

Automatic Context Selection in Explainable Support Vector Machine Classification

Marcelo Loor
Dept. of Telecommunications
and Information Processing
Ghent University
Ghent, Belgium
Marcelo.Loor@UGent.be

Guy De Tré
Dept. of Telecommunications
and Information Processing
Ghent University
Ghent, Belgium
Guy.DeTre@UGent.be

Abstract—The contextualization of knowledge models has recently been proposed for improving the interpretability of predictions made by explainable support vector machine classifiers. An advantage attributed to the improved classifiers is the capability of selecting the context of the resulting predictions. However, this advantage is inapplicable in situations where a data set with unknown context is provided for evaluation. To make the advantage applicable in such situations, a method for automatic context selection is proposed in this paper. The method uses a reduced number of labeled samples contained in the data set with unknown context to select, from an arrangement of contextualized models, the model whose context is the most similar to the unknown context. The selected context is then used for making the predictions of the remaining elements in the provided data set. Illustrative examples show how the proposed method helps to improve the interpretability and accuracy of the resulting predictions.

Index Terms—explainable SVM, contextualized predictions, augmented fuzzy sets

I. INTRODUCTION

Decision makers willing to use *artificial intelligence* (AI) systems for supporting their decisions can be deterred from doing so when the reasons behind predictions made by those systems are difficult to understand [1]. AI systems consisting of *support vector machines* (SVMs) [2], [3] have been considered as such [4], [5]. To change that and make SVM predictions easier to understand, the contextualization of the resulting predictions through *augmented appraisal degrees* (AADs) [6] has been proposed in [7] and [8]. Analogous to [9], in those studies AADs have been used for handling the membership [10] and nonmembership [11] levels along with the most influential aspects that are found during the evaluation process performed to predict the class(es) an object should belong to. An advantage of this is that those AADs can be used for building *contrastive explanations*, which are statements that try to explain why an object should be part of a class instead of another [12].

After realizing that *contextless* knowledge models have been used for contextualizing the predictions in the aforementioned approach, the authors in [13] have studied how the contextualization of knowledge models can help to improve the interpretability of SVM predictions. Thus, they have proposed an explainable SVM classifier, named XSVM@ctx, which uses

a parallel arrangement of contextualized SVM models for yielding contextualized predictions. An advantage attributed to XSVM@ctx is the fact that decision makers can select the context of the resulting predictions. For instance, a veterinarian can request XSVM@ctx to consider the Cocker Spaniel breed while trying to identify different behaviors in a group of Cocker Spaniels. However, that advantage is applicable only when a particular context, say the Cocker Spaniel breed in the previous example, is known in advance. In this regard, the selection of the context for a data set whose context is unknown becomes a key challenge to support the interpretability of the resulting predictions in contextualized SVM classification.

To solve that challenge, in this paper we propose a novel method by which the (most similar) context for a data set (with unknown context) is automatically selected by means of a reduced number of labeled objects contained in that data set. In this method, the labels (or classes) of the reduced number of objects are first predicted using each of the contextualized SVM models contained in the parallel arrangement built by XSVM@ctx. Then, the context of the model having the highest number of agreements between the predicted and actual labels is selected. As might be noticed, the idea behind the method is that two models sharing the same context would yield the most similar predictions [14].

An important aspect of the proposed method is that the identification of an initial context makes the advantage attributed to XSVM@ctx applicable. Thus, decision makers can direct their attention to the relevant aspects that appear when XSVM@ctx puts a prediction in a more specific context. For instance, after identifying that the behavioral data correspond to the Cocker Spaniel breed, veterinarians can confine their attention to the aspects explaining why a particular behavior should be labeled as ‘behavior A’ instead of ‘behavior B’.

The necessity of a reduced number of labeled objects might be seen as a potential drawback of the proposed method. Even so, we consider the cost of context screening in a reduced number of objects insignificant in comparison to the benefit of improving the interpretability and accuracy of the subsequent predictions. For instance, the cost associated to the analysis and labeling of the behavioral data of some dogs performed by a veterinarian might be of minor importance in comparison

to the improvement of the contrastive explanations of the upcoming predictions.

As shown in the examples given above, the motivation behind the proposed method is to offer reliable and trustable predictions that allow decision makers to arrive at informed decisions. To show how this motivation has influenced the implementation of the method, the paper has been structured as follows. A short summary of the mathematical model behind the contextualization of SVM predictions is presented in Section II. The method, which is the main contribution of this work, is described in Section III. Three illustrative examples of the application of the proposed method are presented in Section IV. The work is concluded in Section V.

