to use ensemble learning for processing more time series data. An XGBoost classifier based on shapelet features (XG-SF) was developed in [9] to enhance time series data classification accuracy. The designed technique failed to apply on big data platform. A fragment alignment distance algorithm was developed in [10] for time series data classification through similarity detection. But, classification accuracy was not improved.

The existing classification techniques have limitations such as lesser prediction accuracy, more complexity and higher false-positive rate and so on. In order to overcome the issues, PKFS-TBADC is introduced. The contribution of PKFS-TBADC technique is described as follows,

PKFS-TBADC technique is introduced to achieve higher prediction accuracy by applying polynomial kernel function. Kernel measures similarity between objective and features in input dataset with degree of polynomials. Features are mapped from input space into two subsets namely relevant and irrelevant. The relevant feature subset is selected for classification and irrelevant feature subset is removed to minimize time and space complexity.

PKFS-TBADC technique uses Trimmed Bootstrap Aggregating algorithm for accurate time-series data classification. Trimmed Bootstrap Aggregating classifier uses C4.5 decision tree as weak learner to classify data. Ensemble technique trimmed weak classifier with higher error rate to improve prediction accuracy

The rest of paper is structured as. In Section 2, problem definition and detail explanation of proposed PKFS-TBADC technique are presented. In Section 3, experimental is performed with time series dataset. The results of various parameters are discussed in Section 4. In section 5, reviews related works in time series data analysis. Finally, conclusion is presented in Section 6.

2. METHODOLOGY

POLYNOMIAL KERNEL FEATURE SELECTION BASED TRIMMED BOOTSTRAP AGGREGATING DATA CLASSIFICATION

Time Series data analysis is a challenging problem in data mining due to increasing nature of time series data availability. The number of features and instances presented in dataset may increase prediction risk. Dimensionality reduction minimizes number of irrelevant features in dataset. The relevant feature selection and irrelevant feature removal is significant one for dimensionality reduction. Based on this motivation, Polynomial Kernel Feature Selection based Trimmed Bootstrap Aggregating Data Classification (PKFS-TBADC) technique is introduced.

PKFS-TBADC technique comprises the two major processes namely feature selection and classification. The feature selection is performed using polynomial kernel function. Secondly, Trimmed Bootstrap Aggregating technique is used for classifying time series data with selected features for predicting future events. Bootstrap aggregating is a machine learning ensemble technique designed to improve accuracy of statistical classification. The two processes of PKFS-TBADC technique are described in following subsections.

2.1 POLYNOMIAL KERNEL FUNCTION BASED FEATURE SELECTION

Feature Selection is process of choosing relevant features for predicting future events. Kernel function measures similarity between feature vector and objective function (i.e. prediction). The process of polynomial kernel-based feature selection is shown in figure 1

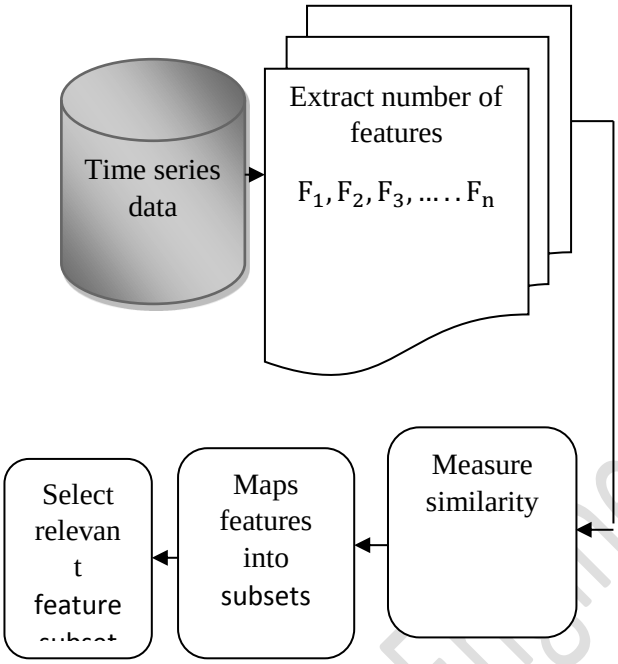

Figure 1 shows polynomial kernel-based feature selection through similarity measure. Let us consider time series dataset and number of features F1,F2,F3,…,Fn. The similarity is measured as,

