Improving ELM-Based Time Series Classification by Diversified Shapelets Selection

Qifa Sun[1], Qiuyan Yan[1,2(✉)], Xinming Yan[1], Wei Chen[1], and Wenxiang Li[3]

[1] School of Computer Science and Technology, China University of Mining Technology, Xuzhou 221116, China
sunqifa@live.com, {yanqy, yanxm, chenw}@cumt.edu.cn
[2] School of Safty Engineering, China University of Mining Technology, Xuzhou 221116, China
[3] School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
liwx2006@hotmail.com

Abstract. ELM is an efficient neural network which has extremely fast learning capacity and good generalization capability. However, ELM fails to measure up the task of time series classification because it hard to extract the features and characters of time series data. Especially, many time series has trend features which cannot be abstracted by ELM thus lead to accuracy decreasing. Although through selection good features can improve the interpretability and accuracy of ELM, canonical methods either fails to select the most representative and interpretative features, or determine the number of features parameterized. In this paper, we propose a novel method by selection diversified top-k shapelets to improve the interpretability and accuracy of ELM. There are three contributions of this paper: First, we put forward a trend feature symbolization method to extract the trend information of time series; Second, the trend feature symbolic expressions are mapped into a shapelet candidates set and a diversified top-k shapelets selection method, named as **Div-TopkShapelets,** are proposed to find the most k distinguish shapelets; Last, we proposed an iterate ELM method, named as **DivShapELM**, automatically determining the best shapelets number and getting the optimum ELM classifier. The experimental results show that our proposed methods significantly improves the effectiveness and interpretability of ELM.

Keywords: Extreme Learning Machine · Time series classification · Shapelets · Diversified query

1 Introduction

Extreme Learning Machine (ELM for short), based on single-hidden layer feedforward neural networks (SLFNs), was proposed for addressing the slow speed of traditional neural networks. ELM has the extremely fast learning capability and good generalization capability through assign the weights connecting inputs to hidden nodes randomly. The weights between hidden nodes and outputs are learned in a single step,

J.-H. Lee and S. Pack (Eds.): QShine 2016, LNICST 199, pp. 446–456, 2017.
DOI: 10.1007/978-3-319-60717-7_44

which essentially amounts to learning a linear model. In terms of these superiorities, a plethora of methods on ELM optimizing and applying have been designed [1–3].

In spite of so many advantages, there still are some shortcomings of ELM. As we all known, the black-box character of neural networks prevent ELM measuring up to time series data classification by itself. The characters of noisy and high-dimensional increase the complexity and decrease the performance of canonical ELM classifiers. Meanwhile, many time series data has the trend characters, and the problem of trend analysis in time series has attracted significant recently. However, ELM classifiers cannot focus on the trend character, whereas lead to performance decreasing. A possible solution to resolve this issue is to improve the interpretability of ELM by feature selection. A set of good features not only can remove the noises and reduce the dimension of time series, but also can express trend character better.

In this paper, we tackle these issues by a diversified representative and interpretative feature selection–based framework, and the selected feature named shapelets. Shapelet was introduced by Ye and Keogh [4] as a primitive for time series data mining, which is supervised segments of time series that are highly descriptive of the target variable [5]. As a popular data analysis technique, shapelets-transformation classification methods have achieved a high momentum in terms of research focus [6–8] and widely applied in Landcover Classification [9], lung cancer predicting [10], and sensor-based human activity recognition [11].

There are two challenges in applying shapelets extraction method in ELM. First, the original shapelets candidates set are so huge and there are many redundant shapelets which decreasing the accuracy and efficiency of classification. Second, current proposed shapelets selection schemes all fail to extract trend behavior from time series. In order to address these two issues, in this paper, we propose a novel shapelets selection method to adapting ELM to time series data classification. First, a trend feature symbolization method is proposed to extract the trend features in time series. Second, combined with trend symbol expressions, we calculate all the shapelet candidates and get rid of all the similar and redundant shapelets in candidates based on diversity graph. Third, we use the shapelets transformed data to iteratively training ELM and get the optimal classifier. The experimental results show that the proposed approach significantly improves the effectiveness and efficiency of ELM and most time series classifiers.

