Lossless Visual Knowledge Discovery in High Dimensional Data with Elliptic Paired Coordinates

Rose McDonald
Dept. of Computer Science
Central Washington University
Ellensburg, WA 98926, USA
Rose.McDonald@cwu.edu

Boris Kovalerchuk
Dept. of Computer Science
Central Washington University
Ellensburg, WA 98926, USA
BorisK@cwu.edu

***Abstract*—Data with more than two or three dimensions are difficult for humans to conceptualize and facilitate knowledge discovery. Novel Elliptic Paired Coordinates (EPCs) allow for multidimensional data to be represented in 2-D without loss of multidimensional information. In addition, EPC halves the required visual elements in the graph in comparison with parallel and radial coordinates. This research explores the effectiveness of constructing predictive machine learning models interactively using EPC visualizations. For this research EllipseVis, an interactive software system, was developed to process high-dimensional datasets, create corresponding EPC visualizations, and build predictive classification models based on dominance rules. The EllipseVis system allows both interactive and automatic discovery of areas that are located with a high percentage of single-class dominance. The experiments using it on benchmark datasets suggest EPC approach is a promising method for discovering predictive models with high coverage and precision that could be useful in many fields allowing for visually appealing dominance rules to be easily interpreted in the application domains.**

Keywords*—Machine learning, data visualization, knowledge discovery, elliptic paired coordinates.

I. INTRODUCTION

Analysis of large and complex multidimensional data benefits from data visualization due opportunity to use unique human perceptual abilities to explore visuals. However, the efficient use of visualization in Machine Learning (ML) is difficult. ML exploration is fundamentally multidimensional, but quite commonly visualization methods do not preserve all n-D information or represent it in the form that is beyond human perceptual abilities [1]. The purpose of this research is conducting experimental studies on datasets using novel **Elliptic Paired Coordinates (EPCs)** [1] and evaluate their effectiveness for interactive and automatic visual knowledge discovery and machine learning. Several current studies attempt to deal with this challenge in different ways. One of the common ways is mapping *n-D points to 2-D points* preserving some n-D information such as n-D distances/similarities using MDS, SOM, t-SNE, PCA and other unsupervised methods [6]. Another approach is *mapping n-D points to 2-D graphs* [1,2,5]. Then both approaches attempt to solve predictive classification tasks in the respective visual representation of n-D data [7]. This paper develops a graph approach with EPC. The selection of EPC of this exploration is based on three considerations: (1) it *preserves* all n-D information, (2) it can capture *non-linear dependencies* in the data being non-linear data transform, and (3) it is more *compact* (requires less visual elements) than methods such as parallel and radial coordinates.

II. ELLIPTIC PAIRED COORDINATES

A. Concept

The EPC belong to the class *curvilinear coordinates*, where all coordinate axes are located on ellipses as shown in Fig. 1. We present and use a version of EPC denoted as EPC-H in [1]. Below for shorter notation we omit H. This EPC uses two times less edges and nodes than parallel and radial coordinates leading to less clutter. Fig. 1 shows an example of 4-D point P = (0.3,0.5,0.5,0.2) in EPC as a short green arrow. For comparison, Fig. 2 shows the same point P in well-known radial and parallel coordinates that present it as graphs with two times more nodes (4 nodes). In Fig. 1, the blue ellipse C_E holds four coordinate curves X_1-X_4.

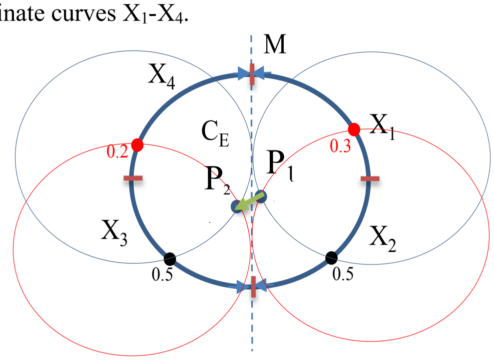


Fig. 1. 4-D point P = (0.3,0.5,0.5,0.2) in 4-D EPC as green arrow $P_1 \rightarrow P_2$. Red marks separate coordinates in the blue coordinate ellipse.

