IXVC: An interactive pipeline for explaining visual clusters in dimensionality reduction visualizations with decision trees

Adrien Bibal [*,1], Antoine Clarinval[1], Bruno Dumas, Benoît Frénay

PReCISE, Faculty of Computer Science, NADI, University of Namur, Rue Grandgagnage 21, B-5000 Namur, Belgium

A B S T R A C T

High-dimensional data with many features are usually challenging to represent with standard visualization techniques. Usually, one has to resort to dimensionality reduction techniques such as PCA, MDS or *t*-SNE to represent such data. Such dimensionality reduction techniques make it possible to highlight the high-dimensional structures of data. In many of such visualizations, comparable instances appear to form visual clusters. However, no feedback is directly given by these techniques to the user about the features that make the instances cluster together in the visualization. As such, the interpretation of which features define a given visual cluster is a complicated task. In this paper, we propose a novel interactive approach (called Interactive eXplanation of Visual Clusters — IXVC) to explain dimensionality reduction visualizations by mapping their clusters to explanations provided by decision trees. The decision trees use features in high-dimensional data to explain two-dimensional clusters, filling the gap between the dimensionality reduction visualization and the original data.

1. Introduction

In machine learning, dimensionality reduction (DR) techniques are designed to reduce the number of features of the original high-dimensional (HD) data. Reducing the number of features to, e.g., two dimensions, provides low-dimensional (LD) data that can be visually presented to users. DR techniques are used in many different fields. For instance, multidimensional scaling (MDS) [1] is used in psychology to explore data or validate hypotheses [2,3] (e.g., see Koch et al. [4]). Principal component analysis (PCA) [5] is another famous technique that can produce visualizations if the first components are kept.

One of the main differences between MDS and PCA is their *interpretability*, which is defined by the intrinsic capacity of a model to be understandable [6,7]. In the context of DR visualizations, the link between the dimensions of a PCA visualization (called *principal components*) and the corresponding HD data is commonly considered interpretable, as the principal components are linear combinations of the HD features. By looking at the weights in the linear combinations, features from HD data that are used for defining the LD dimensions can be identified. On the contrary, the mapping between the HD dimensions and the LD dimensions produced by nonlinear DR (NLDR) techniques is not always clear. This lack of interpretability is an issue, but methods can be developed to *explain* such black-box models or mappings [8]. In

machine learning, explaining a black-box model or mapping is defined as the use of external resources, e.g., interpretable models, to provide insights about the black box behavior [8].

In some cases of NLDR visualizations (e.g., with *t*-SNE [9]), the dimensions have no meaning and therefore cannot be used as a basis for explanation. Instead, explanations must rely on visual clusters present in the visualization. However, there are issues pertaining to visual cluster analysis such as arbitrary cluster shapes and the analyst's intuitiveness injected in the explaining process. Currently, none of the approaches proposed in the literature address most issues related to explaining NLDR through visual clusters. This paper aims to fill this gap and studies the following research question:

> If visual clusters clearly appear in a given NLDR visualization, can we explain these visual clusters based on the original dimensions?

In order to handle this research question, corresponding to explaining black-box DR mappings through visual clusters, an interactive pipeline is proposed. Our pipeline, called *Interactive eXplanation of Visual Clusters* (IXVC), explains the link between clusters visually present in LD and the original HD features by using a decision tree [10]. Decision trees are considered for providing explanations because they

---

can nonlinearly predict the visual clusters while being interpretable. Furthermore, decision trees stay interpretable even in the case where the original data are high-dimensional (as opposed to, e.g., linear models), as the decisions in the trees consider features one by one. This makes the proposed solution scalable in terms of the number of HD dimensions. The pipeline is interactive and therefore involves the analyst in the selection of clusters to be explained. IXVC is implemented in a web application that has been used for the pipeline evaluation.

In order to present IXVC, Section 2 reviews the explanation of DR visualizations through dimensions and clusters. Section 3.1 motivates the need for the explanation of visual clusters. Section 3 then introduces IXVC. Section 4 presents the tool that implements IXVC, as well as examples of use. A user-based experiment has been conducted for evaluating the pipeline and the tool, and is presented in Section 6. A discussion on the limitations of IXVC is presented in Section 6.4. Directions for future work are proposed in Section 8 and Section 9 concludes the paper.

2. Explaining DR visualizations

Dimensionality reduction (DR) is the process of reducing the number of features that are available in high dimension. DR techniques are often used for data exploration through visualization techniques, which is made easier when the number of features is reduced. For instance, scatter plots can be used when the number of dimensions is reduced to two. Fig. 1 presents an example of a DR visualization.

Through the DR process, information is inevitably lost. In multidimensional scaling (MDS), for instance, the measure of this loss, called the *stress*, is defined as the difference between pairwise distances between instances in HD and in LD. More formally, let d_{ij}^{HD} be the distance between the instances i and j in HD, and d_{ij}^{LD} the distance between the instances i and j in LD, the Kruskal's stress [1] is defined as

$$Stress = \sqrt{\frac{\sum_{ij} \left(d_{ij}^{HD} - d_{ij}^{LD}\right)^2}{\sum_{ij} d_{ij}^{HD^2}}}.$$

The DR loss of information, called *DR errors* in this paper for the sake of generality, is an essential element to consider while interpreting or explaining DR visualizations. Indeed, because of DR errors, some instances are not positioned correctly in LD, with respect to their position in HD. These DR errors make the task of analyzing the visualization more difficult. Some visual techniques, as reviewed by Nonato and Aupetit [11], have already been developed to hint the presence of DR errors in visualizations (see, e.g., [12–15]).

Interpreting a particular DR means understanding the mapping between the instances in HD and the corresponding instances in 2D. When the mapping is not interpretable, techniques can be used to explain it. There are two main ways to interpret or explain DR mappings [1]. First, the mapping can be interpreted or explained by focusing on the interpretation of the two new dimensions. The literature concerned by the interpretation and the explanation of the reduced dimensions is developed in Section 2.1. Second, the visual clusters in the 2D visualization can also be used to find an interpretation or explanation, as developed in Section 2.2.

2.1. Explaining DR visualizations using dimensions

Among the two ways to interpret DR visualizations, using the LD dimensions is the most widespread in the literature. Some DR techniques, such as *principal component analysis* (PCA) [5], are considered interpretable because the mapping between the HD and LD dimensions is defined by linear combinations of the HD features. One classical way to link reduced dimensions from a linear DR and the original HD features is by using axis legends [17]. These legends are often represented as bar charts representing the contribution of HD features for each new
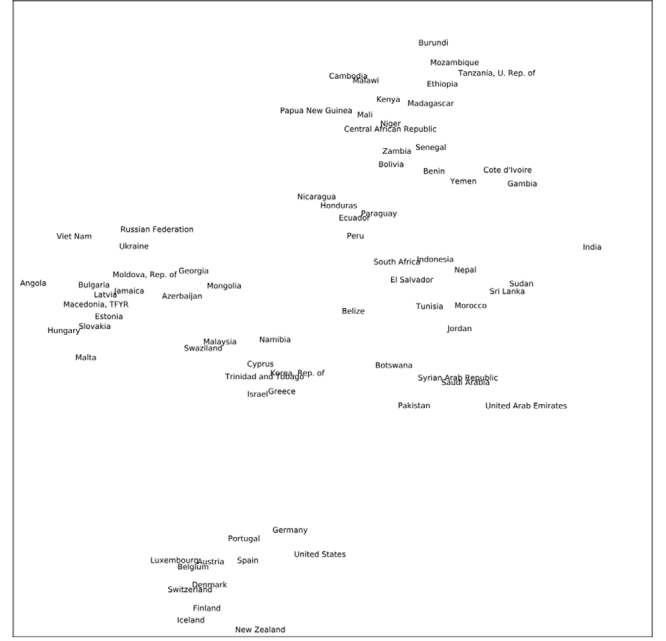

Fig. 1. DR visualization (generated by *t*-SNE [9]) of the 2006 Human Development Report [16]. This visualization is composed of 76 countries from the dataset. Based on the HD socio-economic features, the DR derives two dimensions. Even if visual clusters can intuitively be identified, it is not clear how the HD features have been used to generate them.

dimension. In the case of linear DR, these contributions are contained in the weights. Another way to visualize the contribution of the HD features to the embedding constructed by a linear DR technique is by using biplots [18,19]. Biplots are plots to visualize the instances, such as in traditional scatter plots, but the HD features as well. The HD features are visualized as vectors in the biplot, where their direction and length are based on the contribution of each HD feature to the two reduced dimensions.

In the case of nonlinear DR (NLDR), the contribution of each HD feature to the reduced dimensions is, most of the time, not given through parameters, which makes NLDR mappings hard to interpret [20]. For transferring biplots to the NLDR case, Coimbra et al. make uniform perturbations in the values of each HD feature, while setting the unperturbed HD features to their mean value [21]. By doing so, they obtain curved axes representing the tendencies of HD features in the 2D plot. Following the same idea, Cavallo and Demiralp propose to draw *prolines* for each point and each feature of interest in a scatter plot [22]. A proline is drawn in LD by creating new samples by varying the value of a feature of interest, while all others are fixed, and then by computing the projection of all generated samples to LD. The proline corresponds to the line that connects the projection in 2D of all created samples.

Coimbra et al. also present axis legends for NLDR, based on their curved biplot axes [21]. They define the height of the bars corresponding to the contributions of the HD features to the reduced dimensions by a combination of how parallel the projected curve is to each scatter plot axis, and of how linear the curve is.