$$k\,(F_i, O) = (F_i.O + b)^n \quad (1)$$

In (1), K(Fi,O) denotes kernel function i.e. similarity between features Fi and objective function 'O', Fi denotes feature vector in input space, b represents constant i.e. b>0, n denotes degree of polynomial. Based on similarity, kernel function is used for mapping features in input space into feature space. The mapping is expressed as,

$$\varphi : F_i \rightarrow F_s \quad (2)$$

In (2),φ denotes mapping function, Fi represents feature in input space, Fs denotes features in subset. In feature space, two subsets are obtained such as relevant (S1) and irrelevant (S2). The relevant features are selected for time series data prediction and In (2),φ denotes mapping function, Fi represents feature in input space, Fs denotes features in subset. In feature space, two subsets are obtained such as relevant (S1) and irrelevant (S2). The relevant features are selected for time series data prediction and removed other features to minimize time as well as space complexity. The polynomial kernel based feature selection algorithm is described as follows,

Input: Time series dataset D_t,
Output: select relevant features
Begin
1. **Extract the** number of features $F_1, F_2, F_3, \ldots F_n$ from D_t
2. **for** each feature F_i
3. Measure the polynomial kernel $k\,(F_i, O)$
4. Map the features into subsets $\varphi : F_i \rightarrow F_s$
5. Select relevant features subset
6. Remove the relevant features subset
7. **end for**

end

Algorithm 1 polynomial kernel-based feature selection algorithm

As given in the above algorithm, the feature selection process is carried out for predictive analytics using polynomial kernel function. The polynomial kernel measures the similarity between the features subset and the objective. The more similar features are mapped into the relevant feature subsets. The irrelevant feature subsets are removed therefore the complexity involved in classifying the results is said to be less and hence the accurate prediction is achieved.

2.2 TRIMMED BOOTSTRAP AGGREGATIVE CLASSIFIER

After selecting features, data classification is done with Trimmed Bootstrap Aggregative classifier. Bootstrap Aggregative classifier is an ensemble classifier used to improve classification accuracy by constructing weak learners. The weak learner is base classifier which provides less accurate classification results. Bagging is a strong classifier which provides accurate results. The proposed PKFS-TBADC technique uses ensemble classifier for time series data prediction. The structure of ensemble classifier is shown in figure 2.

