Ordinal versus nominal time series classification

David Guijo-Rubio[1] [0000−0002−8035−4057],
Pedro Antonio Gutiérrez[1] [0000−0002−2657−776X],
Anthony Bagnall[2] [0000−0003−2360−8994], and
César Hervás-Martínez[1] [0000−0003−4564−1816]

[1]Department of Computer Sciences, Universidad de Córdoba, Córdoba, Spain
{dguijo,pagutierrez,chervas}@uco.es
[2]School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich, United Kingdom.
ajb@uea.ac.uk

Abstract. Time series ordinal classification is one of the less studied problems in time series data mining. This problem consists in classifying time series with labels that show a natural order between them. In this paper, an approach is proposed based on the Shapelet Transform (ST) specifically adapted to ordinal classification. ST consists of two different steps: 1) the shapelet extraction procedure and its evaluation; and 2) the classifier learning using the transformed dataset. In this way, regarding the first step, 3 ordinal shapelet quality measures are proposed to assess the shapelets extracted, and, for the second step, an ordinal classifier is applied once the transformed dataset has been constructed. An empirical evaluation is carried out, considering 7 ordinal datasets from the UEA & UCR Time Series Classification (TSC) repository. The results show that a support vector ordinal classifier applied to the ST using the Pearson's correlation coefficient (R^2) is the combination achieving the best results in terms of two evaluation metrics: accuracy and average mean absolute error. A final comparison against three of the most popular and competitive nominal TSC techniques is performed, demonstrating that ordinal approaches can achieve higher performances even in terms of accuracy.

Keywords: Time Series · Ordinal Classification · Ordinal regression · Shapelet Quality Measures

1 Introduction

Time Series Ordinal Classification (TSOC) refers to a prediction problem where the objective is to classify time series with an ordinal label, i.e. the set of labels includes a natural order relationship. In this context, ordinal classification [12] covers those supervised problems where the target variable is discrete and includes a natural order relationship among the labels. Ordinal classification problems can be found in several fields, such as meteorological prediction [10, 11], or medical research [19], among others.

On the other hand, time series consists of data points collected chronologically. In the last years, a countless number of novel approaches in the nominal

Time Series Classification (TSC) field have been presented. According to [2], TSC has been tackled from several points of view, depending on the discriminatory features the approach is trying to find. One of these techniques are shapelets [25], phase independent subsequences of the original time series able to differentiate between classes, i. e. a class can be distinguished depending on whether the shapelets could be found in the original time series or not. Further research was done by Hills *et al.* in [13], where the Shapelet Transform (ST) was firstly proposed, in which the best k shapelets (ordered by using a shapelet quality measure) are used to build a transformed dataset in which the attributes are the distances between the shapelets and the original time series. After that, an effective classifier can be applied to the transformed dataset.

Focusing on the proposal of Hills *et al.* in [13], the ST pipeline can be divided into two main steps: 1) the shapelet extraction procedure and 2) the classifier learning using the transformed dataset as input. Regarding the first step, the best k shapelets are selected by using a shapelet quality measure. In order to adapt this approach to the ordinal setting, in this paper, we propose 3 different metrics to measure the ordinal quality of the shapelets, and we compare them against the state-of-the-art Information Gain metric. The second step is adapted by considering an ordinal classifier, instead of using a nominal one, with the objective of exploiting the natural order relationship of the labels. We compare the results obtained against two nominal state-of-the-art techniques.

In this way, the main objectives of this paper are to establish a baseline for TSOC using ST and to demonstrate that, for those ordinal datasets included in the most popular TSC repository, ordinal approaches are able to achieve better performance than standard TSC techniques in terms of accuracy.

2 Background

Time series is a series of data points arranged in time, i.e. the values of the time series are chronological. In a more formal way, the i-th time series object is defined as $\mathbf{T}_i = \{t_{i1}, t_{i2}, \ldots, t_{in}\}$, where n is the length of the time series (note that we only consider equal-length time series). Therefore, a time series dataset is composed of N time series, being defined as $\mathbf{T} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_N\}$.