Turkay et al. [23] developed an interactive analysis of DR dimensions using dual views of the instances and of the HD features. The idea is to select instances that are then used to compute a scatter plot in which the points are features from the HD space and where the dimensions are chosen statistics. Users can then choose points in this scatter plot, i.e. HD features, which are then the features used to compute a new DR visualization of the instances.

Similarly, Yuan et al. [24] propose a dual view composed of a DR visualization of the HD features and a DR visualization of the instances.

The two visualizations are linked to one another as selecting elements in one visualization is reflected on the other visualization (e.g., points representing the features selected on one visualization are used to compute a new visualization of the instances). While the works of Turkay et al. and Yuan et al. allow users to have a sense of the impact of the original features on a DR visualization of instances, they do not provide an approximation of the DR mapping that uses the HD features to build the two dimensions of the visualization.

Another way to deal with the interpretability issue of NLDR is by transforming the mapping to be linear. For instance, Gisbrecht et al. apply a linear kernel to the NLDR algorithm *t*-SNE [9], in order to make the mapping linear [25].

External resources can also be used to explain the LD dimensions. For instance, social scientists often use *property fitting* (PROFIT) [26] to find trends in a visualization (for an example of use of PROFIT, see Koch et al. [4]). These trends are created using linear combinations of features that have not been used to make the visualization (i.e. external features). *Best interpretable rotation* (BIR) is another solution that explains DR visualization dimensions by using linear combinations of external features [27,28].

2.2. Explaining DR visualizations using clusters

As shown in Fig. 1, visual clusters can be formed in a DR visualization. Explaining a DR visualization mapping using its visual clusters is another way to explain the HD-to-2D mapping. This task is related to the combination of the *verify clusters* task of Brehmer et al. and the *name clusters* task [29]. Nonato and Aupetit call this task *discover relation between visual pattern and original dimensions* [11]. The visual clusters that serve as basis for an explanation can either be identified manually by a user or automatically by a clustering algorithm.

It can be useful to use clustering algorithms in HD to get some clues about the visual clusters in 2D. For instance, if *k*-means [5] is used on HD instances and the labels of the HD clusters found are shown in 2D, it is possible to use the relation between the HD features and the centroids of the HD clusters to get some insights about the visual clusters. However, first, the visual clusters may not correspond to the HD clusters found by *k*-means. Second, the HD centroids provided to the user by *k*-means are defined in terms of each and every HD feature, which makes them difficult to interpret in practice. Wenskovitch et al. review different ways to combine DR techniques with clusters automatically found by clustering algorithms [30].

One recent example of such a combination is Clustrophile 2 [31], which lets users select the DR technique and the clustering algorithm applied in HD in order to visualize HD clusters. Some explanations about the HD clusters are provided alongside visualizations, such as the importance of each feature in each cluster under the form of a heatmap and a decision tree predicting the HD clusters with the HD features. While Clustrophile 2 has some connection with our work, it is designed to explain clustering algorithms applied on the HD data, while our problem is the explanation of the HD-to -2D mapping through the visual clusters produced by the DR method used. Again, the key distinction being that the 2D visual clusters do not necessarily correspond to HD clusters automatically found by, e.g., *k*-means applied in HD. Our work focuses on explaining the visual clusters that users see in the DR visualization and do not understand, hence the need for a manual clustering.

When the visual clusters are identified, an explanation of these clusters can be provided. Most of the time, the explanation is provided by experts (e.g., Lebel et al. [32]). This makes the *name clusters* task subjective [2]. One drawback of this approach is the extra knowledge experts may inject during the explanation that does not come from the data used to generate the visualization. Furthermore, even if they do not inject extra knowledge (e.g., by restricting their explanation to the original HD features), it is still difficult for a user to explain how the HD features are combined to form clusters in 2D.

Approaches exist in the literature to detect and rank features according to their significance regarding a classification, a regression or a clustering procedure [33]. For instance, *recursive feature elimination* (RFE) can be used to remove features one by one iteratively by using the model coefficients in the case of linear models or the feature importance score in, e.g., random forests [33]. In the case of clustering, several different metrics of feature interest can be used to prune the initial set of features [34] (see, e.g., [35–38]).

da Silva et al. propose to rank HD features according to an euclidean ranking and a variance ranking [39]. For each instance in the dataset, a set of neighbors is chosen by the user in LD and the euclidean distance between each instance and their LD neighbors is computed for each HD feature. The visualization is then colored by following the top ranked HD features in the different neighborhoods.

After automatically detecting clusters using a grid in LD, Kandogan proposes to rank HD features by labeling each cluster according to a score associated to each HD feature [40]. This score is computed for each automatically detected LD cluster and for each HD feature as a linear combination of measures on properties of the LD cluster (e.g., its density). The weights of the linear combinations, i.e. the importance of the properties, are set by the user.

Joia et al. use a *singular-value decomposition* (SVD) on the transposed matrices containing the HD features of instances contained in each automatically found cluster to compute the importance of HD features for those clusters [41].

Rauber et al. propose to let the user select a group of instances and a ranking of HD features is provided following a *discrimination* criterion (i.e. how individual HD features explain the separation of selected instances from the rest) or a *coherence* criterion (i.e. how the *compactness* of the selected instances are explained by individual HD features) [42]. However, the HD features are not combined for the explanation.

Parisot et al. use an evolutionary algorithm in order to find the dataset preprocessing that leads to a new dataset for which a clustering result is easier to interpret [43]. In order to find such a new dataset, the objective of the evolutionary algorithm is to find a small decision tree that is used to explain the clustering of the preprocessed dataset, while having a clustering on the preprocessed dataset that is as similar as possible to the clustering on the original dataset.

van Ham et al. consider a scatter plot made from two HD features and use a decision tree to explain a selection of instances in the scatter plot by using HD features that are not used in the scatter plot [44]. Their decision tree is a binary classification tree that has the task of explaining the selected instances versus all others, which does not address the issue of explaining visual clusters in DR visualizations.

Contrasting clusters in PCA (ccPCA) [45] adapts contrasting PCA to find the HD features that best explain the contrast between a particular cluster versus the others in a visualization.

t-viSNE is a tool that includes different techniques for getting insights about *t*-SNE visualizations [46]. In particular, the authors use the first component of a PCA on user-selected points in order to know the main HD features that describe the selected points. However, (i) this selection has no link with the projection, as it explains the HD points instead of how the HD points are projected in LD, and (ii) the explanation is linear because of the PCA, while the projection is nonlinear. In order to take into account these limits, another tool that considers polylines is proposed in t-viSNE. The idea is to draw lines in the visualization and then rank the HD features according to how they explain, for each LD dimension, the order of the LD points on the polylines. Concerning the problem we address in our work, the shortcomings of this tool are that (i) the dimensions are explained instead of the clusters and (ii) the relative importance of the HD features is known, but not how they are combined to explain the dimensions.

A recent approach based on linear segments from Ma and Maciejewski has been proposed to explain NLDR visualizations [47]. The idea, based on LIME in classification [48], is to locally and linearly explain

the NLDR visualization. The two drawbacks of such a technique is that (1) the visualization is explained locally and not globally by a model and (2) the linear models do not scale well with the number of HD features. Because of this last issue, only the top features used in each model are presented.

The idea of generating a DT from a user selection of clusters in a scatter plot, as done in this work, was previously considered by Ware et al. [49]. In our work, however, instead of representing two HD features, the scatter plot is the result of a DR process and the DT is used to approximate the DR mapping. Moreover, the decision tree of Ware et al. is built manually by defining splits through the scatter plot, whereas it is automatically generated from the selection of clusters in our method.

3. Interactive explanation of clusters using decision trees

NLDR visualizations are used without any clue about the mapping between the visualization and the original features that were used to generate it. In order to introduce the importance of explaining this mapping through visual clusters, Section 3.1 presents some challenges related to the state-of-the-art neighborhood preservation techniques. Section 3.2 then proposes answers to the challenges discussed in Sections 2.2 and 3.1. Section 3.3 builds on the proposed answers and presents IXVC, a machine learning pipeline that helps data analysts to explain DR visualizations through visual clusters.

3.1. Challenges when explaining DR through clusters

Explaining DR visualizations through their visual clusters is needed, but the existing solutions that could be used face several challenges. In order to make these challenges explicit, we take neighborhood preservation techniques (*t*-SNE, UMAP, etc.), as an example. The objective of these techniques is to preserve HD proximity by making neighbors two instances in 2D if they are neighbors in HD. More precisely, the closer the instances are in HD, the more these techniques try to put them close in 2D. An example of a *t*-SNE visualization is shown in Fig. 1. Each following challenges is given a name in square brackets.

Explanations Based on Dimensions may not be Possible [Dimensions not Explainable (DNE)] First, given its focus on neighborhoods in HD, neighborhood preserving techniques naturally tend to accentuate clusters in 2D, which makes them good candidates for the explanation through clusters. Furthermore, the dimensions of the visualizations produced by such techniques have no meaning [50] and cannot be used as a basis for explanations. Because of that, only the cluster approach for explaining can be used, and no technique for explaining through dimensions can make neighborhood preserving techniques more interpretable [50], unless they are modified (e.g., Gisbrecht et al. [25]).

LD Clusters have Arbitrary Shapes [Arbitrary Shapes (AS)] Second, visual clusters resulting from neighborhood preserving techniques can have complex shapes, which make clustering algorithms with predefined cluster shapes, such as *k*-means, not suitable. Indeed, the predefined shapes of such clustering algorithms restrain the possible cluster explanations to the clusters that can be possibly formed by the clustering algorithm.

LD Clusters can be Misleading [Misleading Clusters (MC)] Third, cluster analysis performed manually by experts, or automatically by clustering algorithms, can be misleading because of the propensity of neighborhood preserving techniques to show clusters. Indeed, despite their strength in the detection of real HD clusters, these techniques are also known for sometimes presenting clusters in 2D that do not exist in HD [50].