II. PRELIMINARIES

To classify an object, usually the level to which the object belongs to one or more classes is first evaluated. Then, the resulting evaluations are used for predicting the most suitable class(es) for that object.

Mathematically such an evaluation can be represented as a *membership grade* in the framework of *fuzzy set theory* [10]. A membership grade, say $\mu_A(x)$, is a number denoting the level to which an object x is member of a set A . This number is usually chosen from the unit interval $[0, 1]$ with 0 and 1 denoting, in that order, the lowest and the highest membership grades. For instance, an evaluation in which the membership of x in A is rather high can be written as $\mu_A(x) = 0.85$.

When an evaluation has been contextualized and the conditions that arise while performing the evaluation are known, it can be better represented as an AAD [6]. An AAD, say $\hat{\mu}_{A@K}(x)$, is a generalization of a membership grade that, in addition to denoting the level to which x is member of A , it indicates why x has been evaluated as such from a particular perspective or knowledge, say K . Due to this aspect, AADs have been proposed for recording the evaluations performed by SVM classifiers and, thus, improving the interpretability of SVM predictions [7], [8]. For instance, the evaluation “the green part in Fig. 1(b) suggests that the handwritten vowel in Fig. 1(a) is an ‘e’ with a computed grade of 0.43 according to K ” can be characterized by $\hat{\mu}_{e'@K}(x) = \langle \mu_{e'@K}(x), F_{\mu_{e'@K}}(x) \rangle$, where $\mu_{e'@K}(x) = 0.43$ and $F_{\mu_{e'@K}}(x)$ is a collection consisting of the references to the green pixels that have been relevant to the evaluation – the interested reader is referred to [9], [18] for practical examples that explain how to model the concept of AADs.

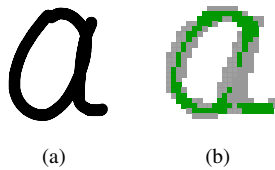


Fig. 1. Is the handwritten vowel an ‘e’?

To handle what is relevant to the evaluations performed by SVM classifiers, one can use the *most influential support*

vectors (MISVs) [8]. The *positive* MISV has aspects that favor the membership of x in A according to K . Likewise, the *negative* MISV has aspects that are against the membership of x in A according to K . While the positive MISV can be characterized by a *membership AAD*, say $\hat{\mu}_{A@K}(x)$, the negative MISV can be characterized by a *nonmembership AAD*, say $\hat{\nu}_{A@K}(x)$. Such MISVs can be used for building *contrastive explanations* [12] like the one shown in Fig. 2¹. As might be noticed in this figure, from a semantical perspective while the positive MISV represents the most relevant example illustrating why the evaluated object should belong to a given class, the negative MISV represents the most relevant counterexample that is against the membership of the object in that class.

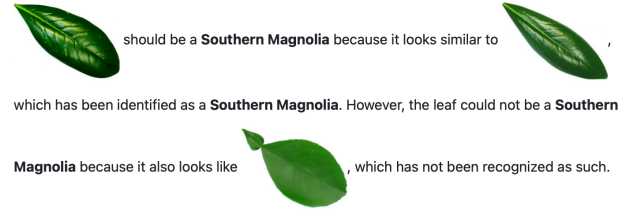


Fig. 2. Example of a contrastive explanation.

Aiming to improve the interpretability of such contrastive explanations, in [13] the authors have studied how the contextualization of models can help to obtain the most appropriate MISVs. Thus, they have proposed XSVM@ctx, which is an explainable SVM classifier that uses contextualized data sets to build an arrangement of contextualized knowledge models. Let $D = \{D_{@ctx_t} | t \in \{1, \dots, T\}\}$ be a collection of T contextualized data sets; and let $K = \{K_{@ctx_t} | t \in \{1, \dots, T\}\}$ be a collection of the T knowledge models associated to D . Under these considerations, the evaluation of the proposition “ x is member of A according to $K_{@ctx_t}$ ” can be denoted by $\hat{\mu}_{A@ctx_t}(x) = \langle \mu_{A@ctx_t}(x), F_{\mu_{A@ctx_t}}(x) \rangle$. The authors have shown that the interpretability can be improved when x and the positive MISV used for obtaining $F_{\mu_{A@ctx_t}}(x)$ share a similar context. In other words, the interpretability of an explanation built with $\hat{\mu}_{A@ctx_t}(x)$ can be improved if $x \in D_{@ctx_t}$. For instance, given that the vowel in Fig. 3(a) has been written by subject $S1$, it is expected that the interpretability of an explanation built with (the positive MISV related to) the vowel written by $S1$ in Fig. 3(b) is better than the interpretability of an explanation built with the vowel written by $S2$ in Fig. 3(c).