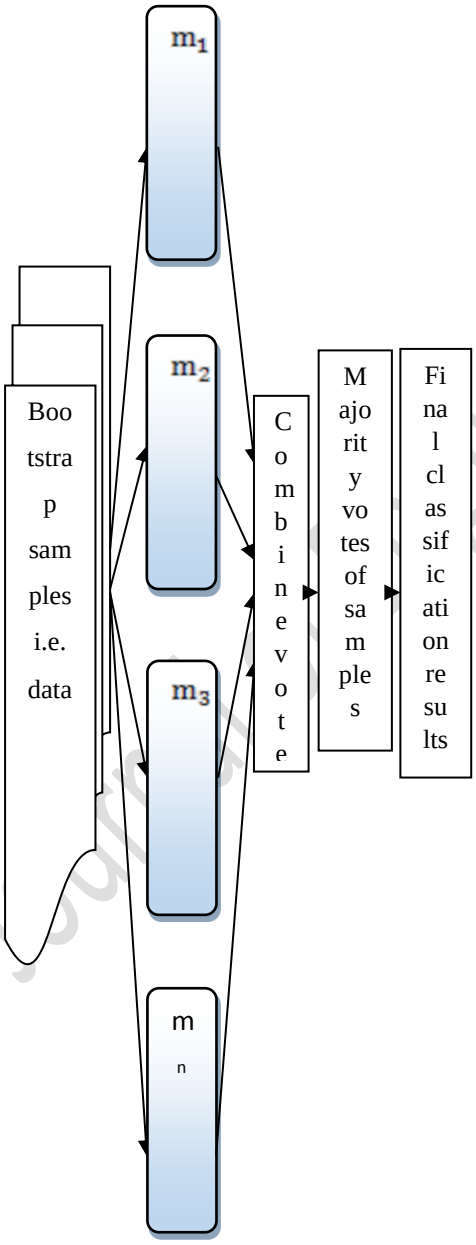

Figure 2 structure of the Trimmed bootstrap aggregating classifier

Figure 2 illustrates the structure of the Trimmed bootstrap aggregating classifier for classifying the given input data with higher accuracy. Let us consider the training sets $\{x_i,y_i\}$ where x_i denotes a bootstrap samples (i.e. input data with the extracted features) and y_i represents a classification results. The bootstrap aggregating classifier initially constructs an empty set of 'm' weak learners $m_1,m_2,m_3,\ldots m_m$. The trimmed bootstrap aggregating classifier uses the c4.5 decision tree as a weak learner to classify the time series data. C4.5 algorithm starts with root node. At every node of the tree, C4.5 selects the feature value i.e. data that efficiently classified into one class or the other. The splitting strategy of the algorithm is based on the mutual information. The attribute value with the value of mutual information is selected and creates the decision for data classification.

Let us consider the training feature value i.e. data $dt_1,dt_2,dt_3,\ldots dt_n$ as the input of the bootstrap aggregating classifier. The root node of the tree measures the mutual information between the training feature value and the testing feature value to perform the classification. The mutual information is accurately and efficiently shows the dependence between the training feature value and testing feature value. This kind of dependence metric is used to quantify the correlation and it is used as the judgment criterion of decision. The mutual information is mathematically calculated as follows

$$MI(dt_i, ds_j) = \sum p_r(dt_i, ds_j) \log \frac{p_r(dt_i,ds_i)}{p_r(dt_i)\, p_r(ds_i)} \quad (3)$$

In (3), MI denotes mutual information, dt_i denotes training feature value, ds_j represents the testing feature value, $p_r(dt_i,ds_j)$ denotes joint distribution probability, $p_r(dt_i)$ and $p_r(ds_i)$ denotes a marginal probability. Based on mutual information measure, root node classified data into different classes.

$$MI(dt_i, ds_j) = \begin{cases} 1, & dt_i, ds_j \text{ are dependent} \\ 0, & dt_i, ds_j \text{are independent} \end{cases} \quad (4)$$

By using (4), mutual information with category is 1 denotes two data is dependent. The mutual information returns '0' denoting two data is independent. Based on dependence measure, time series data are classified into two or more classes. Weak learner output has some training error in classification. Therefore, set of weak learner results are combined to make strong classification. The strong classification results obtained as,

$$y_i = \sum_{i=1}^{m} w_i\,(x) \quad (5)$$

In (5), y_i represents the ensemble classification results, $w_i(x)$ denotes an output of the decision tree classifier. After combining, the out of sample error is calculated for each classifier outcomes. The out of sample error is calculated as the difference between the expected and empirical (i.e. observed) error. The error rate is mathematically calculated as follows,

$$E_s = E_{ex} - E_{od} \quad (6)$$

From (6) E_s denotes an out of sample error, E_{ex} represents the expected error, E_{od} denotes an observed error. Based on the error value, the classifiers are arranged as follows,

$$E_s\big(w_1\,(x)\big) \le E_s\big(w_2\,(x)\big) \le E_s\big(w_3\,(x)\big) \le \cdots . E_s\big(w_m\,(x)\big) \quad (7)$$

From (7), the trimmed bootstrap aggregating classifier trims off the "worst" classifiers which having the largest error rates. Then the aggregating classifier takes an average of classification results and provides most accurate classifiers. The final output of trimmed bootstrap aggregating classification is given below,

$$y_i = \underset{m}{\text{avg}}\; w\,(x) \quad (8)$$

In (8), y_i represents the predicted output of final classification results, avg denotes an average function to find accurate classification of samples whose decision is known to m^{th} classifier. Therefore, bagging classifier improves prediction process and minimizes false positive rate. The algorithmic process of proposed ensemble classification results are shown below,