All the issues are shared by the state-of-the-art NLDR algorithms, such as *t*-SNE [9], UMAP [51] and LargeVis [52]. To the best of our knowledge, no techniques addressing all these issues exist in the literature. IXVC, the solution proposed in this paper to the task *discover relation between visual pattern and original dimensions* [11] tackles these issues.

3.2. Answers to the cluster explanation challenges

A first issue, mentioned in Section 2.2, is the intuitive explanation of clusters. This issue arises when data analysts use their intuition to explain clusters that are made of errors from the DR algorithm. Having access to objective reasons behind visual clusters, based on the HD features, would help data analysts to overcome this issue of intuitive assessment.

The issue called [AS], presented in Section 3.1, concerns the possibility for 2D clusters to have arbitrary complex shapes. The hypotheses made on the form of clusters by clustering algorithms may not be suitable when clusters take complex shapes. However, in the case of a 2D visualization analysis, data analysts can draw the limits of visual clusters themselves (e.g., with a hand-made selection).

The issue [MC] is about the role of DR errors in cluster explanations. In this case, having a feedback on DR errors may help data analysts to explain the mapping. Indeed, if individual errors, for each instance in 2D, are provided to data analysts, it would be possible to decide whether to discard or not some instances during the visual analysis. This issue is important, since the presence of instances erroneously placed in visual clusters because of DR errors can mislead the analyst.

The issue [DNE], related to the absence of meaning of the visualization dimensions, forces the use of the visual cluster to understand the HD-to -2D mapping. When dimensions can be explained, like with MDS visualizations, techniques can be used to approximate the mapping between the original features and the dimensions of the visualization. However, when explaining the visualization dimensions do not make sense, like in *t*-SNE and UMAP [50], a way to approximate the mapping between the original features and the clusters in the visualization must be found.

Ideally, in addition to an easier explanation of visual clusters, providing feedback on visual clusters would also help data analysts to decide to take several actions. For instance, in order to improve the interpretation, they may want to choose a more appropriate DR algorithm, to change the *hyperparameters* (or *meta-parameters*) of the DR algorithm or to remove instances that make the DR process difficult.

3.3. IXVC: the interactive machine learning pipeline

In this section, the interactive pipeline developed for explaining clusters in DR visualizations, *Interactive eXplanation of Visual Clusters* (IXVC), is presented.

The pipeline is used in the context of a DR visualization exploration. As such, the first step is to consider a particular DR visualization **V** ($n \times 2$) built from a dataset **X** ($n \times d$), which corresponds to the scatter plot ① in Fig. 2. The error made by the DR algorithm for each of the *n* instance in the visualization should be provided, in order for the data analyst to unselect elements that have a DR error that is too high. Unselecting elements, in our context, only means that instances that are erroneously projected are hidden in the visualization, so that they do not interfere with users analysis.

The second step is to manually select the visual clusters for which the analyst wants an objective explanation (see P₁ in Fig. 2). All of the *n* instances in the visualization do not have to be selected and the number *p* of 2D clusters can be arbitrary.

Next, a decision tree (DT) is built based on the *p* 2D clusters provided in the second step (see DT ②a in Fig. 2). The 2D cluster memberships serve as labels for the decision tree and the *d* original features of **X**, the HD data, are the criteria on which the decisions in the decision tree are made. As mentioned in Section 2.2, one can note that decision trees can be used to explain HD clusters (see, e.g., [43,53,54]). However, our task is different because the goal is not to cluster data in HD, but to understand a given DR visualization through its 2D visual clusters. Therefore, we propose to understand the visualization by interactively querying the meaning of visual clusters of interest.
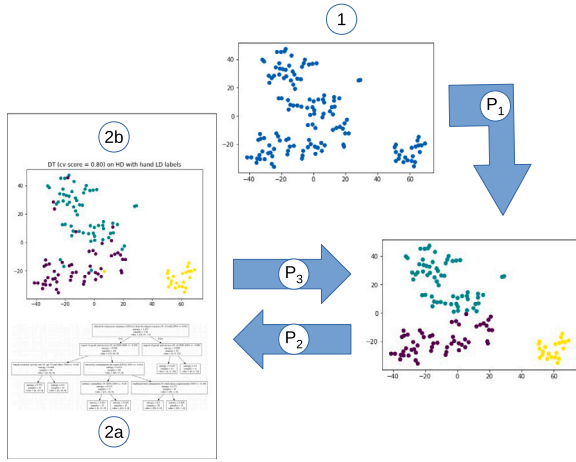
Fig. 2. From a given DR visualization ①, the data analyst manually selects visual clusters (P₁) (the colors correspond to the selected clusters), and a visualization of the errors made by a decision tree (2a) explaining the manual clusters using the HD features is provided (2b). Given the provided feedback, a new manual clustering can be performed by the data analyst (P₃).

A *k*-fold cross validation is then used to select the best hyperparameters for the decision tree (see (2a) in Fig. 2). By doing so, the decision tree provides its best possible solution for interpreting the DR visualization **V** by explaining the *p* manually selected clusters by using the *d* original features. While the first information provided by the decision tree to the analyst is the explanation in terms of the original features, the second information comes from the prediction errors of the decision tree (see the scatter plot (2b) in Fig. 2). Indeed, visualizing the errors made by the DT when predicting the selected visual clusters allows the analyst to see where, in the visualization, the clustering of the analyst cannot be explained. This information may hint that the DT cannot help in the explanation of the selected visual clusters, but can also hint that the selected 2D clusters do not represent real HD clusters and, so, that the instances need to be clustered in a different way in 2D.

Finally, the analyst may either stop the analysis or choose to cluster the instances differently. In the latter case, the analyst proceeds with the second step again by selecting other visual clusters (see (P₃) in Fig. 2), which are then explained with a new decision tree. In the former case, the analyst stops the analysis because the multiple iterative explanations of clusters have provided enough information to the analyst. IXVC is summed up in Algorithm 1.

Sacha et al. [55] developed a process model describing interactive DR with seven scenarios of interaction. Although our work proposes to combine user interaction and DR, this interaction takes place after the DR was performed rather than during its computation. In our approach, the data, feature space, and the DR in the process model of Sacha et al. [55] are considered as given, and the interaction augments the data in order to train a DT that will, in turn, augment the visualization (see Fig. 3). These augmentations will help the analyst understand how clusters are mapped onto the embedding. Scenarios S1 (i.e. data selection) and S2 (i.e. annotation and labeling) from Sacha et al. [55] are supported in the explaining process, as users can filter instances and define visual clusters (hence, assigning a label to instances) to build the DT.

Note that as the goal is to help understanding visual clusters in a visualization, and not an automatically computed clustering, users can draw the frontiers of the clusters they see and assess the explanation

---

Algorithm 1: IXVC

> **Data:** An input matrix **X** and a visualization **V**
> **Result:** An explanation of **V** that satisfies users
> **while** *users are not satisfied with the explanation* **do**
> > ask users to group some instances in visualization **V** into *p* clusters **c**;
> > **foreach** *fold f in k folds of* **X** **do**
> > > **foreach** *hyperparameter values* **hp** **do**
> > > > $DT_{-f, \mathbf{hp}} = \text{train\_decision\_tree}(\text{data}=\mathbf{X}_{-f},$
> > > > $\text{target}=\mathbf{c}_{-f}, \text{hyperparameters}=\mathbf{hp});$
> > > > $\text{score}_{-f, \mathbf{hp}} = \text{predict}(DT_{-f, \mathbf{hp}}, \mathbf{X}_f)$
> > > **end**
> > **end**
> > $\mathbf{hp}^* = \underset{\mathbf{hp}}{\arg\max} \ \text{mean}(\text{score}_{\cdot, \mathbf{hp}});$
> > $DT^* = \text{train\_decision\_tree}(\text{data}=\mathbf{X}, \text{target}=\mathbf{c},$
> > $\text{hyperparameters}=\mathbf{hp}^*);$
> > show DT^* to users;
> > show $\text{predict}(DT^*, \mathbf{X})$ in the visualization **V** to users;
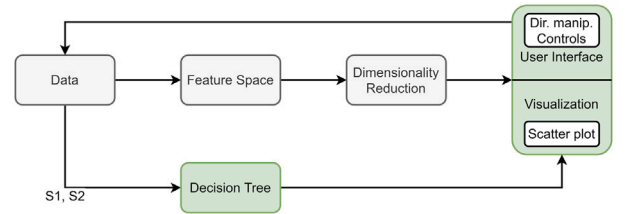> **end**

Fig. 3. In our approach, the data, feature space, and the DR in the process model from Sacha et al. [55] are considered as given, and the interaction augments the data in order to train a DT that will, in turn, augment the visualization.

received via the DT. This particular setup explains why our pipeline is interactive (users must be in the loop), as well as iterative (users can try other explanations in order to expand their understanding of the visualization).

4. Interactive explanation interface

This section introduces the interface implemented to evaluate IXVC. Fig. 4 shows the IXVC interface. The top part of the interface shows the DR scatter plot from which the user selects the clusters (top-left), a list of the selected clusters (top-middle) and a scatter plot (top-right) showing the predictions resulting from the DT (bottom of the interface).

The IXVC interface is implemented as a web application running on a Python web server. The visuals were developed in Javascript using the D3.js library [56]. The Python web server handles the execution of the machine learning algorithms, such as the decision tree, using scikit-learn [57]. For the evaluation of IXVC, visualizations generated by *t*-SNE are used.