As previously stated, an advantage attributed to XSVM@ctx is that one can select the context of the resulting predictions. Thus, since in the previous example it is known that the vowel in Fig. 3(a) has been written by subject $S1$, one can select the model that has been built for $S1$, say $K_{@S1}$, (or another model that is similar to it) to evaluate that vowel. A different situation would have been if the origin (or context) of the aforementioned vowel is unknown. In this case, a mechanism

¹The leaves in this example are part of the data set presented in [15].

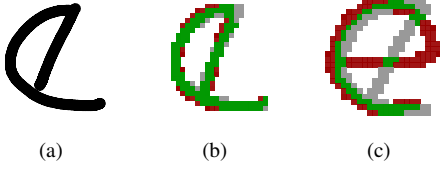


Fig. 3. According to subjects $S1$ and $S2$, is the handwritten vowel an ‘e’?

for identifying that origin is needed to take full advantage of XSVM@ctx. Hence, in the next section we present our novel method for identifying the context of a data set whose context is unknown.

III. AUTOMATIC CONTEXT SELECTION

As was mentioned in the introduction, offering reliable and interpretable SVM predictions with relevant facts that allow decision makers to make informed decisions is a motivational aspect in this work. Since the selection of a suitable context in XSVM@ctx is needed for offering such predictions, in what follows we describe a method by which the most appropriate context for a data set with unknown context is obtained.

Let L be a collection of labels (or classes) associated with the topics under study. Let $K_{@ctx_t} = \{K_{l@ctx_t} | (l \in L)\}$ be an arrangement of knowledge models about the topics in L that have been built using a contextualized data set $D_{@ctx_t}$. Let X be a collection of objects that share an unknown context. Let X_L be a subset of X whose labels are known or have been assigned. Let Y_L be a collection containing the labels associated to X_L . Let $f_p(x, K)$ be a function that predicts the most suitable label for an object $x \in X$ according to the arrangement of knowledge models represented by K . Let $f_a(Y, P)$ be a function that counts the number of agreements between a label $y \in Y$ and a predicted label $p \in P$. Under these considerations, the most suitable context for X is obtained by means of

$$ctx^* = \arg \max_{ctx_t} \{f_a(Y_L, P_{@ctx_t}) | \forall t \in \{1, \dots, T\}\}, \quad (1)$$

where

$$P_{@ctx_t} = \{f_p(x, K_{@ctx_t}) | (x \in X_L)\} \quad (2)$$

and T represents the number of contextual groups.

A general view of the above process is shown in Fig. 4. Notice that the output of the context selection process, which is given by Equation 1, becomes an input of the evaluation process in XSVM@ctx. Notice also that one of the inputs of the context selection process, namely the arrangement of contextualized knowledge models $\{K_{@ctx_1}, K_{@ctx_2}, \dots, K_{@ctx_t}\}$, is an output of the learning process in XSVM@ctx as shown in Fig. 5. These observations clearly illustrate how the proposed method can be incorporated to the XSVM@ctx classifier.

The other input of the context selection process, namely the reduced set of labeled objects denoted by both X_L and Y_L , can be provided by decision makers. This means that the context of the resulting predictions is somehow determined

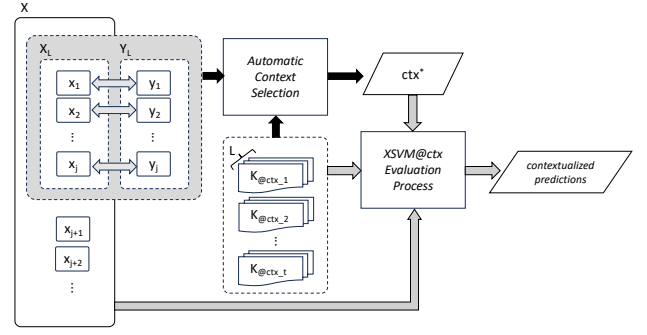


Fig. 4. General view of the automatic context selection process.

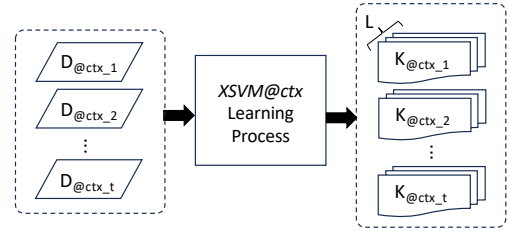


Fig. 5. Contextualized SVM learning process.

by a decision maker, which makes the advantage attributed to XSVM@ctx applicable.