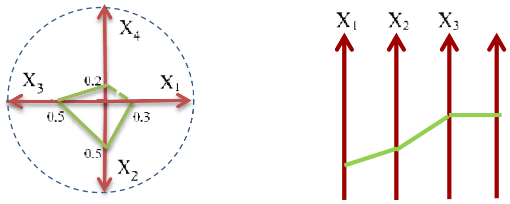
The green arrow ($P_1 \rightarrow P_2$) is constructed by using 4 ellipses of the size of the blue coordinate ellipse. These four ellipses touch the middle vertical line M and move along it to get different 4-D points. Below we explain this process. The thin red ellipse on the right is defined by the three requirements: (1) to go through $x_1 = 0.3$, (2) to touch line M and (3) be of the size of the blue coordinate ellipse C_E. Visually it is done by copying C_E and shifting the copy to the right to touch line M and then sliding it along M to reach $x_1 = 0.3$ on coordinate X_1.

While there two such ellipses, the control of its up and down move allows to select one that will cross a blue ellipse described next. The same process is conducted for $x_2 = 0.5$ with the thin blue ellipse. The point P_1 where these red and blue ellipses cross each other in C_E is the point that represents pair (x_1,x_2) = (0.3,0.5). Next this process is repeated for $x_3 = 0.5$ and $x_4 = 0.2$ to produce the crossing point P_2 of two respective ellipses on the left that represents pair (x_3,x_4) = (0.5,0.2). Then the two crossing points P_1 and P_2 are connected to form the short green

This research was partially funded by CWU College of the Sciences Summer Undergraduate Research Experience (COTS-SURE).

978-1-7281-9134-8/20/$31.00 ©2020 IEEE
DOI 10.1109/IV51561.2020.00053

arrow that serves as lossless visualization of 4-D point P = (0.3,0.5,0.5,0.2).



(a) (b)

Fig. 2. 4-D point P = (0.3,0.5,0.5,0.2) in radial (a) and parallel coordinates (b).

This process is reversible allowing restoring the 4-D point P from the ($P_1 \rightarrow P_2$) arrow. Visually it is done sliding the thin red ellipse on the right to cross point P_1. The point where this ellipse crosses C_E ellipse gives the point on X_1 coordinate, which is x_1 = 0.3 The same process is conducted for the blue circle on the right to find its crossing with X_2 coordinate to find x_2 = 0.5. We repeat it for P_2, thin red and blue ellipses on the left.

B. Elliptic Paired Coordinates Algorithm

The EPC algorithm for n-D data has the following steps:

1. Create the central ellipse, C_E, with a vertical line M bisecting it.
2. Divide the circumference of C_E into sectors by the number of dimensions $X_1, X_2, \ldots, X_n$ to be represented.
3. Normalize the data on the scale [0,1].
4. Graph the data points x_i along the circumference of the C_E in the appropriate sector for each data point.
5. Pair data points by dimension, starting at (x_1,x_2). If there is an odd number of dimension pairs, use the horizontal bisector N instead.
6. Create the thin red and blue side ellipses, such that they touch the line M. Have the top of red ellipses intersect the data point and the bottom of blue ellipses, alternating between red and blue each dimension in sequence.
7. Find the intersects between each pair, connect them.

III. PC-Based Visual Knowledge Discovery System

A. Dominance Rectangular Rules Algorithm in EPC

The next process after n-D data points are visualized in EPC losslessly is finding the rectangular areas where a single class dominates other classes. Ideally these areas should contain only cases of a single class. This search can be conducted interactively or automatically.

Let R is a rectangle then the rule r can have two forms:
Point Rule r: If a node of graph **x*** of n-D point **x** is *in* R then **x** is in class A.
Intersect Rule r: If graph **x*** of n-D point **x** *intersects* R then **x** is in class A.

An intersect rule captures a non-linear relation of 4 attributes that form a line that crosses R, and a point-based rule captures a non-linear relation of 2 attributes (node) in R.

In the experiments below we discover both types of rules. Steps of the **algorithm**:

1. Visualize data in EPC;
2. Set up parameters of the dominant rectangles: coverage, precision thresholds and type (intersect or point);
3. Search for dominant rectangles;

3.1. Automatic search: set up size of the rectangle, shift rectangle with a given step over the EPC display, compute coverage and precision for each rectangle, record rectangles that satisfy thresholds.;

3.2. Interactive search: draw a rectangle of any size and at any location n in the EPC display, collect coverage and precision for this rectangle;

4. Remove/hide cases that are in the accepted rules and continue search for new rules if desired.

The automatic search has *dominance search parameters*: coverage (recall in the class) and precision thresholds. Only rules that exceed thresholds are delivered to the user. In most of the experiments we used at least 10% coverage and 90% precision.