Input: Number of training data i.e. bootstrap samples $dt_1, dt_2, dt_3, \ldots . dt_n$,
Output: Improve prediction accuracy
Begin
1. Construct 'm' decision trees with training data $dt_1, dt_2, dt_3, \ldots . dt_n$
2. Find the mutual information between dt_i and ds_j
3. **If** $(MI(dt_i, ds_j) = 1)$ **then**
4. Dependence between dt_i and ds_j
5. **else**
6. Independence between dt_i and ds_j
7. **end if**
8. Combine a set of weak learners $\sum_{i=1}^{m} w_i\,(x)$
9. **For** each $w_i\,(x)$
10. Calculate the out of sample error 'E_s'
11. **end for**
12. Arrange $w_i\,(x)$ based on out of sample error E_s
13. Trim $w_i\,(x)$ with high error
14. Average all the classification results $y_i = \text{avg}_m\; w\,(x)$
15. Obtain strong classification results

end

Algorithm 2 trimmed Bootstrap Aggregating Data Classification

Algorithm 2 describes algorithmic process of trimmed bootstrap aggregating classifier. The ensemble classifier constructs set of weak classifiers with training data. The weak learner measures mutual dependence between training feature value and testing feature value. Based on mutual dependence, weak learner classifies time series data into different classes. The weak classification results are summed to make strong one. The out of sample error is calculated for each classification results. Weak learners are ordered based on error value. The ensemble classifiers trim highest error classification results and obtain strong results by averaging least error classification results to

improve the prediction accuracy and minimize false positive rate.

3.EXPERIMENTAL EVALUATION AND PARAMETER SETTINGS

An experimental evaluation of PKFS-TBADC technique and existing techniques namely MVCNN [1] and MV-kWNN [2] are carried out using Java language. The time-series data are collected from Hurricanes and Typhoons, 1851–2014 dataset (https://www.kaggle.com/noaa/hurricane-database) for cyclone prediction through the classification. The dataset consists of 22 attributes and 49,105 instances for Atlantic Ocean and 26,138 instances for eastern Pacific Ocean. The tropical cyclones in Atlantic Ocean are predicted through the experimental evaluations with time series data taken in the range from 1000-10000.

4.PERFORMANCE ANALYSIS AND RESULTS DISCUSSION

The comparative analysis of proposed PKFS-TBADC technique and existing MVCNN [1] and MV-kWNN [2] are discussed in this section. Performance analyses of PKFS-TBADC technique are compared with existing results with parameters, namely prediction accuracy, false positive rate, time complexity and space complexity with help of table or graph.

4.1 IMPACT OF PREDICTION ACCURACY

Prediction accuracy is measured as the ratio of numbers of (no. of) data correctly classified to the total number of data. It is measured in percentage and mathematically calculated as,

$$prediction\ accuracy = \left(\frac{No.of\ data\ correctly\ classified}{No.of\ data}\right) * 100 \quad (9)$$

Table 1 prediction accuracy versus the number of data

No. of data	Prediction accuracy (%)		
	PKFS-TBADC	MVCNN	MV-kWNN
1000	95	91	89
2000	94	90	88
3000	93	88	86
4000	92	86	84
5000	91	85	83
6000	90	84	82
7000	89	83	80
8000	88	82	78
9000	87	79	77
10000	86	78	76

Table 1 describes the performance results of prediction accuracy versus number of time series data taken in the range from 1000 to 10000. The ten different results of three methods PKFS-TBADC technique, MVCNN [1], MV-kWNN [2] are reported in table 1. The table values show that the prediction accuracy is considerably improved than the existing techniques. This improvement is achieved by applying trimmed bootstrap classifier. The ensemble classifier uses the empty set of weak learners to classify the data based on mutual information. The strong classification results are obtained by combining weak classifier which improves the cyclone prediction accuracy in Atlantic Ocean. The prediction accuracy is improved by 7% and 10% using PKFS-TBADC technique as compared to MVCNN [1], MV-kWNN [2] respectively.