The rest of this paper is organized as follows. Section 2 brief introduce the conceptions of ELM and shapelets. Section 3 elaborate the proposed method. Experimental analysis is reported in Sect. 4. Finally, Sect. 5 concludes this paper.

2 Related Works

In this section, we will introduce some basic conceptions about ELM and shapelets which are used in this paper.

2.1 ELM

Extreme Learning Machine (ELM) is a generalized single hidden-layer feedforward network. In ELM, the hidden layer node parameter is mathematically calculated instead of being iteratively tuned; thus, it provides good generalization performance at thousands of times faster speed than traditional popular learning algorithms for feedforward neural networks.

Suppose there are N arbitrary distinct training instances $(x_i, t_i) \in R^N \times R^M$, where x_i is one $N \times 1$ input vector and t_i is one $M \times 1$ input vector. If a SLFNs with L hidden nodes can approximate these N samples with zero error, it then implies that there exist β_i, a_i and b_i, such that:

$$f_L(x_j) = \sum_{i=1}^{L} \beta_i G(a_i, b_i, x_j) = t_j,\ j = 1.\ldots..N \tag{1}$$

Where $a_i = [a_{i1}, a_{i2}, \ldots, a_{in}]^{\mathrm{T}}$ and $b_i = [b_{i1}, b_{i2}, \ldots, b_{in}]^{\mathrm{T}}$ denote the weight and bias of the ith hidden layer node, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^{\mathrm{T}}$ is the weight vector connecting the ith hidden node to the output nodes. Then Eq. (1) can be written compactly as:

$$\mathrm{H}\beta = \mathrm{T} \tag{2}$$

Where

$$\mathrm{H}(a_1, \ldots..a_L, b_1, \ldots..b_L, x_1, \ldots\ldots x_N) = \begin{pmatrix} G(a_1, b_1, x_1) & \ldots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{pmatrix}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \ldots \\ \ldots \\ \beta_L^T \end{bmatrix}_{L \times M} \quad \text{and}\ \mathrm{T} = \begin{bmatrix} t_1^T \\ \ldots \\ \ldots \\ t_N^T \end{bmatrix}_{N \times m}$$

Here, $G(a_i, b_i, x_j)$ denotes the activation function which is used to calculate the output of the ith hidden node for the jth training instance. In ELM, many nonlinear activation functions can be used, including sigmoid, sine, hardlimit and radial basis functions. H is called hidden layer output matrix of the network, where with respect to inputs $x_1, x_2 \ldots x_N$ and its jth row represents the output vector of the hidden layer with respect to input x_j.

ELM assigned values to parameters a_i and b_i randomly according to any continuous samplings distribution. Equation (2) then becomes a linear system and the output weight β are estimates as:

$$\widehat{\beta} = \mathrm{H}^{\dagger}\mathrm{T}, \tag{3}$$

where $\mathrm{H}^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix H. $\mathrm{H}^{\dagger} = (\mathrm{H}^{T}\mathrm{H})^{-1}\mathrm{H}^{T}$ if $\mathrm{H}^{T}\mathrm{H}$ is nonsingular or $\mathrm{H}^{\dagger} = (\mathrm{H}^{T}\mathrm{H})^{-1}$ if $\mathrm{H}\mathrm{H}^{T}$ is nonsingular. Here, $\widehat{\beta}$ is the minimum-norm least squares solution of Eq. (2).

2.2 Shapelets

Shapelets are discriminative patterns in time series that best predict the target variable when their distances to the respective time series are used as features for a classifier. The original shapelets-based classifier embeds the shapelets discovery algorithm in a decision tree method, and information gain is adopted to assess the quality of candidate shapelets [4]. Shapelets transformed classification method was proposed to separate the processing of shapelets selection and classification [8]. In this category, distances of time series to shapelets can be viewed as new classification predictors. It has been shown by various researchers that shapelets-derived predictors boost the classification accuracy [6]. In addition, shapelets also provide interpretive features that help domain experts to understand the differences between the target classes.

