# Efficient time series data classification using sliding window technique based improved association rule mining with enhanced support vector machine

**Article** *in* International Journal of Engineering & Technology · June 2018

DOI: 10.14419/ijet.v7i2.33.13890



2 authors:

D. Senthil



Suseendran G.

VELS INSTITUTE OF SCIENCE, TECHNOLOGY & ADVANCED STUDIES (VISTAS), CHE…

*International Journal of Engineering & Technology*, 7 (2.33) (2018) 218-223

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Efficient time series data classification using sliding window technique based improved association rule mining with enhanced support vector machine

**D. Senthil [1] *, G. Suseendran [2]**

*[1] Ph.D., Research Scholar, Department of Information and Technology, School of Computing Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, India*
*[2] Assistant Professor, Department of Information and Technology, School of Computing Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, India, Chennai, India*
**Corresponding author E-mail: senthildphd@gmail.com*

## Abstract

Time series analysis is an important and complex problem in machine learning and statistics. In the existing system, Support Vector Machine (SVM) and Association Rule Mining (ARM) is introduced to implement the time series data. However it has issues with lower accuracy and higher time complexity. Also it has issue with optimal rules discovery and segmentation on time series data. To avoid the above mentioned issues, in the proposed research Sliding Window Technique based Improved ARM with Enhanced SVM (SWT-IARM with ESVM) is proposed. In the proposed system, the preprocessing is performed using Modified K-Means Clustering (MKMC). The indexing process is done by using R-tree which is used to provide faster results. Segmentation is performed by using SWT and it reduces the cost complexity by optimal segments. Then IARM is applied on efficient rule discovery process by generating the most frequent rules. By using ESVM classification approach, the rules are classified more accurately.

***Keywords***: *Time Series Data; Sliding Window; Indexing; IARM; ESVM; Segmentation.*

## 1. Introduction

Time series is an essential class of fleeting information items and it can be effectively gotten from logical and budgetary applications. A period arrangement is a gathering of perceptions made sequentially. The idea of time arrangement information incorporates: expansive in information measure, high dimensionality and important to refresh ceaselessly. Also time arrangement information, which is described by its numerical and persistent nature, is constantly considered in general rather than individual numerical field. The expanding utilization of time arrangement information has started a lot of innovative work endeavors in the field of information mining. The bounteous research on time arrangement information mining in the most recent decade could hamper the section of intrigued specialists, because of its multifaceted nature [1].

Time series data miningt information mining structure is a basic commitment to the fields of time arrangement investigation and information mining as of late. Techniques in light of the time arrangement information mining system can effectively portray and foresee complex, no occasional, unpredictable, et cetera [2]. The strategies conquer confinements to be specific stationarity and linearity necessities of conventional time arrangement examination systems by adjusting information digging ideas for dissecting time arrangement.

The chief properties of time series are such as stationarity, linearity, trend, seasonality, cycle variation and irregular fluctuation. Stationarity as a property of time series data suggests of uniformity in pattern, free of trend or seasonality. Linearity indicates the state of the time series where datasets are free of biases and the input and output maintain a linear function [3]. Trend is a long-term tendency in time series. It shows the decrease or increase in the mean value of the forecast variable over a period of time. Seasonality in time series manifests itself in the form of periodic fluctuating pattern. Cyclic variation means oscillatory movement about the trend level having phases from peak to contradiction to trough to expansion. Irregular fluctuation is erratic and random without having any specific pattern. It is short-term by nature and remains unanticipated.

Data mining is the examination of information with the objective of revealing shrouded designs. Information Mining includes an arrangement of techniques that robotize the logical revelation process. Its uniqueness is found in the kinds of issues tended to those with expansive informational indexes and mind boggling, concealed connections [4]. Information mining originates from the attention on finding concealed example. The hypotheses of time arrangement incorporate investigating direct, stationary time arrangement, forecast and so forth. This strategy uses, for example, bunching, order, affiliation lead mining and probabilistic graphical reliance models to distinguish concealed and helpful data from substantial databases.