4.2 IMPACT OF FALSE-POSITIVE RATE

False-positive rate (FPR) is measured as ratio of no.of data incorrectly classified to total no. of data for cyclone prediction. FPR is measured in terms of percentage (%) and computed as,

$$False\ positive\ rate = \left(\frac{No.of\ data\ incorrectly\ classified}{No.of\ data}\right) * 100 \quad (10)$$

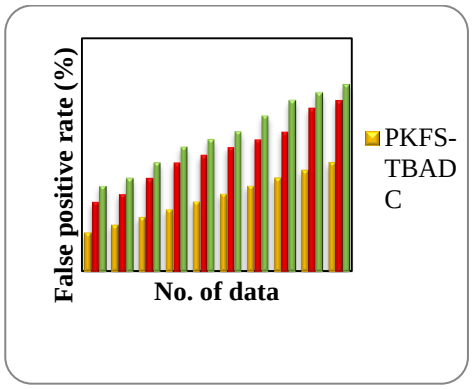

Figure 3 Comparison of the false-positive rate

Figure3 illustrative that the performance results of the false-positive rate with respect to a number of data taken from the dataset. The graphical results infer that the false positive rate is found to be minimized using PKFS-TBADC technique than two existing methods. This is because of using trimmed bootstrap classifier. The trimming ability of classifier provides accurate classification results than conventional technique. The proposed ensemble classifier combines all weak classifier results and measures out of sample error. The larger error rates of the weak classifier results are trimmed. The classification results with a lesser error are selected as final results. This helps in reducing the false positive rate. The false positive rate is minimized by 39% and 47% when compared to existing MVCNN [1], MV-kWNN [2] respectively.

4.3 IMPACT OF TIME COMPLEXITY

Time complexity is referred to as the amount of time taken for cyclone prediction through the classification of the weather data. The time complexity of data classification is calculated as,

$$TC = No.of\ data * T\ (classifying\ single\ data) \quad (11)$$

From (11), TC represents the time complexity, 'T' denotes a time for classifying single data .The time complexity is measured in the unit of milliseconds (ms).

Table 2 time complexity versus no. of data

No. of data	Time complexity (ms)		
	PKFS-TBADC	MVCNN	MV-kWNN
1000	26	29	32
2000	30	34	36
3000	32	36	39
4000	34	41	44
5000	38	43	45
6000	41	45	48
7000	45	48	53
8000	50	53	56
9000	53	55	61
10000	58	61	65

Table2 given above shows the experimental results of time complexity with the number of data in the range from 1000 to 10000. The results clearly show that the time complexity of PKFS-TBADC technique is minimized than conventional classification techniques. This improvement is achieved by performing the feature selection. The dataset with more features causes more time complexity in cyclone prediction. PKFS-TBADC technique uses polynomial kernel to map the features in the input space into subsets through the similarity measure. The ensemble technique classifies the data with the selected relevant features resulting in minimizes the time complexity. From this result, time complexity using PKFS-TBADC technique was lesser by 9% when compared to state-of-the-art methods.

4.4 IMPACT OF SPACE COMPLEXITY

Space complexity is measured as an amount of memory space taken by the algorithm for cyclone prediction. The formula for calculating the space complexity is measured as follows,

$$space\ complexity = no.of\ data * space\ (storing\ one\ data) \quad (12)$$

From (12), the space complexity is measured in terms of megabytes (MB).

Figure 5 given below illustrates the graphical representation of the space complexity of ten different results.