Research challenges of the shapelets technology include that shapelets candidates selection is time consuming and large quantities of redundant shapelets decreases the accuracy of classification. Some works tried to relieve this issue by introducing clustering [7] or pruning method [12, 13] to reduce the redundancy. But the research issue is still an open question.

3 Proposed Method

In this section, we give the similar shapelets conception, diversified top-k shapelets conception and using diversity graph based on the similar shapelets to get the diversified top-k shapelets results. We detailed four parts of our work (1) a trend feature symbolization algorithm, (2) a mapping algorithm transform the trend symbolic expression to shapelets candidates, (3) a method of extraction diversified top-k shapelets, and (4) transforming the data based on diversified top-k shapelets and classification using ELM. The following contents will discuss above three contribution separately.

3.1 Trend Feature Symbolization

In this section, we proposed a trend feature symbolization method, which express the subsequence as a tuple list, named as TFSAList, and each tuple contains two elements: the gradient k and the feature symbol u. The detailed algorithm of find TFSAList is shown in Algorithm 1.

```
Algorithm 1 CreateTFSAList
Input: Dataset: D
       An instance in D: Tid
       The number of instance in D: N
       The number of subsequence in each Tid: num
       The angle threshold: γ
       Trend feature tuple: TT
Output: TFSAList of D
i=0
For id = 1 to N
   L= Tid.length / num
   S1←Tid [0, L-1]
   K1=gradient (S1)
   i= i+1
   TT [i].U= L
   TT [i].K= K1
   For j = 2 to Tid.length-L+1
     S2=←T [j, j+L-1]
     K2=gradient (S2)
     If (|K1-K2|>γ)
         i=i+1
         TT [i].U = TT [i-1].U+1
         TT [i].K = K2
         j = TT [i].U
         S1←T [j, j+L-1]
         K1 = gradient (S1)
       Else
         K1 = (K1+K2) /2
         TT [i].U = TT [i].U +1
   End For
   TFSAList [id] ←Symbolization (TT)
End For
Return TFSAList
```

3.2 Map Trend Symbolic Expression to Shapelets

In this section, we map the trend symbolic expressions (TFSAList) generated in Algorithm 1 to shapelet candidates. The detailed procedure is shown in Algorithm 2.

```
Algorithm 2 MaptrendToshap
input: trend symbolic expression list TFSAList
       data set: D
output: shapelets candidiates
1: for i=1 to TFSAList. Length
2:   score[i] = GenerateScore (TFSAList[i])
3: end for
4: TFSA_sort=Sort (TFSAList, Score)
5: subsequence_set= Map (TFSA_sort, D)
6: for i=1 to subsequence_set.size
7:    info=CallInfoGain (Subsequence_set[i])
8:    shapeletsCandidates. Add (Subsequence_set[i], Info)
9: end for
```

3.3 Extraction Diversified Top-k Shapelets

Diversified query aims to find the objects which are relevant to query according to scores, however, are not similar to others. Considerable works have focused on the diversify top-k query, but they almost applied on a typical circumstance. In our work, we hope to find a general method to find diversified top-k shapelets, so we got the idea from [11] and acquire diversified top-k shapelets based on diversity graph. In algorithm 3, the diversified top-k shapelets selection method, named as **DivTopkShapelets** is elaborated.

```
Algorithm 3 DivTopkShapelets
input:  shapelets candidates: ShapeletsCandidates
        diversified number: k
output: top k number of diversified shapelets: kShapelets
1: shapGraph=genGraph (ShapeletsCandidates)
2: kShapelets = Ø, n =V(Graph).size
3: kShapelets.add (v1)
4: while (|kShapelets|<k)
5:    for i=2 to n
6:       if (Graph[i] ∩kShapelets = Ø)
7:         kShapelets.add (vi)
8:       end if
9:    end for
10: return kShapelets
```

3.4 Diversified Top-k Shapelets Based ELM Classification

After getting diversified top-k shapelets, we can use these optimal shapelets to transform testing datasets. Actually, each shapelet equal to a special feature, every time series T can be transformed an instance have k number of features by calculating the distance between time series T and every shapelet. Meanwhile, in order to get the best classification accuracy and also to get rid of the independence on the parameter k, we set k in an interval [1,κ] which means the iterate times to model ELM classifier. Then we use the ELM to learning the transformed dataset and evaluate every diversified top-k shapelets candidate. The typical k value with the largest prediction accuracy is selected. In this paper, we refer to the proposed ELM classification method as **DivShapELM**. DivShapELM is a general method to improve the ELM learning accuracy by select diversified and trend features of time series.