The time series depends on past estimations of the staying variable yet not on illustrative factors, which may influence the arrangement. In actuality, circumstances, meteorological information that are seen on hourly premise, every day stock costs, and clinical perceptions are untouched arrangement [5]. Additionally, a period arrangement database is a grouping database, where it might contain successions of requested levels with or without solid idea of time. There are two primary goals of time arrangement examination, for example, recognizing the idea of wonder spoke to by the groupings of perceptions and determining future estimations of the time arrangement variable.

In [6] grouped bunching strategies produced for giving different static information into five noteworthy classes: dividing techniques, various leveled techniques, thickness based techniques, lattice based strategies, and model-based techniques. Three of the five noteworthy classifications of bunching techniques for static information, particularly parceling strategies, progressive techniques, and model based strategies, have been used specifically or changed for time arrangement grouping.

Sliding window is a transitory estimate over the real estimation of the time arrangement information [7]. The measure of the window and fragment increments until the point when we achieved the less mistake guess. In the wake of choosing the primary section, the following fragment is chosen from the finish of the main portion. The procedure is rehashed until record-breaking arrangement information are portioned. It is a notable time arrangement information division technique, in which a fragment with a blunder limit and settled window measure is made when the change point is come to [8]. In real information, for example, climate information, sliding window is unacceptable on the grounds that proper blunder edge and change point are required to evade data misfortune.

Association rule mining (ARM) broadly utilized for mining vast databases is initially acquainted with handle showcase bushel information for shopper acquiring designs in different application zones. In [9] utilized new approach of mining transient affiliation rules. In customary affiliation administer mining calculations, if the estimation of least help is set too high, one may lose heaps of profitable standards. In any case, in the event that it is set too low, numerous inconsequential standards will be mined, and it is difficult to recognize which ones are significant. When thinking about transient factors, an itemset may not be visit over the whole database but rather might be visit in some particular interims. They utilized a fleeting affiliation control digging calculation for interim successive examples, which can naturally create the majority of the interims without utilizing any space information.

## 2. Related works

In [10], Sengupta et al (2013) researches the convenience of time arrangement grouping systems to diminish the computational multifaceted nature of savvy framework improvement issues. The strategies center around the request side issue of neighborhood stockpiling measuring for sustainable combination, while featuring the significance and general materialness of these methods. It assembles and conveys an online choice emotionally supportive network to energize the arrangement of housetop sun based. In any case it has issue with cost intricacy for extensive dataset.

In [11] Niennattrakul et al (2007) talked about the bunching of interactive media time arrangement utilizing K-means and K-medoids calculations with dynamic time distorting and showed that K-implies is substantially more bland grouping technique when Euclidean separation is utilized, however it neglected to give amend comes about when dynamic time twisting is utilized as separation measure fit as a fiddle of the time arrangement. As the aftereffects of their investigations, they have affirmed that dynamic time twisting ought not be utilized as subroutine with K-implies calculation and K-medoids with dynamic time distorting gives attractive outcomes.

In [12] Chandrakala et al (2008) talked about a thickness based bunching technique in bit include space for grouping multivariate time arrangement information of fluctuating length. This strategy can likewise be utilized for grouping any sort of organized information, gave a portion which can deal with that sort of information is utilized. It presents heuristic strategies to locate the underlying estimations of the parameters utilized as a part of our proposed calculation. To demonstrate the adequacy of this technique, this strategy is connected to two diverse online manually written character informational indexes which are multivariate time arrangement information of changing length, as a certifiable application. The execution of this technique is contrasted and the phantom bunching and portion k-implies grouping strategies. Other than taking care of anomalies, this technique executes and additionally the unearthly grouping strategy and outflanks the part k-implies bunching technique.

In [13] Viet et al (2013) depicts the utilization of another multidimensional file tree, M-tree, for quick recovery in wavelet changed time arrangement and looks at its execution as a record structure for this changed time arrangement portrayal. The major inventive property of M-tree is that question execution in M-tree can be improved to lessen the quantity of separation calculations. Trial comes about uncover it has more opportunity for list building.