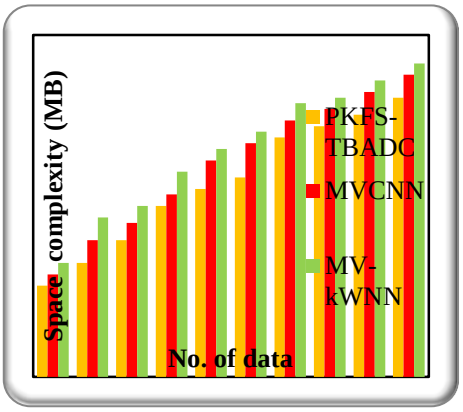

Figure 4 comparison of space complexity

Figure4 depicts graphical representation of space complexity using PKFS-TBADC technique, MVCNN [1], MV-kWNN [2]. The results inferred that space complexity is found to be comparatively lesser using the PKFS-TBADC when compared to existing classification techniques. This is due to application of polynomial kernel-based feature selection and ensemble classification. By applying two algorithms, dimensionality of data is minimized. The more relevant features and their data used for classification process. Followed by, ensemble classifier performs cyclone prediction with lesser complexity. The average results shows space complexity of PKFS-TBADC technique is minimized by 10% than tMVCNN [1] and 17% when compared to MV-kWNN [2].

5. RELATED WORKS

The hidden-unit logistic model (HULM) was designed in [11] for multivariate time-series classification. But, performance of the data classification was not improved. A Hybrid appliance classification technique was developed in [12] for time series data with features extraction. Though the technique minimizes the false positive rate, the time complexity was not minimized. A multi-objective optimization method was designed in [13] for classifying the time series data. But, the prediction was not improved with lesser time consumption. A Distributed FastShapelet Transform algorithm was designed in [14] to improve time series data classification based on the MapReduce. Though the algorithm minimizes the complexity, the accuracy was not improved. A recurrent kernel extreme reservoir machine (RKERM) was developed in [15] for improving the prediction accuracy using time series data. But, performance of error rate was not minimized.

A self-constructing recurrent fuzzy neural network (SCRFNN) was developed in [16] for improving the time-series data prediction accuracy with minimum error rate. The complexity involved in the prediction remained unsolved. An improved Elman neural network was introduced in [17] for time series data prediction with minimum error. But, prediction accuracy was minimized as it failed to perform the relevant feature selection. An evolutionary attention-based LSTM was developed in [18] for multivariate time series prediction. However, the performance of the algorithm was not improved. A novel distributed processing framework was designed in [19] for Big Data probabilistic classification and regression. But, prediction accuracy was not improved. An efficient confidence measure based classification of time series was performed in [20]. However, accurate classification was not performed with minimum time.

6. CONCLUSION

The novel technique called PKFS-TBADC is developed for big-time series data forecasting with higher accuracy and minimum complexity. By designing a polynomial kernel, the features which are more relevant to the prediction are selected and other features are removed. Therefore, this assists to minimize the complexity in the time series data prediction. Next, the ensemble

classification technique is applied to predict future outcomes by constructing the number of weak learners. The weak learners are used for classifying the bootstrap samples based on mutual information. Then the results of the weak classifier are summed and to make a strong by trimming the classifiers with higher out of sample error. As a result, the accurate classification is performed resulting in improves the prediction results. The performance of the PKFS-TBADC technique and existing classification techniques is estimated with different metrics such as prediction accuracy, false-positive rate, time complexity, and space complexity. The experimental results illustrate that when comparing with the state-of-the-art methods, the PKFS-TBADC technique achieves higher prediction accuracy and minimum time, space complexity as well as false-positive rate.

REFERENCES

[1] Chien-Liang Liu, Wen-Hoar Hsaio, and Yao-Chung Tu, "Time Series Classification with Multivariate Convolutional Neural Network", IEEE Transactions on Industrial Electronics Volume 66, Issue 6, 2019, Pages 4788 - 4797

[2] R.Talavera-Llames, R.Pérez-Chacón, A.Troncoso, F.Martínez-Álvarez, "MV-kWNN: A novel multivariate and multi-output weighted nearest neighbours algorithm for big data time series forecasting", Neurocomputing, Elsevier, 2019, Pages 1-40

[3] Shuqin Wang, Gang Hua, Guosheng Hao, and Chunli Xie , "A Cycle Deep Belief Network Model for Multivariate Time Series Classification", Mathematical Problems in Engineering, Hindawi, Volume 2017, October 2017, Pages 1-7

[4] Fazle Karim, Somshubra Majumdar, Houshang Darabi, Samuel Harford, "Multivariate LSTM-FCNs for time series classification", Neural Networks, Elsevier, Volume 116, 2019, Pages 237–245