4 Experiments

In this section, we study the performance of **DivTopkShapelets** and **DivShapELM** by evaluating its efficiency and effectiveness. The algorithms are coded in C++. All experiments are conducted on a 2.0-GHz HP PC with 1G memory running Window XP and using Weka framework with Java. The UCR time series datasets [15] were used in our experiments.

4.1 Accuracy Comparison

In this section, we verified the accuracy improvement for DivTopkShapelets and DivShapELM separately. Firstly, in Sect. 4.1.1, we select two similar works: ClusterShapelet and ShapeletSelection, to compare with DivTopkShapelets. We hope to find whether DivTopkShapelets can select more representative and representative shapelets than compared methods. Secondly, in Sect. 4.1.2, we compared the accuracy of DivShapELM with state-of-the-art ELM to clarify if our proposed method can improve the effectiveness of ELM.

4.1.1 Accuracy Comparison with ClusterShapelet and ShapeletSelection

In Table 1, we compared the relative accuracy of DivTopkShapelets between ClusterShapelet on six different classifier and on fifteen datasets. The 'average' column means average relative accuracy on six different classifier on one typical dataset. From Table 1 we can draw the conclusion that DivTopkShapelets can enhance the accuracy on all these six classifiers and the most is 10.80% accuracy improved on Naïve Bayes. Especially, DivTopkShapelets overhead ClusterShapelet 30.87% on ECGFiveDays dataset.

In Table 2, we compared the relative accuracy between DivTopkShapelets and ShapeletSelection. We can see that compared with ShapeletSelection, DivTopkShapelets can enhance the accuracy on all these six classifiers and improved the average accuracy on ten datasets. On Adiac dataset, DivTopkShapelets has the best improvement, the accuracy improved 20.80%. For classifiers, DivTopkShapelets

Table 1. Relative accuracy between DivTopkShapelets algorithm and ClusterShapelet algorithm

Data	C4.5	1NN	Naïve Bayes	Bayesian network	Random forest	Rotation forest	Average
Adiac	14.07	17.90	17.90	11.76	21.48	21.48	17.43
Beef	−3.33	−3.33	13.33	−16.67	3.33	−3.33	−1.67
ChlorineConcentra	−1.02	−2.84	−8.46	−1.74	−3.67	−2.06	−3.30
Coffee	10.71	0.00	0.00	3.57	3.57	3.57	3.57%
DiatomSizeReducti	0.98	6.86	−3.59	0.33	−4.58	−6.86	−1.14
ECG200	18.00	4.00	9.00	3.00	12.00	7.00	8.83%
ECGFiveDays	**48.43**	28.11	25.0	24.62	**29.85**	**29.15**	**30.87**
FaceFour	2.93	27.60	29.04	12.20	14.33	18.08	17.36
Gun_Point	2.67	9.33	6.67	8.00	6.00	6.67	6.56
MedicalImages	2.63	1.58	−0.26	−4.74	5.00	1.45	0.94
MoteStrain	3.19	19.49	**31.79**	9.35	6.79	11.42	13.67
RobotSurface	25.62	**37.44**	26.96	**30.12**	20.13	25.12	27.57
SyntheticControl	4.67	5.00	18.33	20.00	7.33	5.67	10.17
Trace	7.00	−2.00	−5.00	0.00	4.00	−2.00	0.33
TwoLeadECG	18.00	8.96	1.23	5.36	10.89	−2.46	6.99
Average improved	10.30	10.54	**10.80**	7.01	9.10	7.53	10.00
Data sets improved	13	12	11	12	13	10	12

enhance the average accuracy most on 1NN classifier, the value improve 6.06% and on the Robot Surface dataset, the DivTopkShapelets enhance 1NN classifier 32.61% on accuracy.