In [14] Winarko et al (2008) examine the reasonableness of various systems which can be connected for post-handling of worldly examples. As a post-preparing strategy, this technique is to store the created designs in a database and later question the database for chose designs. At the point when the database of examples is little a successive output of the database gives a pleasant execution, however as the database develops, the exhibitions fall apart and lists should be developed to speed the questions. It talks about the appropriateness of different ordering systems accessible for set-esteemed characteristics, which can be connected to database of worldly examples. The creators have connected two unique usage of mark records, in particular Sequential Signature Files (SSF) and Bit-Slice Signature Files (BSSF) as files.

In [15] Mokbel et al (2005) talked about ceaseless question processor incorporates: (1) New incremental spatio-transient administrators to help a wide assortment of constant spatio-fleeting inquiries, (2) Extended semantics of sliding window questions to manage spatial sliding windows and also worldly sliding windows, and (3) A common execution structure for versatile execution of an arrangement of simultaneous nonstop spatio-fleeting inquiries. Exploratory assessment indicates promising execution of the consistent question processor of the PLACE server.

In [16], Miao et al (2010) created Apriori-broadened mining occasional worldly affiliation rules (MPTAR) as indicated by the particular periodicity of information. The test on a gathering of money related information demonstrates that the strategy is helpful and effective. It is more critical for enhancing the hypothesis and execution of worldly affiliation manages in information mining.

In [17] Cao et al (2003) presents the help vector machines (SVMs) specialists for time arrangement guaging. The summed up SVMs specialists have two-organize neural system engineering. In the 3rst phase, self-sorting out component outline is utilized as a grouping calculation to parcel the entire information space into a few disconnected locales. A tree-organized design is embraced in the segment to keep away from the issue of foreordaining the quantity of parceled locales. At that point, in the second stage, various SVMs, additionally called SVM specialists. The reenactment demonstrates that the SVMs specialists accomplish signi3cant change in the speculation execution in examination with the single SVMs models. What's more, the SVMs specialists likewise focalize speedier and utilize less help vectors.

## 3. Proposed methodology

### 3.1. Preprocessing using modified K-Means clustering (MKMC)

In this research, modified k-means clustering algorithm is proposed for time series dataset. The time series average/mean is required by the algorithm and must be obtained. Euclidean distance is computed to remove the redundancy attributes, missing value replacement and finding wrong values for the given time series dataset.

MKMC method is designed to improve the time, number of iterations and sum of squared errors and this provides much better result as compare to normal K Mean clustering. In MKMC method, data is reduced by normalized method and then the parameters such as time taken, number of iterations, sum of squared errors are improved by using normalized technique. Normalization index method used for modification. Normalization is an important preprocessing step in time series data to standardize the values of all variables from dynamic range into specific range. Then compute the Euclidean distance between different clusters. After that normalization is done to minimize the sum of squared errors and no. of iteration resulting in less execution time to make the grouping of clusters. The Fig 1 shows the overall block diagram of the proposed system.