When launching the IXVC interface, the user is presented with a scatter plot (located at the top-left part of the interface) generated by running *t*-SNE (without PCA preprocessing). Each instance is represented as a black dot with an associated text label showing its name, thus allowing the identification of individual instances. The individual errors resulting from *t*-SNE are depicted by the opacity of each dot, with the whitest dots representing the highest error. This error for each individual instance *i* is measured using the individual Kullback–Leibler divergence loss

$$D_{KL}(\mathbf{p}_i \parallel \mathbf{q}_i) = \mathbf{p}_i \log(\mathbf{p}_i / \mathbf{q}_i),$$
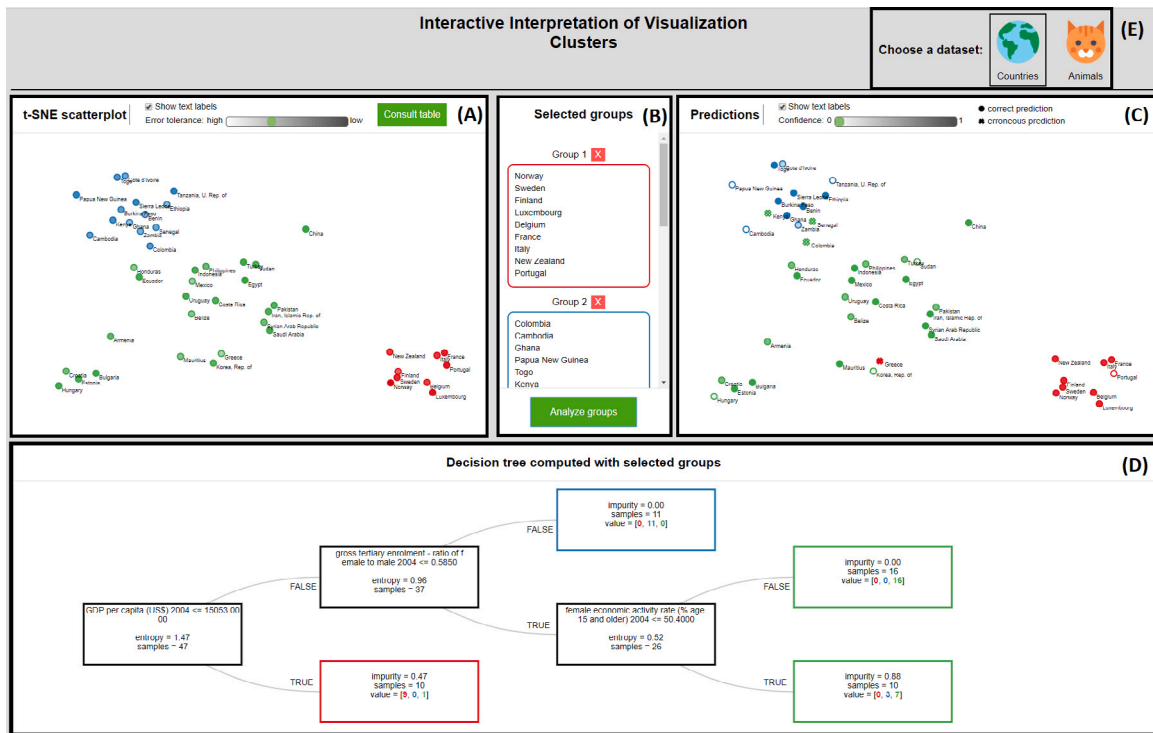
Fig. 4. IXVC interface. Top left scatter plot (A) corresponds to the DR visualization. Instances are colored with respect to the user selection of clusters shown in the top middle part (B). On the bottom, the decision tree (D) explaining the user selection is provided. The colors corresponding to the tree predictions are presented in the top right visualization (C). For the evaluation (see Section 6), users can switch between the *country* and the *zoo* datasets by using the earth and cat icons (E).

where $\mathbf{p}_i$ (resp. $\mathbf{q}_i$) is a vector containing the probability of each n instance to be neighbor of the instance i in HD (resp. 2D). A low divergence for i means that the neighborhood of i in HD is well preserved in 2D. Instances can be filtered out according to a threshold defined by the user on the individual Kullback–Leibler divergence loss (called *DR error* in the interface for the sake of generality). The DR error threshold is labeled as *error tolerance* instead of *loss tolerance* in the interface, as a preliminary evaluation (see Section 6.1) of the interface suggested that it is more meaningful for users and more generic formulated as such.

A major challenge when displaying a scatter plot is the visual clutter that can occur when there are numerous data points to show, which can impede the analyst's work and cause delay in the rendering of the visualization in the interface. Previous work in the literature suggests to implement techniques, including interaction features, in order to tackle visual clutter (e.g., [58,59]) and to ensure that the visualization at hand possesses desirable properties such as scalability and individual data point localization. Ellis and Dix [59] have identified eight properties and eleven clutter reduction techniques that can be used to achieve these desirable properties. In the context of the IXVC interface, three clutter reduction techniques, namely sampling (discussed in Section 6), filtering and opacity, were implemented. This combination allows us to obtain all the desirable properties listed in Ellis and Dix [59] that are necessary to the interface. In particular, the scalability regarding the number of data points (achieved through sampling and filtering), and the ability to discriminate individual points on the visual representation (achieved through opacity) are of utmost importance.

The scatter plot provides a lasso-like interaction allowing users to select visual clusters. The selected instances are subsequently colored alike to mark their belonging to the same cluster following a categorical color scale generated with ColorBrewer [60]. The clusters thus defined by the user are displayed in a pane to the right of the scatter plot. The interface uses the word *groups* instead of *clusters* in order for the evaluation participants to avoid the confusion with clusters that would be obtained from an automatic clustering technique.

When the user is finished selecting visual clusters, a decision tree generated from the cluster selection is displayed under the scatter plot. The decision tree attempts to predict the cluster of each selected instance using the HD features. The representation of the decision tree shows the features selected to build the tree as well as the entropy (named *impurity* in the leaves of the decision tree). The entropy characterizes the distribution of instances by cluster in a specific node. It is equal to 0 if only elements of one cluster are present in the node, and to $\log_2(\text{\# of clusters})$ if elements are spread equally between all the visual clusters to predict. For each leaf in the tree, the number of instances predicted for each cluster is presented.

The prediction for each instance is shown in a second scatter plot in the upper right part of the interface. Whereas the decision tree gives the number of incorrect predictions in each leaf, this scatter plot makes it possible to identify the incorrectly predicted instances in question. The instances are colored according to their predicted cluster and are shaped as a dot if the prediction is consistent with the user selection, and as a cross otherwise. The level of confidence of the predictions made by the decision tree is denoted by the opacity of the points on the scatter plot. Again, instances can be filtered out according to a threshold defined by the user on the minimal confidence provided by the DT.

Based on the DT and the scatter plot showing the predictions, the user can reflect on the explanations and draw a new cluster selection. In turn, he can cycle through the visual cluster explanation pipeline again by generating a new DT with new selected visual clusters. This iterative process is repeated until the user feels that he has a sufficient understanding of the visualization.

5. Case study examples

In this section, two case studies demonstrate the application of the IXVC pipeline. In a first case study, the data analyst works with 50 instances extracted from the 138 countries of the 2006 Human Development Report [16], hereafter called the *country* dataset. The
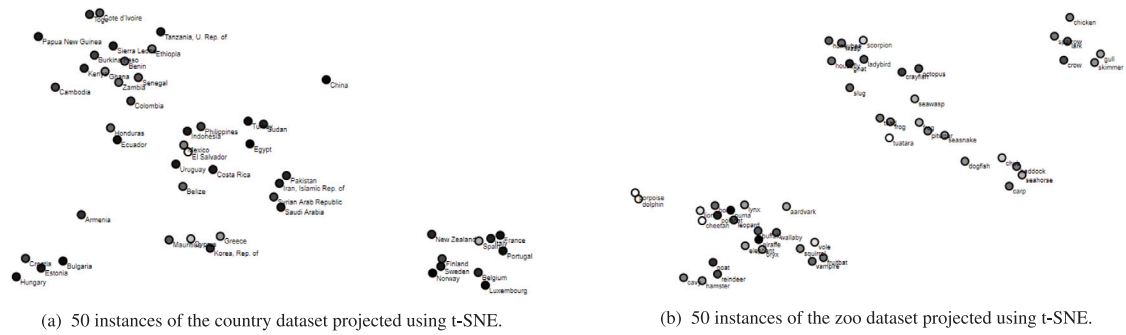
(a) 50 instances of the country dataset projected using t-SNE.

(b) 50 instances of the zoo dataset projected using t-SNE.

Fig. 5. DR visualizations used in our experiments. The luminance of the dots indicates the individual DR errors (the whiter the dots are, the higher the error is).

countries are characterized by 45 socio-economical indicators such as *GDP* and *population growth*.

In the first step, the data analyst discovers the scatter plot shown in Fig. 5a. In the second step, he makes a selection of clusters. In the example of Fig. 6a, three clusters have been selected. In the third step, a decision tree and a second scatter plot are generated based on the cluster selection. The decision tree (Fig. 6c) shows that the red cluster can be flawlessly explained by the *money spent on assisting least developed countries*, with the countries in red spending more. The decision tree separates the 40 remaining instances according to their *GDP*. It also explains the blue cluster as the set of countries above 64.9 billion USD of GDP. Among the 15 concerned instances, 10 are correctly predicted as belonging to the blue cluster. However, 5 instances selected in the green cluster by the user are erroneously predicted as blue. All the remaining instances are predicted to belong to the green cluster. The decision tree uses the *primary exports* feature to separate the 25 remaining instances. Among the countries under 57% of primary export, 2 instances being in the blue cluster are erroneously predicted as green. The scatter plot in Fig. 6b shows the predictions and highlights the errors.