B. EllipseVis: Interactive Software System

EllipseVis is an interactive software system that allows users to visually analyze datasets in EPC and discover classification rules. By moving the camera, and hiding various elements of the visualization, a user can more easily detect patterns to classify data [3].

EllipseVis is also able to make automatic calculations to find classification patterns. By dividing the graph into smaller rectangular sections, EllipseVis checks if most data points within the rectangle belongs to a single class. Such rectangles are called *class dominant rectangles* (or dominant rectangles for short). EllipseVis discovers high-coverage and high-precision rectangles as dominance rules. These rules depend on how many total points are in the rectangle, i.e. how much coverage the rectangle has, as well as how high the percentage of one class is compared to the others, i.e., precision of classification. Dominance rules can also be interactively created by the user in the EllipseVis. Once a rule is created, the data points inside the rule are separated from the rest of the data and not considered for creation of further rules.

The interactive controls include:

- Move Camera (Pan and Zoom)
- Dominance Rectangle Settings: Set minimum coverage and precision; Set rectangle dimensions; Toggle whether the rectangles are calculating data on their points or intersect lines; Automatically find dominance rectangle rules, Clear rectangles.
- Hide or Show elements: intersect lines, dominance rectangles, cycle between showing all lines, lines not within rules, and line within rules, Side, thin red and thin blue ellipses.

IV. Experiments

The experimentation includes studies with several dataset from UCI ML repository [4].

A. Experiment with Iris Data

Fig. 3 shows all 150 cases of Iris 4-D data losslessly. A small insert shows them inside of the 4-D elliptic coordinates. A zoomed part shows discovered rectangles R_1-R_3 of classification rules $\mathbf{r}_1$-$\mathbf{r}_3$. Fig. 4 zooms the overlap area showing two misclassified red cases. For comparison Fig. 5 shows Iris data in parallel coordinates. It is visible that in EPC the data of three classes are more distinct that helps in search for rectangular rules. This distinction is a result of significant non-linear transformations in EPC in comparison with parallel coordinates.

The rule $\mathbf{r}_1$ covers 52 cases (50 correct 2 misclassified cases), the rule $\mathbf{r}_2$ covers 48 cases (all 48 correct cases) and the rule $\mathbf{r}_3$ covers 50 cases (all 50 correct cases). These rules correctly classified 148 out of 150, i.e., 98.67%. In general rules may not cover all cases, therefore we will use the **weighted precision** formula to compute the total precision:

$$\sum_{i=1}^{n}(p_i c_i) / \sum_{i=1}^{n} c_i \qquad (1)$$

where p_i is precision of rule $\mathbf{r}_i$ and c_i is the number of cases covered by $\mathbf{r}_i$

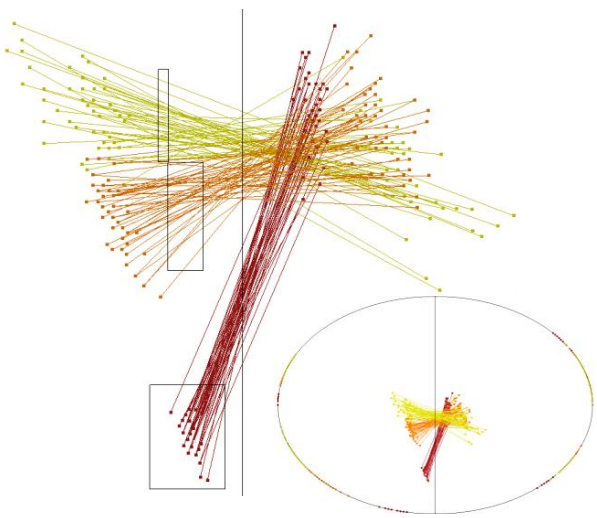


Fig. 3. Three Iris data classes classified with interactively created dominance rules (intersect-based).