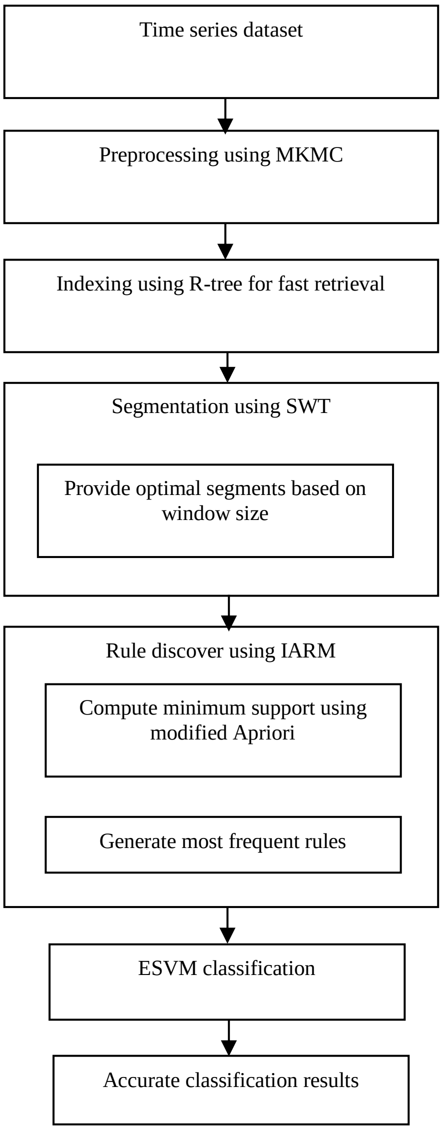


**Fig. 1:** Overall Block Diagram of the Proposed System.

The MKMC tries to separate N information objects into k parcels or groups, where each would have one protest (mean) as its bunch focus, speaking to all information questions inside that group. We at that point dole out whatever remains of the articles to legitimate groups and recalculate new focuses. It rehashes this progression until the point that all bunch focuses are steady. All in all, after every emphasis, the nature of the bunches and the methods themselves will basically be progressed.

MKMC algorithm

Input: Time series dataset

Output: Clusters

Dataset D $= \{\mathrm{d1}, \mathrm{d2}, \ldots \mathrm{d3}\}$

Find the maximum and minimum values of each feature from the dataset.

Normalize real scalar values of datasets with maximum and minimum values using equation

$$\mathrm{v}' = \frac{\mathrm{v}-\min(\mathrm{a})}{\max(\mathrm{a})-\min(\mathrm{a})} \qquad (1)$$

Where min (e) and max (e) are the minimum and the maximum values for attribute A.

Calculate the average score of each data point.

$$\mathrm{di} = \mathrm{d1}, \mathrm{d2}, \ldots.\mathrm{d3}$$

$$\mathrm{di(avg)} = (\mathrm{w1} * \mathrm{d1} + \mathrm{w2} * \mathrm{d2} + \cdots \ldots \mathrm{wm} * \mathrm{dm})/\mathrm{m} \qquad (2)$$

Where d is attribute value, m is number of attributes, w is weight to multiply to ensure fair distribution of cluster. Compute the missing value and redundancy value using (1) and (2)

Sort the data based on average score

Divide the data based on k cluster

Find the closest centroid for each data points and allocate each data points Cluster.

Calculate new centroid for each data

Take the nearest possible data point of the mean as the initial centroid for each data subsets.

In this algorithm, found the number of cluster then assign data points to initial cluster. It finds an effective and useful initial centroid. It helps to determine the number of cluster and which points are the initial centroids. It provides feasible data points which have chance to change current cluster and move to new cluster. Based on the MKMC normalized algorithm, it removes the noise data effectively from the time series dataset. It finds the missing values and repeated values using normalized based MKMC approach. It is used to increase the time series classification accuracy.

### 3.2. Indexing using r-tree algorithm

In this exploration, ordering is finished by utilizing R-tree calculation to enhance the execution. R-tree is a tree information structure ordinarily utilized for putting away and ordering of time arrangement information. It has likewise been utilized as a device for ordering multidimensional information. Since time arrangement can be effortlessly spoken to as an arrangement of n-dimensional information focuses, R-trees have been broadly utilized as a part of earlier research for time arrangement similitude look. It is enhanced for enhanced execution by building a superior quality tree.

The ordering structure is developed from the information time arrangement information beforehand fragmented into an arrangement of subsequences. Contingent upon the application, subsequences can have different degrees of covers. Once the information grouping is fragmented, a R-tree record is built and spared to the database. Information hubs are fit for putting away different information pointers. This outline enables the framework to utilize a similar hub to record different indistinguishable sections with the goal that the extent of the tree can be diminished to enhance general execution. it is as area, measurement and kid hub pointers. It is gone for enhancing the nature of the tree by limiting the cover between the hubs on a similar level. This component is engaged to tackle the issue of ordering and questioning expansive information arrangement accumulations.