Unsatisfied with the 7 erroneous predictions of the decision tree, the data analyst undertakes a second iteration of the pipeline, setting aside the red cluster, for which there was no prediction error. The analyst divides the remaining 40 instances into two significantly changed clusters (Fig. 7a). In this new selection, the former blue cluster (cluster B in Fig. 6a) has been enlarged to include the countries that were selected in the green cluster before, but predicted as blue by the decision tree, such as Columbia, Mexico, Indonesia and Philippines (see Fig. 6b). The new decision tree (Fig. 7b) is trained based on the new selection of clusters and explains the new clusters using the *GDP per capita* feature. It results in only 2 prediction errors instead of the 7 errors that occurred in the first iteration that used the *GDP* to separate the 40 instances. This new explanation, which would have been tedious, or even impossible, to reach without the IXVC pipeline, leaves the data analyst satisfied with the cluster explanation. The selected clusters can be explained by the *money spent on assisting least developed countries* and the *GDP per capita*. 48 instances out of 50 are correctly predicted by the corresponding decision tree.

In a second case study, a data analyst works with an MDS visualization of the Boston dataset [61]. This dataset is composed of house features describing 506 instances. When dealing with the visualization of these 506 instances, the data analyst notices, before using IXVC, that the compact visualization that MDS often provides is separated into two visual clusters (see Fig. 8a). In order to verify that these clusters can be explained with the HD features that were used to build the visualization, the data analyst relies on IXVC. IXVC shows to the data analyst that the DT built to explain the top and bottom visual clusters makes many errors. Therefore, the analyst decides to discard the instances that have a too high projection error and to split the top cluster into two clusters. One of these two new clusters consists of instances in the dense center and constitutes the blue cluster. The other

cluster groups the instances that gravitate around the dense center and forms the green cluster. After having refined the visual clusters as such, the data analyst learns that, indeed, this difficult-to-understand visualization is the representation of different clusters in the HD space. The data analyst conclusion enlightened by the decision tree is that the red cluster corresponds to the neighborhoods with a high rate of criminality, by opposition to the neighborhoods in the blue and green clusters. The blue cluster corresponds to neighborhoods with a low rate of criminality, with a percentage of the land which is residential lower than 57.5%, with few industries nearby, and with houses having an average number of rooms strictly lower than 8. In other words, the blue cluster correspond to city centers with residences and rental businesses. The green cluster contains instances that are variations of the ones in the blue cluster. Indeed, any neighborhood being more residential, with more industries or with bigger dwellings is in the green cluster.

6. Evaluation

This section presents the evaluation of the IXVC pipeline and interface. Note that the datasets that are relevant with the pipeline need to contain understandable features in order to be used with decision trees. For evaluation purposes, two datasets are available for analysis with the interface. First, the *country* dataset presented in Section 5 (see the DR visualization of the dataset used in our experiment in Fig. 5a). Second, the *zoo* dataset [61] characterizes 101 animals with 16 features such as the number of legs and whether they have feathers or not (see the DR visualization of the dataset used in our experiment in Fig. 5b). A table displaying the whole dataset under scrutiny is available to users via a button on the interface. The *t*-SNE perplexities for generating the DR visualizations of the *country* and *zoo* datasets are 6 and 18 respectively. Each time a DT is built to explain visual clusters, the value of the hyperparameter *minimum samples per leaf* is chosen among the 40 values in the range [10, 50[by a 10-fold cross-validation.

For each dataset, 50 instances were randomly sampled. The goal of the evaluation was to evaluate the IXVC pipeline rather than the interface developed to implement it. The participants would have been confronted to a barrier not related to the pipeline, which would have thus tweaked the evaluation results. Showing more than 50 instances at once would hinder the readability of the scatter plot and sampling is one technique commonly suggested to tackle such visual clutter issues [59].

6.1. Preliminary feedback

During the development of the IXVC interface, early feedback has been sought from two researchers that are non-experts, but knowledgeable, in machine learning and information visualization. Their profile was sought to fit a profile of users who would have enough knowledge to use the tool, but not enough to compensate the potential flaws of the interface with their knowledge. The goal of this preliminary feedback was to detect usability flaws in the interface before the evaluation.
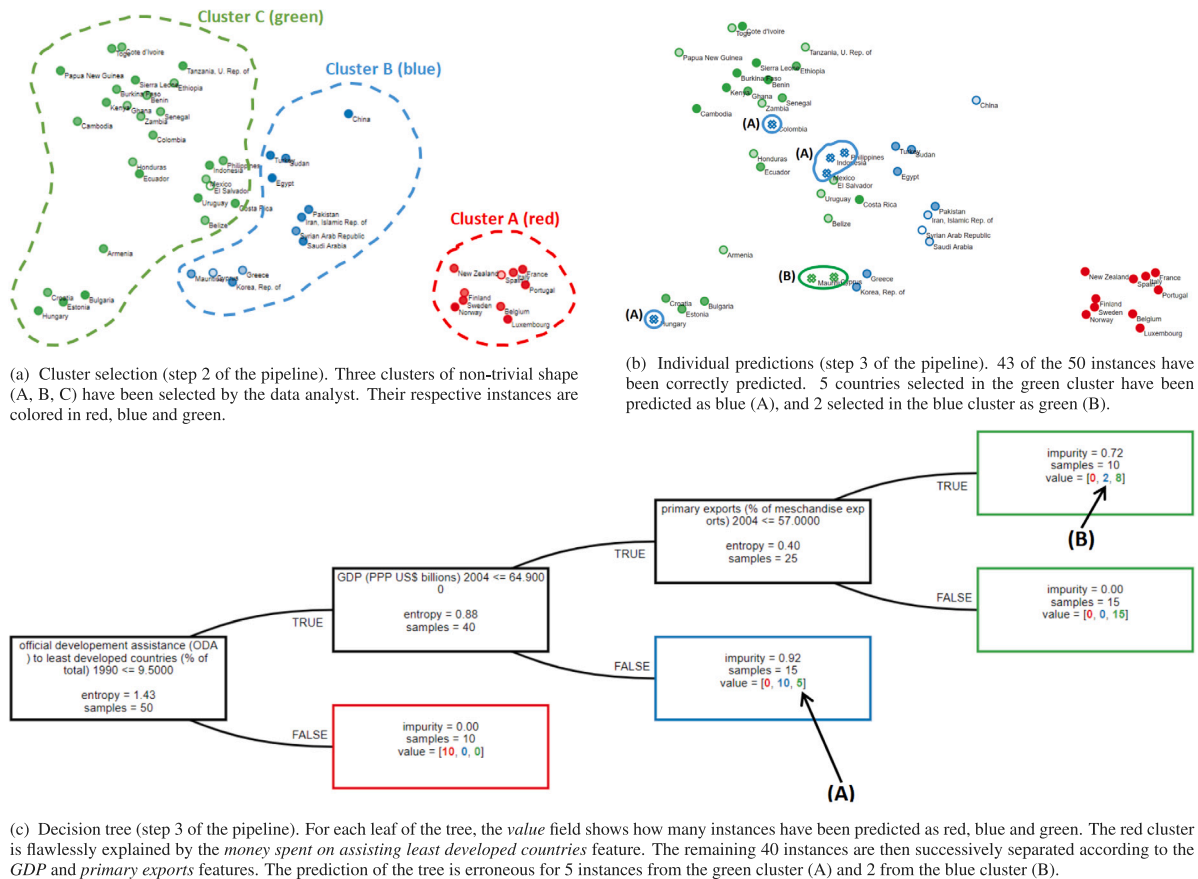
(a) Cluster selection (step 2 of the pipeline). Three clusters of non-trivial shape (A, B, C) have been selected by the data analyst. Their respective instances are colored in red, blue and green.

(b) Individual predictions (step 3 of the pipeline). 43 of the 50 instances have been correctly predicted. 5 countries selected in the green cluster have been predicted as blue (A), and 2 selected in the blue cluster as green (B).

(c) Decision tree (step 3 of the pipeline). For each leaf of the tree, the *value* field shows how many instances have been predicted as red, blue and green. The red cluster is flawlessly explained by the *money spent on assisting least developed countries* feature. The remaining 40 instances are then successively separated according to the *GDP* and *primary exports* features. The prediction of the tree is erroneous for 5 instances from the green cluster (A) and 2 from the blue cluster (B).