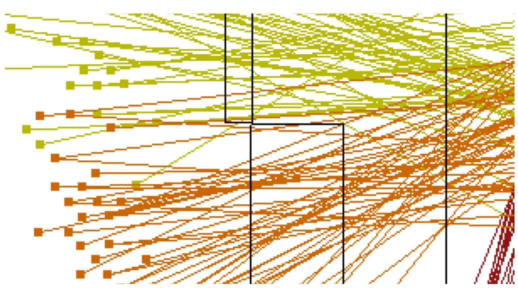


Fig. 4. Zoomed overlapped area from Fig. 3.

B. *Visual Verification of Rules*

An important advantage of EPC is **simplified rule verification** as we show with the Iris example. A common 10-fold Cross Validation (CV) approach randomly splits data to training and validation data to 10 folds to check that the accuracy of prediction on the validation data is acceptable. Besides that, the concept of cross validation can be foreign for the end users EllipseVis allows to simplify or even eliminate it.

Consider Iris data in Fig. 3 and randomly select 90% of lines to training and remaining 10% of them to validation data. If these training cases include the overlap area shown in Fig. 4, then we build rules $\mathbf{r}_1$-$\mathbf{r}_3$ as shows in Fig. 3. The accuracy of these rules is 100% on validation data because they are outside of the overlap area, confirming these rules.

If training data do not include lines in the overlap area, but validation data include them, then the discovered dominance rules (rectangles) can differ from $\mathbf{r}_1$-$\mathbf{r}_3$. For instance, $\mathbf{r}_1$ can be shifted lower or $\mathbf{r}_2$ shifted higher resulted in misclassifying some validation cases. This is a **worst split** of the data to training and validation, because training data *are not representative* to these validation data. The visual EPC representation allows to see it and use this worst split selecting 15 Iris cases (10%) in the overlap area. In the very worst case all of them are misclassified (0 accuracy), but 9 other 10-fold CV splits likely will not contain the overlap area in the validation data due to random split, and will be recognized with 100% accuracy. Thus, the average 10-fold CV accuracy will be 90%. So, finding accurate rules such as $\mathbf{r}_1$-$\mathbf{r}_3$ on the full dataset likely indicates that 0-fold CV result will be similar. Thus, it can be skipped. Instead, we can find visually the worst split and evaluate its accuracy on validation data [1]. It will likely be greater than 0. Therefore the 10-fold average will be greater than 90%.

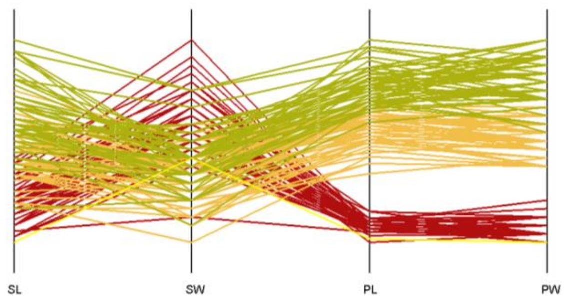


Fig 5. Iris data in parallel coordinates.

C. *Experiment with Wisconsin Breast Cancer Data*

The Wisconsin Breast Cancer (WBC) dataset [4] consists of 683 full 9-D cases: 444 benign (red) and 239 malignant (green) cases as shown in Figs. 6-9. The dimension X_9 is duplicated to X_{10} to create an even number of coordinate be paired. Table 1 and Figs. 6-8 show the *automatically discovered* rules in the EllipseVis system. Together five simple rectangular rules cover **96.34%** of cases with weighted precision **95.13%**. Fig. 9 shows an example where interactive rule discovery is less efficient than automatic requiring two times more rules and larger rectangles.

TABLE 1. RESULTS OF AUTOMATIC WBC RULES DISCOVERY IN EPC (INTERSECT-BASED).

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	B	64.18	98.59
$\mathbf{r}_2$	B	33.10	92.51
$\mathbf{r}_1$ or $\mathbf{r}_2$	**B**	**97.28**	
$\mathbf{r}_3$	M	42.67	92.17
$\mathbf{r}_4$	M	37.23	92.13
$\mathbf{r}_5$	M	14.64	97.14
$\mathbf{r}_3$ or $\mathbf{r}_4$ or $\mathbf{r}_5$	**M**	**94.54**	
All rules	**B, M**	**96.34**	**95.13**

D. *Experiment with Multi-class Glass Data*

The 10-D multi-class Glass Identification dataset consists of 214 cases of 6 imbalanced classes of glass [4]. The U.S. Forensic Science Service gathered this *10-dimensional data* for use in criminal investigations. Fig. 10 shows all Glass classes in EPC. Below we show results of one class vs. all other classes.