### 3.3. Segmentation using sliding window technique

The sliding window technique is used to provide more compact representation though efficient segmentation. Time series segmentation is a fundamental component in the process of analysis and research of time series data. Its relevance should especially be viewed in the context of implications on the creation of a valid model of time series. As a data mining research problem, segmentation focuses on dividing the time series into appropriate, inter-

nally homogenous segments, so that the structure of time series, through pattern and/or rule discovery in the behaviour of the observed variable, could be revealed.
Minimizing the sum of squares of these distances the Euclidean distance also minimizes, which is, precisely, defined as the square root of the sum of squared distances between empirical and approximated values. Therefore, the essence of segmentation problem is reflected in finding the optimal approximation for which the error function, Ep, is minimal. In fact, the optimal segmentation of time series, T, for the defined parameters is defined as the segmentation that results in the lowest segmentation error in relation to other possible combinations of segmentation

$$S_{optimal}(T,k) = arg\, min_{S\in n,k} E_p(T,S) \quad (3)$$

Given the specificity of necessities in the investigation of huge time arrangement informational collections, and assortment of prerequisites that can be characterized in a type of division issue, finding the ideal arrangement

$S_{optimal}$.

In the division stage, the sliding window calculation works by tying down the left purpose of a potential fragment at the primary information purpose of a period arrangement, at that point endeavoring to inexact the information to one side with expanding longer sections. Eventually I, the blunder for the potential fragment is more prominent than the client indicated edge, so the subsequence from the grapple to I-1 is changed into a section. The grapple is moved to area I, and the procedure rehashes until the point that the whole time arrangement has been changed into a piecewise straight estimation
Sliding window algorithm

Seg_T,S=sliding window (T, max_error)

Anchor =1

While not finished segmenting time series

i = 2;

While calculate_error(T[anchor: anchor + i ]) < max_error

i = i + 1;

End;

Seg_TS = concat(Seg_TS, create_segment(T[anchor: anchor + (i-1)]);

Anchor = anchor + i;

End;

Contingent upon the blunder measure utilized, there might be different enhancements conceivable. Since the remaining blunder is monotonically non-diminishing with the expansion of more information focuses, one doesn't need to test each estimation of I from 2 to the last picked esteem. The calculation proposes at first setting I to s, where s is the mean length of the past portions. In the event that the figure is critical (the deliberate mistake is still not exactly max_error) at that point the calculation proceeds to increase I as in the great calculation. Else they start to decrement I until the point when the deliberate mistake is not exactly max_error. This streamlining can enormously accelerate the calculation if the mean length of fragments is vast in connection to the standard deviation of their length. It is utilized to diminish both the space and the computational cost of putting away and transmitting such information.

### 3.4. Rule discovery using IARM

In this exploration, ideal based Apriori calculation is proposed named as enhanced ARM. Apriori calculation is the most broadly utilized calculation for affiliation run mining. In Apriori calculation, right off the bat the competitor itemsets are created. At that point the database is examined for checking the help of these itemsets to create visit 1-itemset. Amid the principal check, 1-itemsets are created by dismissing those itemsets whose help is underneath the edge. In the ensuing passes, the competitor k-itemsets are produced after (k-1)th ignore the database by joining (k-1) itemsets. The pruning of the non-interesting itemsets is finished by the Apriori property which expresses that the subset of an incessant itemset should likewise be visit. In any case, it has impediments as takes after.
It just clarifies the nearness and nonattendance of a thing in value-based databases. If there should arise an occurrence of substantial dataset, this calculation isn't proficient. In Apriori, all things are dealt with similarly by utilizing the nearness and nonappearance of things. Apriori calculation requires substantial number of outputs of dataset. If there should arise an occurrence of substantial dataset, Apriori calculation creates vast number of applicant itemsets. Calculation examine database over and again to search visit item-sets, so additional time and asset are required in expansive number of sweeps so it is wasteful in substantial time arrangement datasets.
Consequently to beat the previously mentioned issue, in this examination, advanced based Apriori calculation is upgraded. It helps in enhancing the productivity of Apriori calculation. It diminished the measure of unique enormous dataset adequately and subsequently exchanges that don't comprise of regular item-sets are of no significance in the following sweeps for looking successive item-sets. It is utilized to decrease the quantity of outputs and evacuates substantial hopeful which drives bring down time intricacy.