Fig. 6. Case study example 1: first iteration of the IXVC pipeline to explain the visual clusters of countries. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
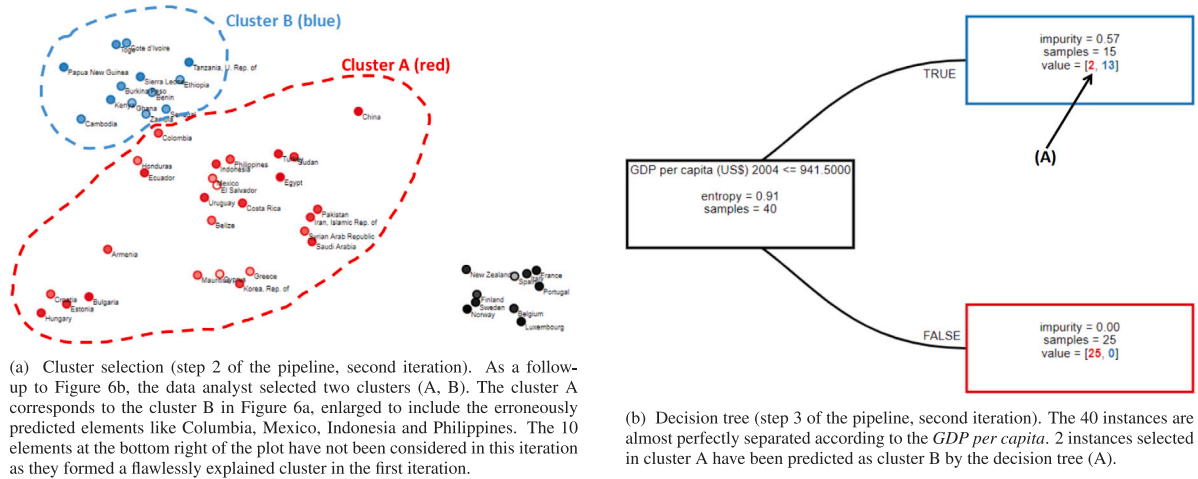

(a) Cluster selection (step 2 of the pipeline, second iteration). As a follow-up to Figure 6b, the data analyst selected two clusters (A, B). The cluster A corresponds to the cluster B in Figure 6a, enlarged to include the erroneously predicted elements like Columbia, Mexico, Indonesia and Philippines. The 10 elements at the bottom right of the plot have not been considered in this iteration as they formed a flawlessly explained cluster in the first iteration.

(b) Decision tree (step 3 of the pipeline, second iteration). The 40 instances are almost perfectly separated according to the *GDP per capita*. 2 instances selected in cluster A have been predicted as cluster B by the decision tree (A).

Fig. 7. Case study example 1: second iteration of the IXVC pipeline to explain the visual clusters of countries. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Overall, the two researchers made a few suggestions such as adding captions and changing labels in order to make the interface clearer. For example, the "group" label was previously "cluster". The two participants explained that "group" would be a clearer term, as using "cluster" could wrongly indicate that IXVC conducts an automated cluster analysis. The term "error" was also suggested by the participants as a replacement for the more technical term of "loss" that was previously used. The two participants also made suggestions regarding the color encoding of the selected groups, which led to the color scale used in the evaluation.

6.2. Evaluation methodology

The objective of the evaluation is to measure whether IXVC helps to conduct an analysis of a DR visualization, and if so, with more objectivity. As a tool supporting analysis through a DR visualization and decision trees, IXVC is destined to users knowledgeable of these techniques. 16 students (13 males and 3 females) following a graduate-level data science program in which DR and decision trees are taught were recruited. The age of the participants ranged from 20 to 53 years (two participants are older students resuming their studies), with a
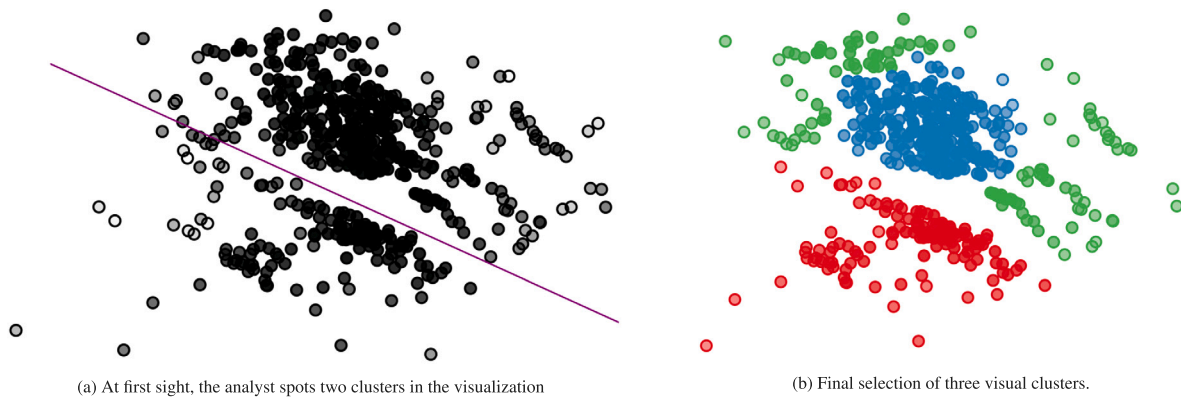
(a) At first sight, the analyst spots two clusters in the visualization

(b) Final selection of three visual clusters.

Fig. 8. Case study example 2: separation of a difficult-to-understand MDS visualization of the Boston dataset into three clusters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

median of 22. The students had previously carried out a class project on intuitive visual cluster explanation (without any tool) with the *country* dataset.

The evaluation consisted of 45-minute sessions following quasi-empirical evaluation practices [62]. The sessions began with a brief introduction to the goal of the IXVC pipeline. No explanation on how the interface works was provided at this point. Two researchers were present throughout the session in order to answer participants' questions and to take note of their remarks as well as observations. Then, the two datasets participants were asked to work on were introduced. The *country* dataset was presented as a set of countries characterized by various socio-economical indicators and the *zoo* dataset was presented as a set of animals described by biological traits. It was essential to provide only the minimum, but necessary, information in order not to guide the explanations towards specific HD features. Observation and questionnaire filling were used to collect data.

6.2.1. Observations

Observations were conducted throughout the sessions by two researchers to detect usability issues and to see whether the analysis behavior of the participants was consistent with the pipeline. Observations were mainly passive with questions from participants answered when asked.

6.2.2. Questionnaire

When participants were finished with the analysis of the two datasets, they were invited to fill a short three-part questionnaire. First, the initial perceived expertise of the participants was measured. The second part of the questionnaire was about the data analysis process with IXVC. Lastly, the general usability of the IXVC interface was measured in order to control the impact of the interface in the evaluation of the pipeline. For this latter part, the System Usability Scale (SUS) questionnaire [63] was used. The SUS is a questionnaire scoring the usability of a system with 10 questions measured on a 5-point Likert scale. It has the advantage of being quick to complete and highly reliable. Following literature recommendations, two adaptations were made to the original SUS. In the *eighth item of the SUS*, "cumbersome" was replaced by "awkward", as the participants are non-native English speakers. The word "cumbersome" in the SUS has been reported to cause confusion [64], especially among non-native English speakers [65]. The *first item of the SUS* measures the extent to which users would like to use a system frequently. Since the participants of the evaluation are students, who only perform cluster explanation in the context of a class, the first item of the SUS formulated as such was not relevant, and would have tweaked the SUS score. Instead, an adapted version of this question was included in the questionnaire, formulated as "I would like to use the tool in the future if I need to analyze a visualization generated by t-SNE". Although the answers to this question are of interest to the

evaluation, it was not included in the computation of the SUS score in order to preserve its reliability. Rather, the score was computed from the 9 other questions, as Lewis and Sauro [66] showed that removing an item inducts a negligible deviation from the results of the 10-item scale and has no impact on the reliability of the score.

6.3. Evaluation results

This section presents the results of the user evaluation. The results obtained from the observations and from the answers to the questionnaire are successively discussed.

6.3.1. Observations

Most participants intuitively followed the pipeline to get explanations about visual clusters. However, some usages of the IXVC interface that differed from how participants were expected to apply the pipeline were observed. First, the analysis process was more exploratory than expected. The pipeline describes an iterative process in which a selection of clusters is refined by adjusting the manual clustering. However, a few participants tended to try many different cluster selections instead of iteratively refining one.

Second, one participant (P11) was attempting to generate a decision tree involving an intuitive feature not necessarily present among the HD features. Indeed, in the scatter plot of the zoo dataset, the participant selected clusters separating aquatic animals from another. At this point, P11 did not consult the dataset table to see if there was indeed a feature distinguishing aquatic animals. In doing so, P11 tried to build a decision tree where this feature appears. P11 made repeated attempts until such a tree was displayed on the screen.

6.3.2. Questionnaire

Overall, participants rated, on a scale from 1 to 5, their knowledge of *t*-SNE as intermediate (median = 3) and of decision trees as good (median = 4). They felt familiar with the dataset on countries (median = 4), but not with the one on animals (median = 2). In both cases, access to the dataset tables during the analysis was useful to the participants (median = 4 and 4.5 for countries and animals).

75% of the participants stated that they prefer the IXVC interface to the tool-free approach they used in the prior class project. Furthermore, 81% reported they would like to use the IXVC interface again if they have to work on a *t*-SNE generated visualization in the future (see Fig. 9a). Participants felt that the interface helped them to gain a better understanding of the datasets (median = 4 (agree) for both datasets, see Fig. 9b). The IXVC interface scored 3.5 (between neutral and agree) as support to a more objective analysis for both datasets (see Fig. 9c).