[5] Zhichen Gong, Huanhuan Chen, Bo Yuan, Xin Yao, "Multiobjective Learning in the Model Space for Time Series Classification", IEEE Transactions on Cybernetics, Volume 49, Issue 3, 2019, Pages 918 – 932

[6] Jiancheng Sun, Yong Yang, Yanqing Liu, Chunlin Chen, Wenyuan Rao, Yaohui Bai "Univariate time series classification using information geometry", Pattern Recognition, Elsevier, Volume 95, 2019, Pages 24-35

[7] Haimin Yang, Zhisong Pan, and Qing Tao, "Robust and Adaptive Online Time Series Prediction with Long Short-Term Memory", Computational Intelligence and Neuroscience, Hindawi, Volume 2017, December 2017, Pages 1-9

[8] Wei Chen and Ke Shi, "A deep learning framework for time series classification using Relative Position Matrix and Convolutional Neural Network", Neurocomputing, Elsevier, Volume 359, 2019, Pages 384-394

[9] Cun Jia, Xiunan Zou, Yupeng Hu, Shijun Liu, Lei Lyu, Xiangwei Zheng, "XG-SF: An XGBoost Classifier Based on Shapelet Features for Time Series Classification", Procedia Computer Science, Elsevier, Volume 147, 2019, Pages 24–28

[10] Miaomiao Zhang and Dechang Pi, "A New Time Series Representation Model and Corresponding Similarity Measure for Fast and Accurate Similarity Detection", IEEE Access, Volume 5, 2017, Pages 24503 – 24519

[11] Wenjie Pei, Hamdi Dibeklioğlu, David M. J. Tax, Laurens van der Maaten, "Multivariate Time-Series Classification Using the Hidden-Unit Logistic Model", IEEE Transactions on Neural Networks and Learning Systems, Volume 29, Issue 4, 2018, Pages 920 - 931

[12] LiuHui, Wu Haiping, Yu Chenming, "A hybrid model for appliance classification based on time series features", Energy and Buildings, Elsevier, Volume 196, 2019, Pages 112-123

[13] U.Mori, A.Mendiburu, I.M.Miranda, J.A.Lozano, "Early classification of time series using multi-objective optimization techniques", Information Sciences, Elsevier, Volume 492, 2019, Pages 204-218

[14] Francisco J.Baldán and José M.Benítez, "Distributed FastShapelet Transform: a Big Data time series classification algorithm", Information Sciences, Elsevier, Volume 496, 2019, Pages 451-463

[15] Zongying Liu, Chu Kiong Loo, Naoki Masuyama, Kitsuchart Pasupa, "Recurrent Kernel Extreme Reservoir Machine for Time Series Prediction", IEEE Access, Volume 6, Pages 19583 - 19596

[16] Qinghai Li and Rui-Chang Lin, "A New Approach for Chaotic Time Series Prediction Using Recurrent Neural Network", Mathematical Problems in Engineering, Hindawi Publishing Corporation, Volume 2016, October 2016, Pages 1-9

[17] Ying Zhang, Xiaoping Wang, Huiming Tang, "An improved Elman neural network with piecewise weighted gradient for time series prediction", Knowledge-Based Systems, Elsevier, Volume 181, 2019, 1047, Pages 1-10

[18] Youru Li, Zhenfeng Zhu, Deqiang Kong, Hua Han, Yao Zhao, "A new method to mitigate data fluctuations for time series prediction", Applied Mathematical Modelling, Elsevier, Volume 65, 2019, Pages 390-407

[19] Georgios Chatzigeorgakidis, Sophia Karagiorgou, Spiros Athanasiou and Spiros Skiadopoulos, "FML‐kNN: scalable machine learning on Big Data using k‐nearest neighbor joins" Journal of Big Data, Springer, Volume 5, Issue 4, 2018, Pages 1-27

[20] Junwei Lv, Xuegang Hu, Lei Li, Peipei Li, "An Effective Confidence-Based Early Classification of Time Series", IEEE Access, Volume 7, 2019, Pages 96113 - 96124