The internal structure of the selection process is shown in Fig. 6. Notice how the knowledge model $K_{@ctx_t}$ associated with context ctx_t is used for making contextualized predictions for each x_j in the subset of labeled objects X_L . These predictions constitute the contextualized collection $P_{@ctx_t}$, which is then compared to the collection Y_L , which contains the labels associated to the objects in X_L . This illustrates how two models sharing a similar context would produce similar predictions [14].

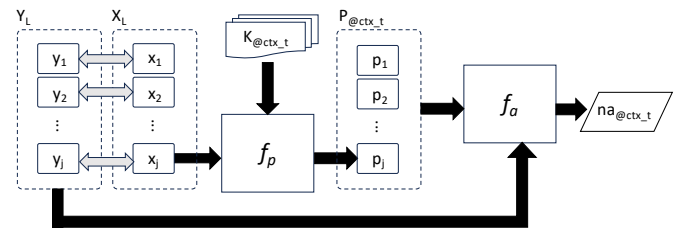


Fig. 6. Number of agreements per context.

A question that arises from the description of the proposed method concerns the size of the reduced set of labeled objects. In principle, one labeled object per class would be necessary, i.e., the size of the reduced set should be at least the number of labels (or classes) in the topic under study. For instance, the veterinarian mentioned in the introduction should identify and label at least one Cocker Spaniel per each behavior under

study. However, this number can increase in situations where the context could not be identified clearly. Such a situation can occur when the number of agreements is the same for two different contexts. In this case, an increase of the number of labeled objects per class would help to differentiate the contexts. If increasing the number does not differentiate them, it is possible that the contexts are similar to each other – the detection of similar contexts in XSVM@ctx is subject to further study.

Another expected question concerns the cost associated with labeling the objects in the reduced set since an expert on the topic under study or specialized tools might be needed. Setting the number of labeled objects as low as possible might be an option for lowering the cost. As was mentioned above, one labeled object per class might be enough. Even though reducing the cost might be seen as option, we consider that the expected improvement in interpretability and accuracy can easily surpass that cost. Examples illustrating such improvements are presented in the next section.

IV. ILLUSTRATIVE EXAMPLES

To show how the proposed method can improve both the accuracy and interpretability of SVM predictions for data sets with unknown context, in this work we have used three different data sets: the toy data set of handwritten vowels presented in [13]; the data set of electromyography (EMG) signals used for recognizing hand gestures presented in [16]; and the data set of hotels bookings presented in [17]. To build the knowledge models associated with these data sets, we use the open source library presented in [18]. We also use this library for implementing our method and obtaining the results presented in this section.

A. Handwritten Vowels

This data set consists of 100 digitized handwritten vowels given by two subjects, which have been identified as $S1$ and $S2$. The data set has been contextualized according to the origin in two subsets of 50 handwritten vowels each, namely $D_{@S1}$ and $D_{@S2}$. Each digitized vowel consists of a square matrix of 28 by 28 pixels whose intensities are linked to the strength of the pen while writing the vowel. Fifty percent of the digitized vowels in each subset (to be precise 5 ‘a’s, 5 ‘e’s, 5 ‘i’s, 5 ‘o’s and 5 ‘u’s) have been used as input of XSVM@ctx for building the contextualized knowledge models for each vowel and for each context (see Fig. 5). For instance, while $K_{e'@S1}$ denotes the knowledge about ‘e’ according to $S1$ that has been built with fifty percent of $D_{@S1}$, $K_{e'@S2}$ denotes the knowledge about ‘e’ according to $S2$ that has been built with fifty percent of $D_{@S2}$. These models have been organized in two arrangements, namely $K_{@S1} = \{K_{a'@S1}, K_{e'@S1}, K_{i'@S1}, K_{o'@S1}, K_{u'@S1}\}$ and $K_{@S2} = \{K_{a'@S2}, K_{e'@S2}, K_{i'@S2}, K_{o'@S2}, K_{u'@S2}\}$ (see Figures 4 and 5). Polynomial kernels of 10 degree with hyper parameters $C = 0.1$ and $\gamma = 2$ have been used with $L2$ -normalized data for building these models.

The remaining fifty percent of the digitized vowels of either $S1$ or $S2$ have been used as the test set X and, thus, they have been considered as if they came from an unknown context. Five digitized handwritten vowels, one per vowel (class), from this test set have been deemed the constituent parts of the collection X_L of labeled objects and have been associated to their corresponding labels in the collection Y_L (see Fig. 4). These collections along with $K_{@S1}$ and $K_{@S2}$ have been used as input of the context selection process to obtain ctx^* as output. The values of ctx^* obtained in several instances of the selection process have been in agreement with the origin of the test data sets. This means that one digitized handwritten vowel per class has been needed to properly identify the context in this example.