In this paper, we are considering ordinal TSC problems: each time series is associated with a label $\mathcal{C}_i \in Y$, where the set of ordinal labels is $Y = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_Q\}$, including $Q > 2$ categories. An order relationship between the labels is found in the problem, i.e. the constraint $C_1 \prec C_2 \prec \ldots \prec C_Q$ should be satisfied.

2.1 Time series shapelets

TSC is a very popular field of research in time series data mining [2]. One of the most recent approaches in this field consists in an ensemble including several modules, each one based on a different transformation applied to the original time series dataset, prior to the classification step. One of the first proposals was the shapelet module. A shapelet [25] is a phase independent subsequence of the original time series. The original approach finds all possible shapelets through an enumerative search, which is particularly slow, and then embeds the shapelets in a decision tree without a significant improvement in performance. From this starting point, several approaches have been published in the literature, including [3, 9, 13], among others. In this paper, we focus on the ST [13], a two-phase approach that uses the extracted shapelets to transform the original dataset, in which the transformed attributes represent the similarity in shape between the original time series and the shapelets obtained, and then applies a classifier to the transformed dataset.

More formally, a shapelet $\mathbf{s}_j = \{s_1, s_2, \dots, s_l\}$ is a subsequence of a time series $\mathbf{T}_j$, where $l \leq n$ and the subscript j is used to explicitly show that the shapelet $\mathbf{s}$ is a subsequence of time series $\mathbf{T}_j$. The main pipeline for the shapelet extraction procedure consists of three parts [13]: first of all, a set of candidates is randomly generated satisfying several constraints, then, the distance between each shapelet and the original time series is computed to, finally, measure the shapelet quality. The last version of ST [3] proposes new constraints for the shapelet extraction, such as balancing the number of shapelets extracted per class. Moreover, the Euclidean distance is used to measure the similarity between the set of shapelets and the original time series; this distance is computed as the minimum of the distances between the shapelet and all the subsequences with the same length of the shapelet. Finally, the Information Gain (IG) [22] is used to assess the shapelet quality and retain those with higher IG. The formulation is detailed in [13].

In order to consider the natural order between the labels, we propose to consider three different shapelet quality measures. The main idea is to extract shapelets able to reduce the misclassification errors involving more jumps in the ordinal scale:

- Ordinal Fisher (OF) score [20] is an ordinal adaptation of the Fisher score [7]. This measure gives higher penalisation for distant classes in the ordinal scale, therefore, distant classes should be associated with higher distances. It is defined as:

$$OF(\mathbf{s}_j) = \frac{\sum_{k=1}^{Q} \sum_{j=1}^{Q} |\mathcal{O}(\mathcal{C}_k) - \mathcal{O}(\mathcal{C}_j)|(\bar{x}_k - \bar{x}_j)^2}{(Q-1)\sum_{k=1}^{Q}(S_k)^2}, \tag{1}$$

  where $\mathcal{O}(\mathcal{C}_q)$ is the position of the category C_q in the ordinal scale, i.e. $\mathcal{O}(\mathcal{C}_q) = q$, and $\bar{x}_k$ and S_k are the mean and standard deviation of the distances according to the evaluated shapelet $\mathbf{s}_j$ when considering only the time series of the class $\mathcal{C}_k$.
- The Pearson's correlation coefficient (R^2) calculates the correlation between $d_{\mathbf{s}_j,\mathbf{T}_i}$ and c_{y_j,y_i}, $i \in \{1, \dots, N\}$, where $d_{\mathbf{s}_j,\mathbf{T}_i}$ are the distances from the shapelet $\mathbf{s}_j$ to the original times series $\mathbf{T}_i$, and c_{y_j,y_i} are the differences of the corresponding class values $c_{y_j,y_i} = |\mathcal{O}(\mathcal{C}_j) - \mathcal{O}(\mathcal{C}_i)|$, where y_j is the class of $\mathbf{T}_j$ (the time series from which the shapelet $\mathbf{s}_j$ is extracted) and y_i is the

class of $\mathbf{T}_i$. In this way, R^2 is defined as:

$$R^2(\mathbf{s}) = \sum_{i=1}^{N} \frac{S(d_{\mathbf{s}_j,\mathbf{T}_i}, c_{\mathbf{s}_j,\mathbf{T}_i})}{S_{d_{\mathbf{s}_j,\mathbf{T}_i}} S_{c_{\mathbf{s}_j,\mathbf{T}_i}}}, \tag{2}$$

where $S(\cdot)$ is the covariance of two variables.