4.1.2 Accuracy Comparison Between DivShapELM and ELM

As shown in Table 3, DivShapELM has obvious advantages than state-of-art ELM on 12 out of 15 datasets. Especially, On ECGFiveDays Dataset, DivShapELM has the accuracy of 90.57%, better then ELM 32.7%. So by introduction the conception of diversified top-k shapelets, we can get the most representative attributes of dataset and also can get rid of the redundant, which can obviously improve the accuracy of ELM and also can enhance the explainable of selected features.

4.2 Time Cost Comparison

DivShapELM has three extra pre-procedures: shapelets candidate selection, diversified shapelets selection time and data transform time. Once the transformed data got, the rest procedure is a usual classification process. Table 4 give the extra time and classification time of DivShapELM and ELM. The time cost of diversified shapelets selection are varied with dataset, but which can be conducted in an offline manner. Apparently, DivShapELM has the less classification time than ELM apparently on 13 out of 15 datasets.

Table 2. Relative accuracy between DivTopkShapelets and ShapeletSelection

Data	C4.5	1NN	Naïve Bayes	Bayesian network	Random forest	Rotation forest	Average
Adiac	16.62	21.23	**17.14**	**16.37**	**26.09**	**23.02**	**20.08**
Beef	−3.33	3.33	16.67	0.00	3.33	10.00	5.00
ChlorineConcentration	0.10	14.24	−10.70	−0.65	13.54	0.78	2.89
Coffee	−3.57	7.14	0.00	0.00	−7.14	7.14	0.59
DiatomSizeReduc	5.56	9.15	−6.86	−4.90	−16.34	−5.89	−3.21
ECG200	4.00	−6.00	2.00	3.00	5.00	0.00	1.33
ECGFiveDays	−0.46	−0.23	−1.28	−0.58	−0.35	0.70	−0.37
FaceFour	3.41	0.00	1.13	−7.95	−1.13	9.09	0.76
Gun_Point	0.67	−2.00	−1.34	−0.67	−3.33	−0.67	−1.22
MedicalImages	1.84	8.55	0.53	−4.08	15.00	5.92	4.63
MoteStrain	−4.63	1.77	−5.11	−5.59	−6.15	−1.03	−3.46
RobotSurface	**16.97**	**32.61**	5.99	8.49	12.64	11.32	14.67
SyntheticControl	3.00	1.67	−0.67	0.67	−1.00	0.00	0.61
Trace	6.00	0.00	−4.00	2.00	0.00	−2.00	0.33
TwoLeadECG	−3.07	−0.52	0.88	−0.35	−6.05	−2.46	−1.93
Average improved	2.87	**6.06**	0.96	0.38	2.27	3.73	2.71
Data sets improved	10	11	8	7	7	10	10

Table 3. Accuracy comparison between DivShapELM and ELM

Data	DivShapELM	ELM
Adiac	**58.70**	41.05
Beef	**67.88**	55.42
ChlorineConcentration	50.30	**59.77**
Coffee	**95.44**	90.60
DiatomSizeReduction	82.76	**88.95**
ECG200	**90.12**	67.15
ECGFiveDays	**90.57**	57.87
FaceFour	**87.34**	80.12
Gun_Point	**95.43**	88.65
MedicalImages	51.23	**62.00**
MoteStrain	**84.70**	66.19
SonyAIBORobotSurface	**94.16**	74.55
SyntheticControl	**97.63**	85.40
Trace	**93.17**	90.77
TwoLeadECG	**94.23**	78.12

Table 4. Run time of DivShapELM and ELM

Data	Candidate selection (s)	Diversified shapelets selection (s)	Data transform (s)	DivShapELM (s)	ELM (s)
Adiac	1277	2.09	0.811	**0.04992**	0.13728
Beef	1026	251.894	0.702	0.0156	0.0156
ChlorineConcentration	2636	3.136	7.363	**0.01716**	0.05148
Coffee	337	234.422	0.436	0.04368	0.00936
DiatomSizeReduction	95	476.331	2.294	**0.02652**	0.0312
ECG200	216	0.905	0.14	**0.0156**	0.02496
ECGFiveDays	84	0.858	0.717	**0.02184**	0.02652
FaceFour	634	66.94	0.671	**0.02184**	0.06864
Gun_Point	151	3.167	0.218	**0.0312**	0.09048
MedicalImages	732	0.219	0.514	**0.00468**	0.05304
MoteStrain	30	0.687	0.483	**0.03276**	0.04368
SonyAIBORobotSurface	28	0.39	0.249	0.01404	0.00624
SyntheticControl	340	0.343	0.25	**0.04368**	0.07488
Trace	1168	268.165	1.201	**0.01716**	0.1638
TwoLeadECG	25	0.265	0.546	**0.039**	0.1017