To have an indication of the improvement of the accuracy, the predictions obtained with the context chosen by our method have been compared with the predictions obtained with a random context selection. The corresponding confusion matrices are shown in Fig. 7(a) and Fig. 7(b) respectively. Notice that while the accuracy of the class ‘u’ is the same regardless of the selected context, the accuracies of the other classes are improved when the context is chosen by our method. This observation suggests that the knowledge models built for the class ‘u’ according to contexts $S1$ and $S2$ are quite similar.

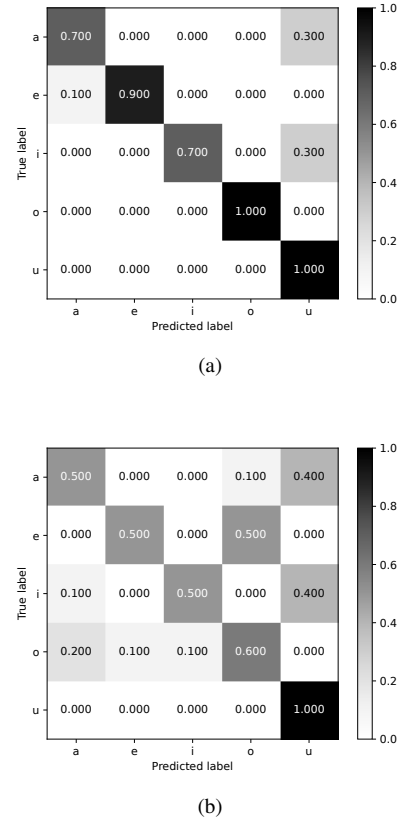


Fig. 7. Handwritten vowels example: predictions obtained with the context chosen by our method (a) vs. predictions obtained with a random context (b).

To have an indication of the improvement in the interpretability of the resulting predictions, the positive MISVs included in the contextualized predictions produced by XSVM@ctx according to contexts $S1$ and $S2$ have been compared to each other. Fig. 8 shows two of such comparisons for two handwritten vowels whose origin is $S1$. One might find easier to explain why the handwritten ‘e’ to the left in Fig. 8(a) has been predicted an ‘e’ by means of the positive MISV in the middle (Context $S1$) than to explain why the drawing is predicted an ‘i’ using the positive MISV to the right (Context $S2$). In a similar way, explaining why the drawing to the left in Fig. 8(b) is an ‘o’ with help of the positive MISV in the middle should be easier than explaining why is an ‘a’ with the positive MISV to the right.

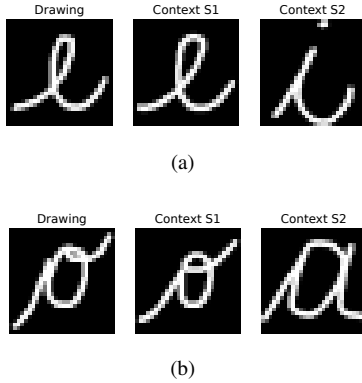


Fig. 8. Handwritten vowels example: Contrasting positive MISVs.

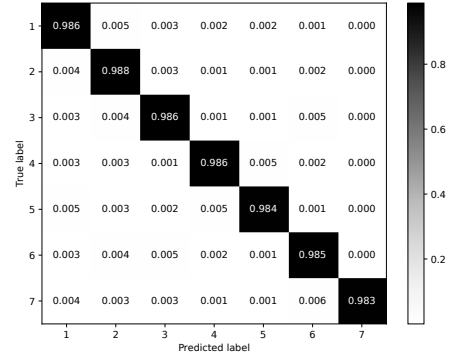
B. Hand Gestures

The original data set consists of 4,237,907 (EMG) signals recorded while 36 subjects were asked to perform the following hand gestures: ‘hand at rest’, ‘hand clenched in a fist’, ‘wrist flexion’, ‘wrist extension’, ‘radial deviations’, ‘ulnar deviations’ and ‘extended palm’, which have been identified, in that order, with natural numbers from 1 to 7. However, 1,512,750 records remain in the curated data set after removing the unlabeled records. The curated data set has been contextualized according to the origin of the signals, i.e., 36 contextualized subsets have been considered. Each record consists of 8 EMG channels, which have been $L2$ normalized. In a similar way to the previous example, 50 percent of the records per gesture in each contextualized subset have been used as training set for building 36 arrangements of knowledge models $K_{@ctx_i}, i \in \{1, \dots, 36\}$, where each arrangement contains the knowledge models built for each gesture, i.e., $K_{@ctx_i} = \{K_{l@ctx_i} | l \in \{1, \dots, 7\}\}$ – recall Figures 4 and 5. RBF kernels with hyper parameters $C = 100$ and $\gamma = 20$ have been used for building these models.