- Similarly, the Spearman's correlation coefficient (ρ) computes the correlation between two categorical or continuous variables, following the idea presented for the R^2 quality measure. Therefore, ρ is defined as:

$$\rho(\mathbf{s}) = 1 - \frac{6\sum_{i=1}^{N}(\mathcal{R}(d_{\mathbf{s}_j,\mathbf{T}_i}) - \mathcal{R}(c_{\mathbf{s}_j,\mathbf{T}_i}))^2}{N(N^2-1)}, \tag{3}$$

where $\mathcal{R}(x)$ is the rank of x in the set of all values obtained.

2.2 Ordinal classification

Once the transformed dataset is constructed (each new attribute j represents the distance between time series i and shapelet j), a classifier is applied to it. One of the main objectives of this paper is to demonstrate that ordinal classifiers can lead to a better performance than nominal ones, given their ability to consider the natural order between the labels. In this way, three different support vector machine techniques have been chosen, using the `ORCA` framework [21][1]:

- In order to perform comparisons, we first consider nominal Support Vector Classifier (SVC) [14] with two options: one versus one formulation (SVC1V1) and one versus all paradigm (SVC1VA). These two nominal classifiers are very popular in the state-of-the-art, given their accuracy for both binary and nominal multiclass problems.
- On the other hand, an ordinal technique is considered: the Support Vector Ordinal Regression (SVOR) [23] methodology, which is the adaptation of SVC to ordinal classification. Specifically, in this paper we have chosen the SVOR version considering IMplicit constrains (SVORIM) [4]. This approach computes the discriminant parallel hyperplanes for the data and a set of thresholds by imposing implicit constraints in the optimization problem.

In order to assess the performance of ordinal classification problems, there are several metrics that can be considered [5]. In this paper, apart from the accuracy, which is the standard evaluation metric for nominal classification, a specific ordinal evaluation metric should be considered to avoid ignoring order information. In this sense, the misclassification errors are not equally penalised, giving more cost to those misclassified patterns in farther classes. Therefore, we have considered the Correct Classification Rate (CCR), also known as accuracy, which is the global performance of a classifier and the Average Mean Absolute Error ($AMAE$) [1], which measures the ordinal classification errors made for each class.

[1] `ORCA` is available in the repository `https://github.com/ayrna/orca`.

3 Experimental results and discussion

This section exposes the ordinal time series datasets considered, as well as the experimental settings used. Moreover, the results achieved for the three classifiers applied to the four versions of ST using different shapelet quality measures are also shown, along with a comparison of the best ordinal ST approach to the main state-of-the-art algorithms in nominal TSC[2].

3.1 TSOC datasets

Table 1 shows 7 datasets appropriately selected from the popular UEA & UCR TSC repository[3], given their ordinal nature. All of them belong to the field of bone age estimation [6], except the *EthanolLevel* dataset, which is obtained from the detection of forget spirits using non-intrusive methods [15].

Apart from the main information of the datasets, the Imbalance Ratio (IR) [18] is also included in Table 1. This feature shows whether the distribution of patterns in the datasets is imbalanced, i. e. most of the patterns belongs to a given class (high values for IR). In these cases, trivial classifiers can achieve high values of accuracy.

Table 1. Characteristics of the datasets used in the experiments.

Dataset	#Classes (Q)	#Train	#Test	Length	%IR
DistalPhalanxOutlineAgeGroup	3	400	139	80	1.532
DistalPhalanxTW	6	400	139	80	1.577
EthanolLevel	4	504	500	1751	0.750
MiddlePhalanxOutlineAgeGroup	3	400	154	80	0.881
MiddlePhalanxTW	6	399	154	80	1.276
ProximalPhalanxOutlineAgeGroup	3	400	205	80	0.951
ProximalPhalanxTW	6	400	205	80	2.203

3.2 Experimental settings

The ST algorithm has been run for one hour during shapelet search. This algorithm has been run with the default values. In the case of ST using IG as shapelet quality measure, an inferior limit of $IG = 0.05$ is used to discard very low-quality shapelets. Furthermore, aiming to reproduce the same behaviour for the remaining shapelet quality measures, the lowest-quality 10% shapelets are also discarded.