1) Represent the database as a matrix where each row represents the transaction of the database and each column represents the item in the database.
2) Count the occurrences of each item in the database.
3) L1=find_frequent_1-itemsets(D);
4) For(k=2;Lk-1≠Φ; k++){
5) Ck=Apriori_gen(Lk-1, min_sup);
6) for each transaction t∈D{
7) Ct=subset(Ck,t);
8) for each candidate c∈Ct
9) c.count++;
10) }
11) Lk={ c∈Ck |c.count≥min_sup };
12) if(k>=2){
13) delete_datavalue(D, Lk, Lk-1);
14) delete_datarow (D, Lk); }
15) }
16) 16.return L value
17) Discover rules based on length and support count
18) find the frequent value and select the best rules
19) prune the unnecessary and redundancy rules
20) Produce the optimal rules for given time series dataset

The above algorithm describes that the most frequent rule generation using IARM algorithm. It is used to extract the high confidence rules from the frequent itemsets.

### 3.5. Classification using ESVM approach

Upgraded Support Vector Machines (ESVM) is utilized for time arrangement forecast. The ESVM classifier is positively contrasted with other arrangement approaches by performing piece based approval. The impact of the quantity of tenets to the arrangement rate is additionally examined for ongoing datasets.
Bolster vector machineis a regulated learning procedure that looks for an ideal hyperplane to isolate two classes of tests. Portion capacities are utilized to outline include information into a higher

measurement space where the information should have a superior conveyance, and afterward an ideal isolating hyperplane in the high dimensional element space is picked.

A very much prepared portion bolster vector machine is improved (ESVM) for ordering the principles in the time arrangement information. In ESVM strategy, the examples are mapped to a higher measurement space by a bit work, and an ideal choice plane is made. There are two points of interest of ESVM: the speculation capacity of the ESVM technique is ideal by augmenting the edge remove and by mapping tests to higher measurement space, it can unravel nonlinear grouping errands.

ESVM method solves binary classification problem by minimizing the function below:

$$\Phi(w,\xi) = \frac{1}{2}(w * w) + C(\sum_{l=1}^{L} \xi_1) \quad (4)$$

With the conditions:

$$\forall \xi_1 \geq 0, w * x_l + b \geq 1 - \xi_1 \text{ if } y_l = +1 \quad (5)$$

And

$$w * x_l + b \leq -1 + \xi_1 \text{ if } y_l = -1$$

Where $w$ is the unknown separating plane, $\xi$ is the soft margin, $x_l$ is the training sample, $y_l$ is the corresponding class of $x_j$, L is the number of training samples, and C is a constant.

By Lagrange method, the minimization problem can be transformed as a task to find the parameter vector $\alpha^0$ maximizing the function

$$W(\alpha) = \sum_{l=1}^{L} \alpha_1 - \frac{1}{2}\sum_{ij}^{L} \alpha_i \alpha_j y_i y_j K(x_i x_j) \quad (10)$$

$$\sum_{l=1}^{L} \alpha_l y_l = 0,\ 0 \leq \alpha_l \leq C$$

Where $x_s$ is a support vector obtained from training, NSV is the number of the support vectors and K(x, xs) is the kernel function. The most frequent and important rules are classified significantly.

## 4. Experimental result

Evaluation of the performance in terms of time complexity, accuracy and rule generation of SVM, ARM and SWT-IARM with ESVM is conducted. To achieve this goal in the experiments both real datasets and synthetic datasets are used.

It tested 35 time series datasets obtained from [18] and [19], consisting of stock market values, chemical and physics phenomenon measurements.