5 Conclusion

In order to improve the interpretable and representable of ELM, we proposed a novel method to select distinct and optimal features through selecting k number shapelets of time series dataset, meanwhile, the selected shapelets can express the trend information as well. We verified the effectiveness and efficiency of ELM and other 6 classifies on 15 datasets. The experiments results that DivTopkShapelets can improve the effectiveness and efficiency on almost all of the classifiers on almost datasets. The DivShapELM method also has outstanding accuracy than state-of-the-art ELM method.

Acknowledgement. Supported by the Natural Science Foundation of Jiangsu Province of China (BK20140192). National Natural Science Foundation and Shanxi Provincial People's Government Jointly Funded Project of China for Coal Base and Low Carbon (No. U1510115), the Qing Lan Project, the China Postdoctoral Science Foundation (No. 2013T60574).

References

1. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K.: Extreme learning machine: a new learning scheme of feedforward neural networks. In: Proceedings of International Joint Conference on Neural Networks (IJCNN 2004), pp. 985–990 (2004)
2. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K.: Extreme learning machine: theory and applications. Neurocomputing **70**, 489–501 (2004)

3. Huang, G.-B., Zhu, Q.-Y., Mao, K.Z., Siew, C.-K., Saratchandran, P., Sundararajan, N.: Can threshold networks be trained directly. IEEE Trans. Circuits Syst. II **53**(3), 187–191 (2006)
4. Ye. L., Keogh, E.: Time series shapelets: a new primitive for data mining. In: Proceedings of 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 947–956. ACM (2009)
5. Ye, L., Keogh, E.: Time series shapelets: a novel technique that allows accurate, interpretable and fast classification. Data Min. Knowl. Disc. **22**(1–2), 149–182 (2011)
6. Mueen, A., Keogh, E., Young, N.: Logical-shapelets: an expressive primitive for time series classification. In: Proceedings of 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1154–1162. ACM (2011)
7. Rakthanmanon, T., Keogh, E.: Fast shapelets: a scalable algorithm for discovering time series shapelets. In: Proceedings of 13th SIAM Conference on Data Mining (SDM) (2013)
8. Lines, J., Davis, L.M., Hills, J., et al.: A shapelet transform for time series classification. In: Proceedings of 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 289–297. ACM (2012)
9. Hills, J., Lines, J., Baranauskas, E., et al.: Classification of time series by shapelet transformation. Data Min. Knowl. Disc. **28**(4), 851–881 (2014)
10. Zakaria, J., Mueen, A., Keogh, E.: Clustering time series using unsupervised-shapelets. In: 2012 IEEE 12th International Conference on Data Mining (ICDM), pp. 785–794. IEEE (2012)
11. Xing, Z., Pei, J., Philip, S.Y., et al.: Extracting interpretable features for early classification on time series. In: SDM, vol. 11, pp. 247–258 (2011)
12. Chang, K.W., Deka, B., Hwu, W.M.W., et al.: Efficient pattern-based time series classification on GPU. In: 2012 IEEE 12th International Conference on Data Mining (ICDM), pp. 131–140. IEEE (2012)
13. Yuan, J.D., Wang, Z.H., Han, M.: Shapelet pruning and shapelet coverage for time series classification. J. Softw. **26**(9), 2311–2325 (2015). (in Chinese)
14. Qin, L., Yu, J.X., Chang, L.: Diversifying top-k results. Proc. VLDB Endow. **5**(11), 1124–1135 (2012)
15. Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., Batista, G.: The UCR Time Series Classification Archive. [DB/OL] (2015). http://www.cs.ucr.edu/~eamonn/time_series_data/