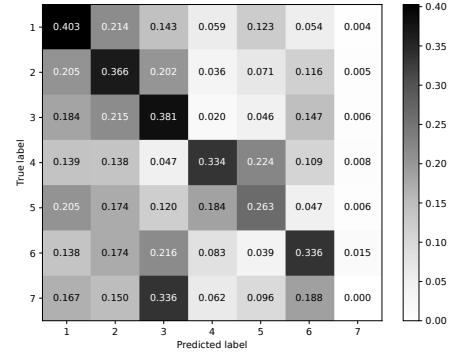
Initially, 7 records (one record per gesture) had been selected from the remaining 50 percent of any of the contextualized test sets to build the collection X_L of labeled records and the collection Y_L with the corresponding labels – recall Fig. 4. Nevertheless, the values of ctx^* obtained in this case

have not been in agreement with the labels in Y_L for all the 36 contexts – recall that $f_a(Y_L, P)$ (see Fig. 6) is a function that counts the number of agreements between a label $y \in Y_L$ and a predicted label $p \in P$. Hence, increasing the number to 3 records per class (i.e., 21 records) have been finally found to be the right number for building the collections X_L and Y_L in any of the 36 contexts.

As was done in the previous example, predictions made with the context chosen by the proposed method have been contrasted to the predictions made with a random context. The results are shown in Fig. 9(a) and Fig. 9(b) respectively. As can be noticed, the accuracy of the predictions made with the context selected by our method exceeds the accuracy of the predictions made with a random context.



(a)



(b)

Fig. 9. Hand gestures example: predictions obtained with the context chosen by our method (a) vs. predictions obtained with a random context (b).

The improvement in the interpretability of the predictions in this example can be observed in Fig. 10. In this case, the signal related to the positive MISV in $K_{@ctx_7}$ (see Fig. 10(a)) is practically equal to the shape of the signal whose gesture is being predicted, which makes easier to explain why the class ‘wrist extension’ has been predicted. In contrast, the signal related to the positive MISV in $K_{@ctx_{26}}$ (see Fig. 10(b)) differs from the shape and the right class of the signal whose gesture is being predicted.

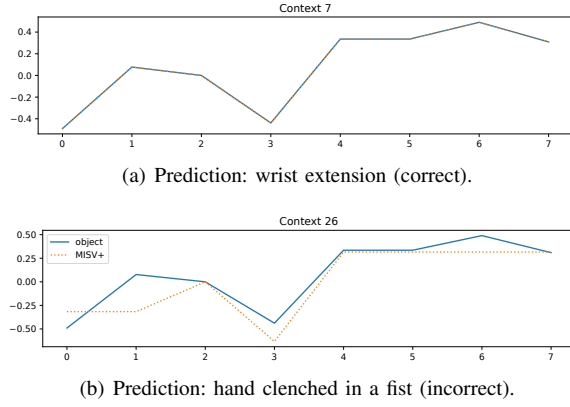


Fig. 10. Hand gestures example: Contrasting positive MISVs.

C. Hotel Bookings

This data set consists of two contextualized subsets of hotel bookings: one corresponds to a resort hotel ($H1$) and the other corresponds to a city hotel ($H2$). While the subset $H1$ contains 40,060 bookings, the subset $H2$ contains 79,330 bookings. Although the original data set consists of 119,390 bookings in total, only 116,958 records remain in the curated data set. Each record in the curated data set consists of the following features: ‘LeadTime’, ‘StaysInWeekendNights’, ‘StaysInWeekNights’, ‘Adults’, ‘Children’, ‘Babies’, ‘IsRepeatedGuest’, ‘PreviousCancellations’, ‘PreviousBookingsNotCanceled’, ‘BookingChanges’, ‘ADR’, ‘RequiredCarParkingSpaces’, ‘TotalOfSpecialRequests’, ‘StayCost’, ‘MarketSegment’, ‘DistributionChannel’ and ‘CustomerType’ – the interested reader is referred to [17] for a detailed description of these features. The categorical features, namely ‘MarketSegment’, ‘DistributionChannel’ and ‘CustomerType’, have been encoded using one-hot encoding scheme. After encoding the categorical features, all the features have been scaled to the unit interval $[0, 1]$. The bookings have been categorized in two groups: bookings that have not been canceled (identified by 0) and bookings that have been canceled (identified by 1). Seventy percent of the bookings per category have been included in the contextualized training sets, and the remaining 30 percent in the contextualized test sets. The contextualized training sets have been used for building the arrangements of knowledge models $K_{@H1} = \{K_{0@H1}, K_{1@H1}\}$ and $K_{@H2} = \{K_{0@H2}, K_{1@H2}\}$. RBF kernels with hyper parameters $C = 100$ and $\gamma = 3$ have been used for building these models. Since the contextualized training sets are unbalanced, class weights computed $w_0 = \frac{n_1}{n_0 + n_1}$ and $w_1 = 1 - w_0$ have been used as additional hyper parameters during the learning process, where n_0 and n_1 are the numbers of bookings identified with labels 0 and 1 respectively.