Regarding the datasets, the standard train and test data splits given in the UEA & UCR TSC repository are used. Moreover, it is worthy of mention that

[2] All the code used in this paper is available in the repository `https://github.com/dguijo/TSOC`.
[3] `http://www.timeseriesclassification.com/`

the models are adjusted using only the training set, whereas the test set is only used to evaluate the learned models.

With respect the classifiers, they have been run once, given their deterministic nature. Moreover, their sensitive hyper-parameters have been adjusted using a nested 10-fold cross-validation approach, considering $AMAE$ as the parameter selection criteria, due to the fact that CCR ignores ordinal information. Given that the three classifiers are SVM-based, the same range of values $\{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$ has been used to adjust both the cost parameter and the kernel width.

Finally, the main code for the ST and for the IG shapelet quality measure was obtained from `sktime` toolkit [17][4].

3.3 Results

In Table 2, the results achieved for the four versions of the ST using different shapelet quality measures are shown. Concretely, the performances of the three classifiers applied to the transforms are presented in terms of CCR and $AMAE$. Furthermore, in order to compare the results in a more global way, we have included the average ranking and the number of datasets in which the respective shapelet quality measure is able to reach to the best performance (#Wins).

As can be seen in Table 2, the ST using R^2 as shapelet quality measure is the one achieving the best results for most of the datasets and classifiers. Specifically, in terms of CCR, the R^2 measure obtains an average ranking of 1.95, followed by ρ (2.36). Regarding number of wins, the ST combined with R^2 reaches to the best results in 11 cases, whereas ST using either ρ or IG ties in 8 cases. On the other hand, in terms of $AMAE$, the ST combined with R^2 also achieves the best results, achieving an average ranking of 1.74 with 11 wins, whereas the second best approach is the standard ST using the IG as shapelet quality measure, with an average ranking of 2.38 and 8 wins. Therefore, it is clear that ST using R^2 as shapelet quality measure achieves the best results without much dependence on the classifier used.

3.4 Comparison against the state-of-the-art algorithms in TSC

In order to establish a comparison against the main state-of-the-art algorithms in TSC, the following three algorithms have been used (which achive the best results up-to-the-knowledge of the authors):

- The Hierarchical Vote Collective of Transformation-based Ensembles (HIVE-COTE) [16] is a meta-ensemble composed of five different modules with several algorithms in each one. These modules rely on the idea of transforming the original dataset prior to classification, such as ST, among others.

[4] `sktime` is available in the repository `https://github.com/alan-turing-institute/sktime`.

Table 2. CCR and $AMAE$ results achieved by the four ST methods (OF, ρ and R^2 are the proposals in this paper).