The SUS score of IXVC is 77 (95% confidence interval is [72,82]), which is above the "good usability" threshold defined by Bangor et al. [67] at 71.4. Fig. 10 shows the Likert distribution for the 9 items
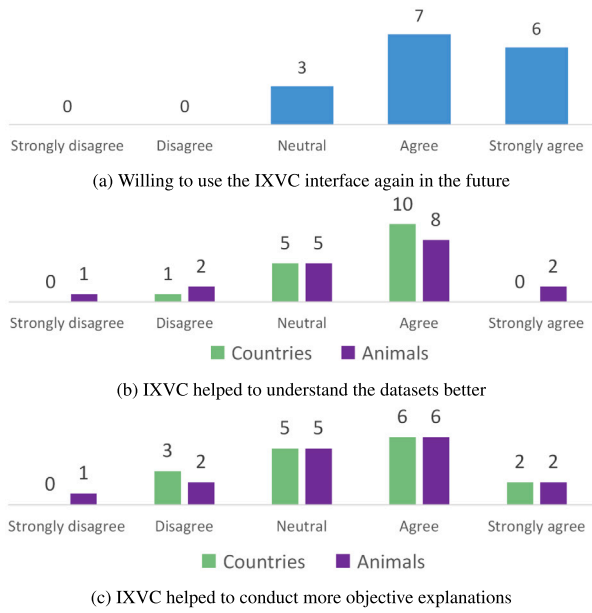
(a) Willing to use the IXVC interface again in the future

(b) IXVC helped to understand the datasets better

(c) IXVC helped to conduct more objective explanations

Fig. 9. Distribution of the answers from the 16 participants.
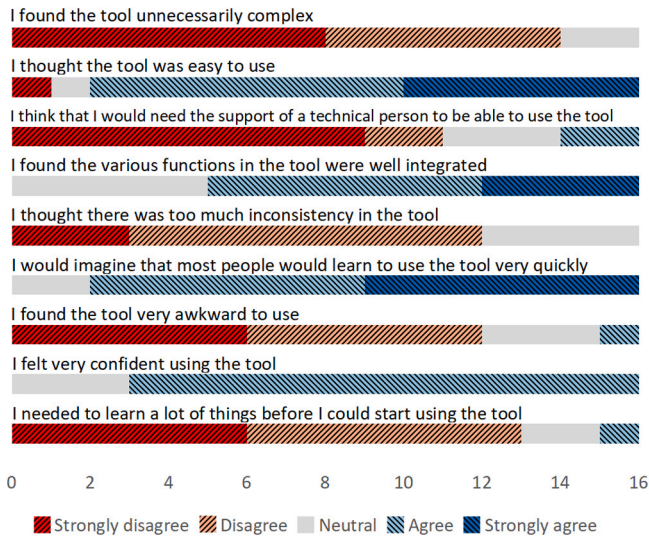

Fig. 10. Likert distribution for the 9 items used to compute the SUS score.

used to compute the SUS score. Lewis and Sauro [68] showed that the data gathered through the SUS can also be used to reliably derive a learnability score that can be interpreted in the same way as the SUS score. It measures the extent to which an interface enables its users to learn how to use it. The learnability score is computed by considering the fourth and tenth items of the questionnaire. For the IXVC interface, the score stands at 78, very close to the SUS score, which indicates that it has a good learnability.

6.4. Discussion

The first observation from the experimental results is that the interface did not alter the evaluation of the pipeline. Indeed, with a SUS score of 77, the interface usability has been considered good, meaning that the implementation has not, for the most part, interfered with the evaluation of the pipeline.

Considering the pipeline itself, it has been considered more useful than the intuitive analysis performed without it by 75% of the participants. This indicates that providing information on the explanation of the selected clusters is important when explaining DR visualizations. Moreover, the participants showed great enthusiasm towards the explanations given by the decision tree and the interactivity of the IXVC interface. In the questionnaire, P1 wrote that "the decision tree is great to visualize how the different groups can be divided and what differentiates them the most". P9 wrote "the real added value is in the decision tree. I think it is very valuable to have an objective reason for clusters".

However, while some participants felt that IXVC brings added-value, the question "I feel that the tool helped me to conduct a more objective analysis" scored between neutral and agree. In an open field of the questionnaire, one participant wrote that "the analysis, in my point of view, is not more objective since there is a great part of [intuitivity] when choosing the clusters". It seems that some participants interpreted the question as "is the whole process objective" and did not understand that the objectivity resides in the explanations given by the decision tree.

Finally, although the familiarity of the participants was quite different between the two datasets, the same results were observed for both. This leads us to conclude that IXVC is beneficial to the DR explainability process, irrespective of the prior knowledge of the data at hand.

7. Limitations of IXVC

In order to explain visual clusters, IXVC uses DTs. These models are interesting because they provide a nonlinear, though interpretable, decision boundary. However, if the visual clusters cannot be easily explained by a decision tree kind of decision boundary, the DT can ultimately be large and uninterpretable. One solution to this limitation would be to integrate other techniques from the literature into the pipeline, such that if the DT is too large to explain the visual clusters, other useful information (e.g., from ccPCA [45]) can be provided.

Another limitation of IXVC is the fact that the convergence of the pipeline (i.e. being satisfied by the explanation of the visual clusters) is rooted in the expertise of the user. This expertise is domain-dependent and is therefore difficult to define as a prerequisite for the use of the pipeline.

8. Future work

The evaluation results and open fields in the questionnaire allow considering future directions to improve IXVC. A first future work is the development of IXVC for making it an educational tool. As mentioned in Section 6, IXVC is destined to users knowledgeable on DR and decision trees. However, several participants pointed out that the use of IXVC, in fact, needs little knowledge of these techniques. Since the participants are students following a data science program, they are eager to use tools that may ease the understanding of techniques such as *t*-SNE. Moreover, the playful character of IXVC was emphasized by one participant.

Another future work can be identified following the observation of P11 described in Section 6.3.1. P11 used IXVC for checking if a feature he had in mind played a role in the visual cluster separation. Based on this use, the IXVC interface could propose all features as clickable buttons, which would show how the selected feature would make it possible to separate the visual clusters. In the IXVC pipeline, this corresponds to providing all possible decision trees with only one decision node. This feature can also be added before the use of decision trees, by coloring the instances in the visualization according to a selected HD feature. This pre-exploration step can indicate to the analyst how to manually cluster the 2D instances in step 2 of IXVC.

Participant P16 suggested to generate several decision trees in order to have the opportunity to consider different possible explanations. This

means that, in another future work, several trees could be suggested in step 3 of IXVC. For instance, after building the first decision tree, a second one could be built by removing, from the possible features to choose for a decision, the feature that is chosen in the first node of the tree.

Finally, an analysis of the use of IXVC to compare DR techniques can be performed. Using, e.g., synthetic datasets, one would be able to check how hard it is to explain visualizations from particular DR techniques based on different conditions. Such a condition can be that the DR technique projects distinct visual clusters that are in fact controlled to be intertwined in HD, or the opposite.

9. Conclusion

In this paper, we proposed an interactive machine learning pipeline called IXVC (for *Interactive eXplanation of Visual Clusters*). The pipeline provides explanations of visual clusters manually selected by a data analyst in a dimensionality reduction (DR) visualization. The explanatory feedback on the manually selected clusters is provided by a decision tree whose decisions are based on the high-dimensional (HD) features. Interactively, the data analyst can thus select clusters in the visualization and receive an explanation of the selected clusters through a decision tree. IXVC is a need for data analysts [29] and handles a task that is called *discover relation between visual pattern and original dimensions* [11].

IXVC was implemented as a web application for its evaluation. The results of the evaluation suggest that using the proposed interactive pipeline helps users to explain how visual clusters in a DR visualization are related to the HD features that have been used to create the visualization, even when the mapping between the high and the low dimensions is not provided. It is also suggested by the evaluation results that the usefulness of the pipeline does not depend on the prior knowledge the analyst has on the dataset.

CRediT authorship contribution statement

Adrien Bibal: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Antoine Clarinval:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Bruno Dumas:** Writing – review & editing, Supervision. **Benoît Frénay:** Writing – review & editing, Supervision.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. Antoine Clarinval, co-first author of this work, is funded by the European Regional Development Fund (ERDF) through the Wal-e-Cities LIV project with award number [ETR121200003138].

Acknowledgments

References

[1] Kruskal JB, Wish M. Multidimensional scaling. Sage; 1978.

[2] Hout MC, Papesh MH, Goldinger SD. Multidimensional scaling. Wiley Interdiscip Rev: Cogn Sci 2013;4(1):93–103.

[3] Jaworska N, Chupetlovska-Anastasova A. A review of multidimensional scaling (MDS) and its utility in various psychological domains. Tutor Quant Methods Psychol 2009;5(1):1–10.

[4] Koch A, Imhoff R, Dotsch R, Unkelbach C, Alves H. The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. J Personal Soc Psychol 2016;110(5):675–709.

[5] Bishop C. Pattern recognition and machine learning. New York: Springer–Verlag; 2006.

[6] Bibal A, Frénay B. Interpretability of machine learning models and representations: an introduction. In: Proceedings of the European symposium on artificial neural networks (ESANN). 2016, p. 77–82.

[7] Lipton ZC. The mythos of model interpretability. Queue 2018;16(3):31–57.

[8] Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv 2018;51(5):1–42.

[9] van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9:2579–605.

[10] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.

[11] Nonato LG, Aupetit M. Multidimensional projection for visual analytics: Linking techniques with distortions, tasks, and layout enrichment. IEEE Trans Vis Comput Graphics 2018;25(8):2650–73.

[12] Aupetit M. Visualizing distortions and recovering topology in continuous projection techniques. Neurocomputing 2007;70(7–9):1304–30.

[13] Lespinats S, Aupetit M. Checkviz: Sanity check and topological clues for linear and non-linear mappings. Comput Graph Forum 2011;30(1):113–25.

[14] Schreck T, Von Landesberger T, Bremm S. Techniques for precision-based visual analysis of projected data. Inf Vis 2010;9(3):181–93.

[15] Martins RM, Coimbra DB, Minghim R, Telea AC. Visual analysis of dimensionality reduction quality for parameterized projections. Comput Graph 2014;41:26–42.

[16] United Nations Development Programme. Human development report. 2006.

[17] Broeksema B, Telea AC, Baudel T. Visual analysis of multi-dimensional categorical data sets. Comput Graph Forum 2013;32(8):158–69.

[18] Gower JC, Hand DJ. Biplots, Vol. 54. CRC Press; 1995.

[19] Greenacre MJ. Biplots in practice. Fundacion BBVA; 2010.

[20] Liu S, Maljovec D, Wang B, Bremer P-T, Pascucci V. Visualizing high-dimensional data: Advances in the past decade. IEEE Trans Vis Comput Graph 2017;23(3):1249–68.