Twenty records (10 bookings per label) have been selected from the contextualized test sets to build the collection X_L of labeled bookings and the collection Y_L with the corresponding labels – this number of records has been found to be the right number in this example after following the steps suggested in

Section III.

Fig. 11 shows the performance of the predictions: while Fig. 11(a) shows the results obtained with the context selected by our method, Fig. 11(b) shows the results obtained with a random context. Notice that, even though there is an improvement in the accuracy of the predictions made with our method, the improvement is small compared to the improvements obtained in the previous examples. A potential explanation for this situation is that the contexts of $H1$ and $H2$ are rather similar and, thus, either context might be chosen for making the subsequent predictions.

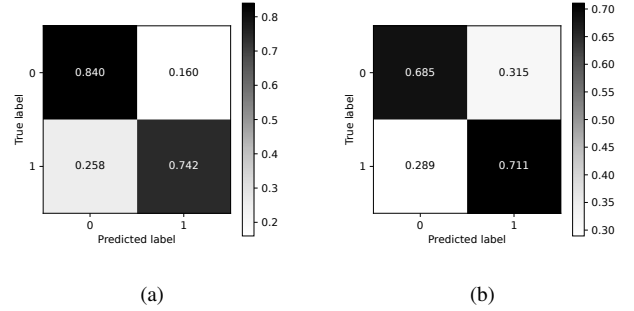


Fig. 11. Hotel bookings example: predictions obtained with the context chosen by our method (a) vs. predictions obtained with a random context (b).

The potential similarity between the contexts $H1$ and $H2$ is also evident when contrasting the corresponding positive MISVs of their predictions as illustrated in Fig. 12. Notice the slight difference between the graphs representing to both MISVs and the graph representing to booking whose label is being predicted.

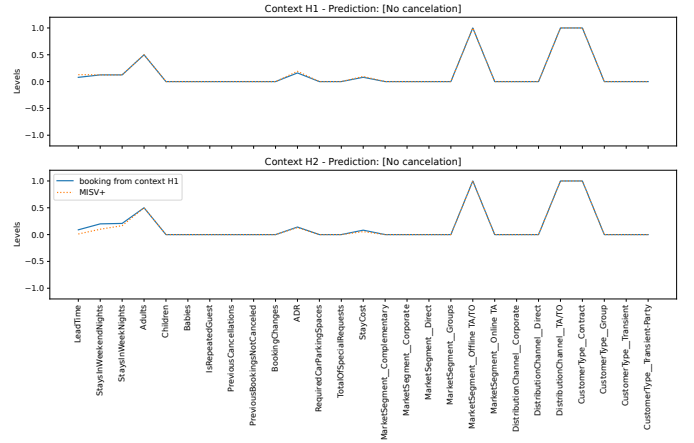


Fig. 12. Hotel bookings example: Contrasting positive MISVs.

D. Discussion

The results obtained in the above examples show that our method is able to select the most appropriate context to be

used in XSVM@ctx and, hence, improve the accuracy and interpretability of the subsequent predictions.

It has also been shown that the number of labeled objects needed to properly identify the context of a data set with unknown context can be very small. This corroborates what was previously mentioned about how insignificant the overall cost of labeling a subset can be in relation to the benefits obtained with the improvement of the accuracy and interpretability.

The small improvement of the accuracy in the third example suggests that the context of the resort hotel is rather similar to the context of a city hotel. A negligible improvement would suggest that the contexts are quite similar and, thus, merging the knowledge models related to these contexts would be a very likely option. This can be observed through the receiver operating characteristic (ROC) curves depicted in Fig. 13, which illustrates the performance of XSVM@ctx with automatic context selection, as well as XSVM@ctx with random context selection – the ROC curve illustrating the performance of contextless XSVM is also depicted as reference. The detection of similar contexts and the conditions that are needed for merging them are subject to further study.

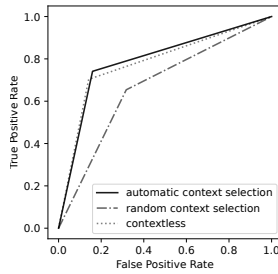


Fig. 13. Hotel bookings example: ROC curves.