		CCR				$AMAE$			
Classifier	Dataset	IG	OF	ρ	R^2	IG	OF	ρ	R^2
SVORIM	DistalPhalanxOutline	**75.54**	74.82	**75.54**	**75.54**	**0.2277**	0.2665	0.2443	**0.2277**
	DistalPhalanxTW	68.35	*69.06*	65.47	**69.78**	**0.4671**	0.5045	0.5264	*0.4822*
	EthanolLevel	**71.40**	46.00	62.00	*62.40*	**0.2938**	0.6067	0.3988	*0.3973*
	MiddlePhalanxOutline	62.99	62.99	**63.64**	**63.64**	**0.5484**	*0.5521*	0.5791	0.5676
	MiddlePhalanxTW	**56.49**	54.55	53.90	**56.49**	1.0137	1.0308	*1.0039*	**0.9851**
	ProximalPhalanxOutline	*86.34*	84.88	*86.34*	**87.32**	*0.1824*	0.2254	0.1978	**0.1744**
	ProximalPhalanxTW	74.63	*76.10*	**79.02**	*76.10*	0.5371	0.4989	*0.4521*	**0.4198**
SVC1V1	DistalPhalanxOutline	**75.54**	74.82	**75.54**	74.82	**0.2277**	0.2334	**0.2277**	0.2334
	DistalPhalanxTW	*69.06*	*69.06*	66.91	**70.50**	0.5600	*0.5046*	0.5440	**0.4614**
	EthanolLevel	**69.00**	48.80	58.20	*61.00*	**0.3301**	0.6795	0.4632	*0.4294*
	MiddlePhalanxOutline	61.04	61.04	*61.69*	**62.34**	0.5827	0.5775	*0.5737*	**0.5636**
	MiddlePhalanxTW	**59.09**	56.49	**59.09**	**59.09**	*0.8785*	0.8962	0.8963	**0.8541**
	ProximalPhalanxOutline	*85.85*	**86.34**	*85.85*	*85.85*	*0.1858*	**0.1820**	0.2016	*0.1858*
	ProximalPhalanxTW	76.59	*78.54*	**80.98**	72.68	0.5104	0.4836	**0.4536**	*0.4569*
SVC1VA	DistalPhalanxOutline	**75.54**	74.10	*74.82*	74.10	**0.2277**	0.2572	*0.2546*	0.2778
	DistalPhalanxTW	66.19	*68.35*	67.63	**69.06**	0.5972	*0.5158*	0.5702	**0.4893**
	EthanolLevel	**67.60**	47.20	56.40	*58.80*	**0.3444**	0.7630	0.5178	*0.4794*
	MiddlePhalanxOutline	*62.34*	61.04	*62.34*	**63.64**	*0.5636*	0.5723	0.5699	**0.5561**
	MiddlePhalanxTW	55.19	49.35	**57.79**	*56.49*	1.0677	1.1290	*0.9689*	**0.9541**
	ProximalPhalanxOutline	*85.85*	85.37	*85.85*	**86.34**	*0.1858*	0.1896	*0.1858*	**0.1820**
	ProximalPhalanxTW	75.61	*76.59*	**80.98**	*76.59*	0.5362	0.4852	**0.3825**	*0.4446*
	Average ranking	2.48	3.22	*2.36*	**1.95**	*2.38*	3.24	2.64	**1.74**
	#Wins	*8*	1	*8*	**11**	*8*	1	3	**11**

- InceptionTime [8] is an ensemble of deep Convolutional Neural Networks (CNN) models, inspired by the Inception-v4 architecture. In this model, several filters of different lengths are applied simultaneously to an input time series.
- Time Series Combination of Heterogeneous and Integrated Embedding Forest (TS-CHIEF) [24] is an ensemble classifier integrating the most effective embeddings of time series, using tree-structured classifiers.

All these three ensembles are highly competitive in terms of accuracy, although HIVE-COTE is the one achieving the best performance in terms of CCR. However, the main advantages of InceptionTime and TS-CHIEF are their scalability and efficiency.

Table 3 shows the comparison carried out in terms of CCR, given that it is the main goal of TSC. Specifically, the results shown for the ST are those in which the Pearson's correlation coefficient (R^2) is used as the shapelet quality measure, considering different classifiers applied to the transformed dataset: SVC1V1, SVC1VA and SVORIM. Moreover, the results shown for the InceptionTime and TS-CHIEF algorithms are those presented in the original papers (though we have run the TS-CHIEF algorithm for the *EthanolLevel* dataset, given that it is included in the cited work). For HIVE-COTE, they were obtained using the last version of the algorithm, because it has been recently improved.

Table 3. Comparison in terms of CCR of different classifiers applied to the ST using R^2 as quality measure against the state-of-the-art algorithms in TSC.

	SVC1V1	SVC1VA	SVORIM	HIVE-COTE	InceptionTime	TS-CHIEF
DistalPhalanxOutline	*74.82*	74.10	**75.54**	**75.54**	73.38	74.10
DistalPhalanxTW	**70.50**	69.06	*69.78*	67.63	68.35	68.35
EthanolLevel	61.00	58.80	62.40	*71.40*	**81.40**	52.80
MiddlePhalanxOutline	*62.34*	**63.64**	**63.64**	59.09	55.19	59.09
MiddlePhalanxTW	**59.09**	*56.49*	*56.49*	55.84	51.30	55.85
ProximalPhalanxOutline	85.85	*86.34*	**87.32**	84.39	84.88	84.88
ProximalPhalanxTW	72.68	76.59	76.10	*80.00*	77.56	**81.46**
Average ranking	*3.00*	3.21	**2.36**	3.79	4.43	4.21
#Wins	*2*	1	**3**	1	1	1