Accuracy

Accuracy is determined as the overall correctness of the model and is computed as the total actual classification parameters ($T_p + T_n$) which is segregated by the sum of the classification parameters ($T_p + T_n + F_p + F_n$). The accuracy is computed as like

$$\text{Accuracy} = \frac{T_p + T_n}{(T_p + T_n + F_p + F_n)} \quad (13)$$

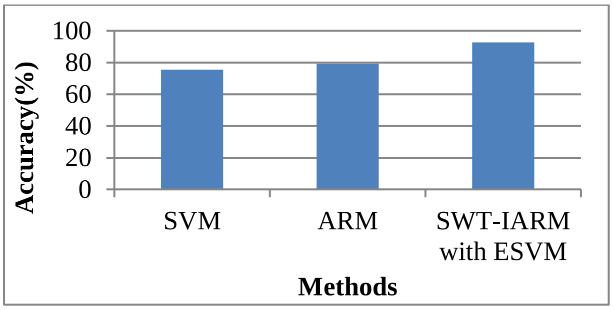


**Fig. 2:** Accuracy Comparison.

From the above Figure 2, it can be observed that the comparison metric is evaluated using existing and proposed method in terms of accuracy. For x-axis the algorithms are taken and in y-axis the accuracy value is plotted. The existing method provides lower accuracy whereas the proposed system provides higher accuracy for the given input. The proposed SWT-IARM with ESVM algorithm selects best rules from the time series data. These rules are applied in ESVM training and testing phase to produce more relevant time series data from given dataset. The result proves that the proposed system attains greater classification results using SWT-IARM with ESVM algorithm. Thus the proposed SWT-IARM with ESVM algorithm is superior to the previous the SVM and ARM algorithms.

Time complexity

The system is better when the algorithm provides lower complexity values

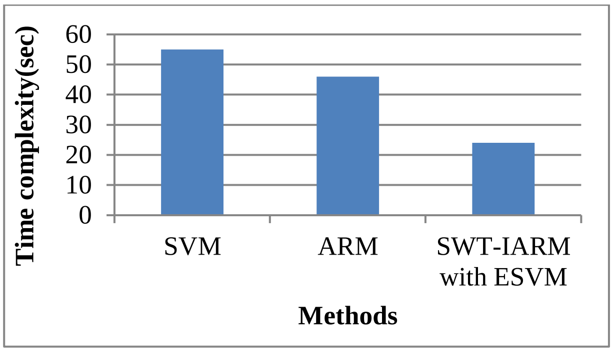


**Fig. 3:** Time Complexity.

From the above Figure 6, it can be observed that the comparison metric is evaluated using existing and proposed method in terms of time complexity. For x-axis the algorithms are taken and in y-axis the time complexity value is plotted. The existing method provides higher time complexity whereas the proposed system provides lower time complexity for the given input. The proposed SWT-IARM with ESVM algorithm selects best rules. These rules are applied in ESVM training and testing phase to produce more relevant data on time series dataset. The result proves that the proposed system attains greater classification results using SWT-IARM with ESVM algorithm. Thus the proposed SWT-IARM with ESVM algorithm is superior to the previous the SVM and ARM algorithms.

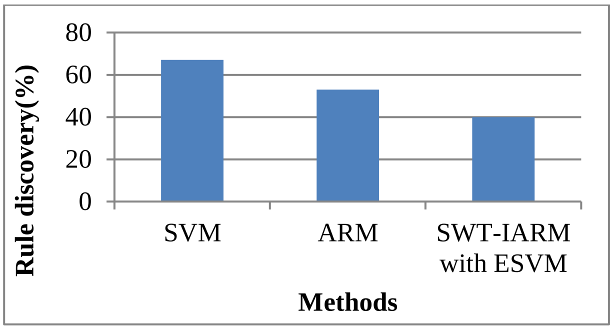


**Fig. 4:** Rule Discovery.

From the above diagram, the rules generated by existing and proposed method are shown. For x-axis the algorithms are taken and in y-axis the rule discovery value is plotted. The proposed SWT-IARM with ESVM algorithm provides lower number of rules and hence it proves that the superior time series classification.