As was mentioned in Section III, XSVM@ctx uses contextualized data sets to build a parallel arrangement of contextualized SVM models. Hence, the test sets used in the previous examples are deemed contextualized as well, but with unknown contexts. Notice that these data sets have been contextualized according to the origin of the data, e.g., the EMG signals have been contextualized according to the subject who performed the hand gestures. We consider that such a contextualization is easily applicable to, e.g., medical data whose context is unknown in order to build personalized knowledge models.

V. CONCLUSIONS

In this paper, we have presented a method whereby the context of a data set with unknown context is automatically selected through a reduced subset of labeled objects.

The method is intended for making an advantage attributed to XSVM@ctx applicable in situations where the context of a data set is unknown. With that advantage, decision makers can focus on the relevant aspects that became visible when XSVM@ctx builds a prediction in a more specific context.

Focusing on the relevant aspects, decision makers would understand in a better way the reasons behind such a prediction. Hence, by choosing the right context the proposed method favors the interpretability of the resulting predictions.

The results obtained in three different examples have shown that our method is able to choose the right context and, thus, not only improves the interpretability but also the accuracy of the resulting predictions.

A small improvement in the accuracy reported in one of the examples suggests that choosing the right context is negligible when the contexts are similar. Detecting similar contexts and the necessary conditions to merge them are subject to further research.

ACKNOWLEDGMENT

This research received support from the Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme, and also from ES-POL Polytechnic University under the research project “Interpretable Artificial Intelligence (XAI) in Group Decision-Making (FIEC-200-2020).”

REFERENCES

- [1] D. Gunning, “Explainable Artificial Intelligence (XAI),” 2017. [Online]. Available: www.darpa.mil/attachments/XAIIndustryDay_Final.pptx
- [2] V. N. Vapnik, *The nature of statistical learning theory*. Springer-Verlag New York, Inc., 1995.
- [3] V. N. Vapnik and V. Vapnik, *Statistical learning theory*. Wiley New York, 1998, vol. 1.
- [4] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, “A survey of methods for explaining black box models,” *ACM Comput. Surv.*, vol. 51, no. 5, Aug. 2018.
- [5] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, vol. 58, pp. 82–115, 2020.
- [6] M. Loor and G. De Tré, “On the need for augmented appraisal degrees to handle experience-based evaluations,” *Applied Soft Computing*, vol. 54, pp. 284 – 295, 2017.
- [7] M. Loor and G. De Tré, “Explaining computer predictions with augmented appraisal degrees,” in *2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (EUSFLAT 2019)*. Atlantis Press, 2019/08.
- [8] —, “Contextualizing support vector machine predictions,” *International Journal of Computational Intelligence Systems*, vol. 13, pp. 1483–1497, 2020.
- [9] M. Loor and G. De Tré, “Contextualizing Naive Bayes Predictions,” in *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, M.-J. Lesot, S. Vieira, M. Z. Reformat, J. P. Carvalho, A. Wilbik, B. Bouchon-Meunier, and R. R. Yager, Eds. Cham: Springer International Publishing, 2020, pp. 814–827.
- [10] L. Zadeh, “Fuzzy sets,” *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.
- [11] K. T. Atanassov, “Intuitionistic fuzzy sets,” *Fuzzy sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.
- [12] T. Miller, “Explanation in artificial intelligence: Insights from the social sciences,” *Artificial Intelligence*, vol. 267, pp. 1–38, 2019.
- [13] M. Loor, A. Tapia-Rosero, and G. De Tré, “Contextual Boosting to Explainable SVM Classification,” in *Fuzzy Logic and Technology, and Aggregation Operators*, S. Massanet, S. Montes, D. Ruiz-Aguilera, and M. González-Hidalgo, Eds. Cham: Springer Nature Switzerland, 2023, pp. 480–491.
- [14] M. Loor and G. De Tré, “Identifying and properly handling context in crowdsourcing,” *Applied Soft Computing*, vol. 73, pp. 203 – 214, 2018.

- [15] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *2007 IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 11–16.
- [16] S. Lobov, N. Krilova, I. Kastalskiy, V. Kazantsev, and V. Makarov, "EMG data for gestures data set - UCI machine learning repository," 2019. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures>
- [17] N. Antonio, A. de Almeida, and L. Nunes, "Hotel booking demand datasets," *Data in Brief*, vol. 22, pp. 41–49, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352340918315191>
- [18] M. Loor, A. Tapia-Rosero, and G. De Tré, "An Open-Source Software Library for Explainable Support Vector Machine Classification," in *2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2022, pp. 1–7.