As can be seen in Table 3, SVORIM achieves the best results or the second best in most of the datasets, with 3 wins and an average ranking of 2.36, considerably better than the rest of approaches. SVC1V1 is the second one with an average ranking of 3.00 and 2 wins. The remaining techniques only have 1 win and their average rankings are much worse. Furthermore, all the classifiers applied to the ST combined with R^2 shapelet quality measure (SVC1V1, SVC1VA and SVORIM) achieve a higher performance than state-of-the-art TSC methods.

Some facts must be outlined from the results shown in Table 3: 1) The combination of the ordinal classifier SVORIM with ST R^2 quality measure obtains the best performance in terms of CCR. 2) Nominal classifiers, SVC1V1 and SVC1VA, are taking advantage of the ordinal information induced by the ST combined with R^2 and also obtain competitive results, better than those of the ensemble approaches. 3) HIVE-COTE and TS-CHIEF results are very similar for almost all the datasets, being HIVE-COTE slightly better. 4) InceptionTime is the algorithm obtaining the worse results, because the datasets include short time series. The only exception is *EthanolLevel*, with length equal to 1751, for which InceptionTime is the one obtaining the best performance.

4 Conclusions

This paper presents a novel approach to Time Series Ordinal Classification using the Shapelet Transform (ST). To obtain the k best shapelets for the ST, 3 different ordinal shapelet quality measures are proposed, exploiting the order of labels: Ordinal Fisher (OF), Pearson's correlation coefficient (R^2) and Spearman's correlation coefficient (ρ). These approaches are then compared against Information Gain (IG), which is the one used by the standard ST.

On the other hand, regarding the second step of ST in which a classifier is applied to the transformed data, this paper proposes the use of an ordinal support vector classifier, which is compared against the corresponding nominal versions.

Finally, a comparison against some of the best state-of-the-art techniques in TSC is included: HIVE-COTE, TS-CHIEF and InceptionTime. In this way, the best ordinal approach presented in this paper (ST using R^2 as shapelet quality measure combined with the support vector ordinal classifier) is able to obtain a better accuracy rank than the alternative nominal TSC techniques.

Possible lines of future research are to include the ordinal information of the labels in other points of the ST process and to adapt other modules of the HIVE-COTE meta-ensemble to ordinal classification.

Acknowledgement. This research has been partially supported by the “Ministerio de Economía, Industria y Competitividad” (Ref. TIN2017-85887-C2-1-P) and the “Fondo Europeo de Desarrollo Regional (FEDER) y de la Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía” (Ref. UCO-1261651), Spain. D. Guijo-Rubio’s research has been supported by the FPU Predoctoral Program from Spanish Ministry of Education and Science (Grant Ref. FPU16/02128).

References

1. Baccianella, S., Esuli, A., Sebastiani, F.: Evaluation measures for ordinal regression. In: Intelligent Systems Design and Applications, 2009. ISDA’09. Ninth International Conference on. pp. 283–287. IEEE (2009)
2. Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E.: The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Mining and Knowledge Discovery **31**(3), 606–660 (2017)
3. Bostrom, A., Bagnall, A.: Binary shapelet transform for multiclass time series classification. Transactions on Large-Scale Data and Knowledge Centered Systems **32**, 24–46 (2017)
4. Chu, W., Keerthi, S.S.: Support vector ordinal regression. Neural Computation **19**(3), 792–815 (2007)
5. Cruz-Ramírez, M., Hervás-Martínez, C., Sánchez-Monedero, J., Gutiérrez, P.A.: Metrics to guide a multi-objective evolutionary algorithm for ordinal classification. Neurocomputing **135**, 21–31 (2014)
6. Davis, L.M., Theobald, B.J., Lines, J., Toms, A., Bagnall, A.: On the segmentation and classification of hand radiographs. International journal of neural systems **22**(05), 1250020 (2012)
7. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern classification. John Wiley & Sons (2012)
8. Fawaz, H.I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D.F., Weber, J., Webb, G.I., Idoumghar, L., Muller, P.A., Petitjean, F.: Inceptiontime: Finding alexnet for time series classification. ArXiv e-prints **arXiv:1909.04939** (2019), `http://arxiv.org/abs/1909.04939`
9. Grabocka, J., Schilling, N., Wistuba, M., Schmidt-Thieme, L.: Learning time-series shapelets. In: Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2014)