[21] Coimbra DB, Martins RM, Neves TT, Telea AC, Paulovich FV. Explaining three-dimensional dimensionality reduction plots. Inf Vis 2016;15(2):154–72.

[22] Cavallo M, Demiralp Ç. A visual interaction framework for dimensionality reduction based data exploration. In: Proceedings of the conference on human factors in computing systems. 2018, p. 635:1–635:13.

[23] Turkay C, Filzmoser P, Hauser H. Brushing dimensions-a dual visual analysis model for high-dimensional data. IEEE Trans Vis Comput Graph 2011;17(12):2591–9.

[24] Yuan X, Ren D, Wang Z, Guo C. Dimension projection matrix/tree: Interactive subspace visual exploration and analysis of high dimensional data. IEEE Trans Vis Comput Graph 2013;19(12):2625–33.

[25] Gisbrecht A, Mokbel B, Hammer B. Linear basis-function t-SNE for fast nonlinear dimensionality reduction. In: Proceedings of the international joint conference on neural networks. 2012, p. 1–8.

[26] Chang JJ, Carroll JD. How to use PROFIT, a computer program for property fitting by optimizing nonlinear or linear correlation. Bell Laboratories; 1968, Unpublished Manuscript.

[27] Bibal A, Marion R, Frénay B. Finding the most interpretable MDS rotation for sparse linear models based on external features. In: Proceedings of the European symposium on artificial neural networks, computational intelligence and machine learning. 2018, p. 537–542.

[28] Marion R, Bibal A, Frénay B. BIR: A method for selecting the best interpretable multidimensional scaling rotation using external variables. Neurocomputing 2019;342:83–96.

[29] Brehmer M, Sedlmair M, Ingram S, Munzner T. Visualizing dimensionally-reduced data: Interviews with analysts and a characterization of task sequences. In: Proceedings of the workshop on beyond time and errors: Novel evaluation methods for visualization. 2014, p. 1–8.

[30] Wenskovitch J, Crandell I, Ramakrishnan N, House L, North C. Towards a systematic combination of dimension reduction and clustering in visual analytics. IEEE Trans Vis Comput Graphics 2017;24(1):131–41.

[31] Cavallo M, Demiralp Ç. Clustrophile 2: Guided visual clustering analysis. IEEE Trans Vis Comput Graphics 2018;25(1):267–76.

[32] Lebel A, Cantinotti M, Pampalon R, Thériault M, Smith LA, Hamelin A-M. Concept mapping of diet and physical activity: uncovering local stakeholders perception in the Quebec City region. Soc Sci Med 2011;72(3):439–45.

[33] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Mach Learn 2002;46(1–3):389–422.

[34] Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3(Mar):1157–82.

[35] Xing EP, Karp RM. CLIFF: clustering of high-dimensional microarray data via iterative feature filtering using normalized cuts. Bioinformatics 2001;17(suppl_1):S306–15.

[36] Mitra P, Murthy C, Pal SK. Unsupervised feature selection using feature similarity. IEEE Trans Pattern Anal Mach Intell 2002;24(3):301–12.

[37] Dy JG, Brodley CE. Feature selection for unsupervised learning. J Mach Learn Res 2004;5(Aug):845–89.

[38] Cai D, Zhang C, He X. Unsupervised feature selection for multi-cluster data. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining. 2010, p. 333–42.

[39] da Silva RRO, Rauber PE, Martins RM, Minghim R, Telea AC. Attribute-based visual explanation of multidimensional projections. In: Proceedings of EuroVA. 2015, p. 134–9.

[40] Kandogan E. Just-in-time annotation of clusters, outliers, and trends in point-based data visualizations. In: Proceedings of the IEEE conference on visual analytics science and technology (VAST). 2012, p. 73–82.

[41] Joia P, Petronetto F, Nonato LG. Uncovering representative groups in multidimensional projections. Comput Graph Forum 2015;34(3):281–90.

[42] Rauber PE, da Silva RRO, Feringa S, Celebi ME, Falcão AX, Telea AC. Interactive image feature selection aided by dimensionality reduction. In: Proceedings of EuroVA. 2015, p. 54–61.

[43] Parisot O, Ghoniem M, Otjacques B. Decision trees and data preprocessing to help clustering interpretation. In: Proceedings of the international conference on data management technologies and applications. 2014, p. 48–55.

[44] van Ham F, Petitclerc M, Pisters R. Guiding multidimensional analysis using decision trees. In: Proceedings of the conference of the center for advanced studies on collaborative research. 2013, p. 200–14.

[45] Fujiwara T, Kwon O-H, Ma K-L. Supporting analysis of dimensionality reduction results with contrastive learning. IEEE Trans Vis Comput Graphics 2019;26(1):45–55.

[46] Chatzimparmpas A, Martins RM, Kerren A. T-viSNE: Interactive assessment and interpretation of t-SNE projections. IEEE Trans Vis Comput Graph 2020;26(8):2696–714.

[47] Ma Y, Maciejewski R. Visual analysis of class separations with locally linear segments. IEEE Trans Vis Comput Graphics 2020;27(1):241–53.

[48] Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?" Explaining the predictions of any classifier. In: Proceedings of ACM SIGKDD. 2016, p. 1135–44.

[49] Ware M, Frank E, Holmes G, Hall M, Witten IH. Interactive machine learning: letting users build classifiers. Int J Hum–Comput Stud 2001;55(3):281–92.

[50] Wattenberg M, Viégas F, Johnson I. How to use t-SNE effectively. Distill 2016. http://dx.doi.org/10.23915/distill.00002, URL http://distill.pub/2016/misread-tsne.

[51] McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. 2018, arXiv preprint arXiv:1802.03426.

[52] Tang J, Liu J, Zhang M, Mei Q. Visualizing large-scale and high-dimensional data. In: Proceedings of the international conference on world wide web. 2016, p. 287–97.

[53] Huang Z. Clustering large data sets with mixed numeric and categorical values. In: Proceedings of the pacific-asia conference on knowledge discovery and data mining. 1997, p. 21–34.

[54] Qiu M, Davis S, Ikem F. Evaluation of clustering techniques in data mining tools. Issues Inf Syst 2004;5(1).

[55] Sacha D, Zhang L, Sedlmair M, Lee JA, Peltonen J, Weiskopf D, North SC, Keim DA. Visual interaction with dimensionality reduction: A structured literature analysis. IEEE Trans Vis Comput Graph 2016;23(1):241–50.

[56] Bostock M, Ogievetsky V, Heer J. D3 data-driven documents. IEEE Trans Vis Comput Graphics 2011;17(12):2301–9.

[57] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12(Oct):2825–30.

[58] Mayorga A, Gleicher M. Splatterplots: Overcoming overdraw in scatter plots. IEEE Trans Vis Comput Graphics 2013;19(9):1526–38.

[59] Ellis G, Dix A. A taxonomy of clutter reduction for information visualisation. IEEE Trans Vis Comput Graphics 2007;13(6):1216–23.

[60] Harrower M, Brewer CA. ColorBrewer.org: an online tool for selecting colour schemes for maps. Cartogr J 2003;40(1):27–37.

[61] Dheeru D, Karra Taniskidou E. UCI machine learning repository. University of California, Irvine, School of Information and Computer Sciences; 2017, URL http://archive.ics.uci.edu/ml.

[62] Hartson R, Pyla PS. The UX book: Process and guidelines for ensuring a quality user experience. Elsevier; 2012.

[63] Brooke J. SUS – A quick and dirty usability scale. In: Usability evaluation in industry, Vol. 189. 1996, p. 4–7, (194).

[64] Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum–Comput Interact 2008;24(6):574–94.

[65] Finstad K. The system usability scale and non-native English speakers. J Usability Stud 2006;1(4):185–8.

[66] Lewis JR, Sauro J. Can I leave this one out?: the effect of dropping an item from the SUS. J Usability Stud 2017;13(1):38–46.

[67] Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. J Usability Stud 2009;4(3):114–23.

[68] Lewis JR, Sauro J. The factor structure of the system usability scale. In: International conference on human centered design. Springer; 2009, p. 94–103.

Adrien Bibal is a postdoctoral researcher at the University of Namur (Belgium). He received an M.S. degree in Computer Science and an M.A. degree in Philosophy from the Université catholique de Louvain (Belgium) in 2013 and 2015 respectively. His Ph.D. thesis in machine learning, completed in 2020 at the University of Namur (Belgium), was on the interpretability and explainability of dimensionality reduction mappings.

Antoine Clarinval received a master degree in computer science in 2017 from the University of Namur, Belgium. He is currently pursuing the PhD degree at the University of Namur. His research interests include smart city education, citizen participation in smart cities, and how public displays and information visualization can support it. In this regard, he is especially interested in traffic data visualization and open data.

Bruno Dumas received his PhD in 2010 from the University of Fribourg, Switzerland. His PhD thesis focused on the creation of multimodal interfaces, following three axes: software architectures, modeling languages and multimodal fusion algorithms. He then worked for three and a half years at the Vrije Universiteit Brussel as a post-doc. His research areas focus on human-machine interaction, multimodal interfaces and more broadly on how the expansion of computing in everyday life influences usage.

Benoît Frénay is associate professor at the Université de Namur. He received his Ph.D. degree from the Université catholique de Louvain (Belgium) in 2013. His main research interests in machine learning include interpretability, interactive machine learning, dimensionality reduction, label noise, robust inference and feature selection. In 2014, he received the Scientific Prize IBM Belgium for Informatics for his PhD thesis on Uncertainty and Label Noise in Machine Learning.