## 5. Conclusion

In this work, the time series dataset is evaluated using efficient techniques. The indexing approach is focused to increase the similarity and faster access.. The time needed to build a data series index becomes prohibitive as the data grows, and may take less time for huge data series using SWT. It used SWT which is used to increase the system efficiency and reduce the space complexity considerably. Then the rule discovery process is done by using IARM method which provides more frequent rules for the time series data. ESVM is applied for obtaining robust classification result on the given time series dataset. In future, can be extend the research to the regions where seasonal variations will be large, since there is a wider scope of significant cost reduction using storage in such areas. The recent hybrid time series prediction algorithms can be proposed to improve the overall performance.

## References


[1] Motoyoshi M, Miura T & Watanabe K, "Mining temporal classes from time series data", *Proceedings of the 11th ACM International Conference on Information and Knowledge Management*, (2002), pp.493–498.

[2] Plant C, Wohlschläger AM & Zherdin A, "Interaction-based clustering of multivariate time series", *Ninth IEEE International Conference on Data Mining*, (2009).

[3] Elfeky MG, Aref WG & Elmagarmid AK, "Periodicity Detection in Time Series Databases", *IEEE Transactions on Knowledge and Data Engineering*, Vol.17, (2005), pp.875–887.

[4] Gorunescu F, *Data Mining Concepts, Models and Techniques*, Springer, Romania, (2010).

[5] Feng X & Huang H, "A Fuzzy Set Based Reconstructed Phase Space Method for Identification of Temporal Patterns in Complex Time Series", *IEEE Transactions on Knowledge and Data Engineering*, Vol.17, (2005), pp.601-613.

[6] Han MK, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, San Francisco, (2001).

[7] Hota HS, Handa R & Shrivas AK, "Time Series Data Prediction Using Sliding Window Based RBF Neural Network", *International Journal of Computational Intelligence Research*, Vol.13, No.5, (2017), pp.1145-1156.

[8] Vafaeipour M, Rahbari O, Rosen MA, Fazelpour F & Ansarirad P, "Application of sliding window technique for prediction of wind velocity time series", *International journal of Energy and environmental engineering (springer)*, Vol.5,(2014), pp.105-111

[9] Yin KC, Hsieh YL & Yang DL, "GLFMiner: global and local frequent pattern mining with temporal intervals", *the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, (2010), pp.2248-2253.

[10] Sengupta N, Aloka S, Narayanaswamy B, Ismail H & Mathew S, "Time series data mining for demand side decision support", *Innovative Smart Grid Technologies-Asia*, (2013), pp.1-6.

[11] Niennattrakul V & Ratanamahatana CA, "On clustering multimedia time series data using k-means and dynamic time warping", *International Conference on Multimedia and Ubiquitous Engineering*, (2007).

[12] Chandrakala S & Chandra Sekhar C, "A density based method for multivariate time series clustering in kernel feature space", *IEEE International Joint Conference on Neural Networks, IEEE World Congress on Computational Intelligence*, (2008).

[13] Viet HH & Anh DT, "M-tree as an index structure for time series data", *International Conference on. Computing, Management and Telecommunications*, (2013).

[14] Winarko E & Roddick JF, "A signature-based indexing method for efficient content-based retrieval of relative temporal patterns", *IEEE Transactions on Knowledge and Data Engineering*, Vol.20, No.6, (2008), pp.825-835.

[15] Mokbel MF, Xiong X, Hammad MA & Aref WG, "Continuous query processing of spatio-temporal data streams in place", *GeoInformatica*, Vol.9, No.4,(2005), pp.343-365.

[16] Miao R & Shen XJ, "Construction of periodic temporal association rules in data mining", *Seventh International Conference on. Fuzzy Systems and Knowledge Discovery (FSKD)*, (2010).

[17] Cao L, "Support vector machines experts for time series forecasting", *Neuro computing*, Vol.51, (2003), pp.321-339.

[18] Wessa. http://www.wessa.net/stocksdata.wasp, 2011.

[19] Data Market. http://datamarket.com/, 2011.