10. Guijo-Rubio, D., Casanova-Mateo, C., Sanz-Justo, J., Gutiérrez, P., Cornejo-Bueno, S., Hervás, C., Salcedo-Sanz, S.: Ordinal regression algorithms for the analysis of convective situations over madrid-barajas airport. Atmospheric Research **236**, 104798 (2020)
11. Guijo-Rubio, D., Gutiérrez, P., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., Hervás-Martínez, C.: Prediction of low-visibility events due to fog using ordinal classification. Atmospheric Research **214**, 64–73 (2018)
12. Gutiérrez, P.A., Pérez-Ortiz, M., Sánchez-Monedero, J., Fernández-Navarro, F., Hervás-Martínez, C.: Ordinal regression methods: survey and experimental study. IEEE Transactions on Knowledge and Data Engineering **28**(1), 127–146 (2016)
13. Hills, J., Lines, J., Baranauskas, E., Mapp, J., Bagnall, A.: Classification of time series by shapelet transformation. Data Mining and Knowledge Discovery **28**(4), 851–881 (2014)
14. Hsu, C.W., Lin, C.J.: A comparison of methods for multiclass support vector machines. IEEE transactions on Neural Networks **13**(2), 415–425 (2002)
15. Large, J., Kemsley, E.K., Wellner, N., Goodall, I., Bagnall, A.: Detecting forged alcohol non-invasively through vibrational spectroscopy and machine learning. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. pp. 298–309. Springer (2018)
16. Lines, J., Taylor, S., Bagnall, A.: Time series classification with HIVE-COTE: The hierarchical vote collective of transformation-based ensembles. ACM Trans. Knowledge Discovery from Data **12**(5) (2018)
17. Löning, M., Bagnall, A., Ganesh, S., Kazakov, V., Lines, J., Király, F.J.: sktime: A Unified Interface for Machine Learning with Time Series. In: Workshop on Systems for ML at NeurIPS 2019
18. Pérez-Ortiz, M., Gutiérrez, P.A., Hervás-Martínez, C., Yao, X.: Graph-based approaches for over-sampling in the context of ordinal regression. IEEE Transactions on Knowledge and Data Engineering **27**(5), 1233–1245 (2014)
19. Pérez-Ortiz, M., Sáez, A., Sánchez-Monedero, J., Gutiérrez, P.A., Hervás-Martínez, C.: Tackling the ordinal and imbalance nature of a melanoma image classification problem. In: 2016 International Joint Conference on Neural Networks (IJCNN). pp. 2156–2163. IEEE (2016)
20. Pérez-Ortiz, M., Torres-Jiménez, M., Gutiérrez, P.A., Sánchez-Monedero, J., Hervás-Martínez, C.: Fisher score-based feature selection for ordinal classification: A social survey on subjective well-being. In: International Conference on Hybrid Artificial Intelligence Systems. pp. 597–608. Springer (2016)
21. Sánchez-Monedero, J., Gutiérrez, P.A., Pérez-Ortiz, M.: Orca: A matlab/octave toolbox for ordinal regression. Journal of Machine Learning Research **20**(125), 1–5 (2019)
22. Shannon, C.E.: A mathematical theory of communication. ACM SIGMOBILE mobile computing and communications review **5**(1), 3–55 (2001)
23. Shashua, A., Levin, A.: Ranking with large margin principle: Two approaches. In: Advances in neural information processing systems. pp. 961–968 (2003)
24. Shifaz, A., Pelletier, C., Petitjean, F., Webb, G.: TS-CHIEF: A scalable and accurate forest algorithm for time series classification. ArXiv e-prints **arXiv:1906.10329** (2019), `http://arxiv.org/abs/1906.10329`
25. Ye, L., Keogh, E.: Time series shapelets: a novel technique that allows accurate, interpretable and fast classification. Data Mining and Knowledge Discovery **22**(1-2), 149–182 (2011)