All other classes vs. class 5. Table 2 shows all three rules discovered. They cover **87.06%** of cases of all other classes with weighted precision **98.29%**.

TABLE 2. 10-D GLASS RULES FOR CLASS 5 VS. ALL OTHERS (POINT-BASED).

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	All but 5	37.31	100
$\mathbf{r}_2$	All but 5	38.81	98.72
$\mathbf{r}_3$	All but 5	10.95	90.91
All rules	All but 5	**87.06**	**98.29**

All other classes vs. class 6. Fig. 11 shows the result of rule discovery to separate cases of all other classes from class 6. The Ellipse system found three rectangles, R_1-R_3. and respective rules $\mathbf{r}_1$-$\mathbf{r}_3$. Total all three rules **cover 99.51%** of cases with weighted **precision 95.59%** (see Table 3).

TABLE 3. 10-D GLASS RULES FOR CLASS 6 VS. ALL OTHERS (POINT-BASED).

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	All but 6	50.24	99.03
$\mathbf{r}_2$	All but 6	18.05	94.59
$\mathbf{r}_3$	All but 6	31.22	90.63
All rules	All but 6	**99.51**	**95.59**

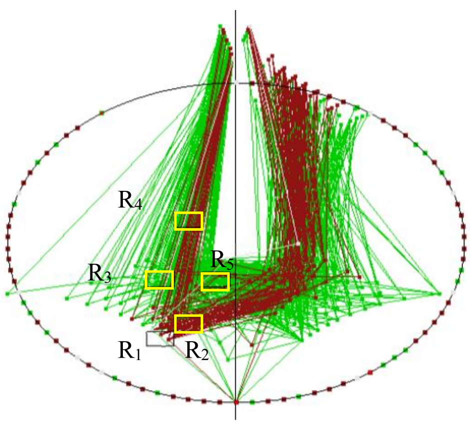


Fig. 6. WBC cases covered by automatically discovered rules $\mathbf{r}_1$- $\mathbf{r}_5$ with rectangles R_1- R_5 : 658 out of 683 cases (96.34%) with 95.13% precsion.

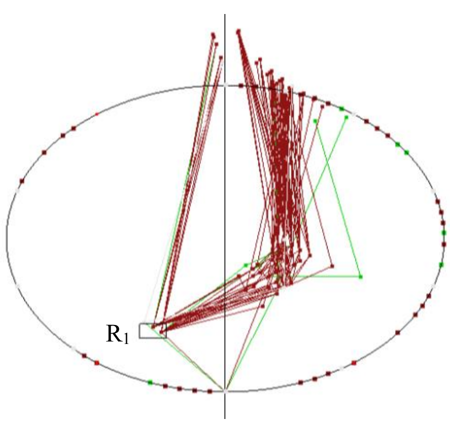


Fig. 7. Cases that satisfy rule $\mathbf{r}_1$ based on the rectangle R_1 that covers 285 cases out of 444 cases of this class (64.18%) with 98.59% precision.

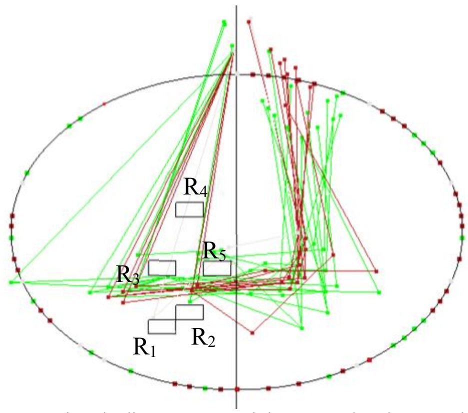


Fig. 8. Automatic rule discovery: remaining cases that do not satisfy rules $\mathbf{r}_1$ - $\mathbf{r}_5$ based on the rectangles R_1- R_5 with 96.34% of total coverage/recall.

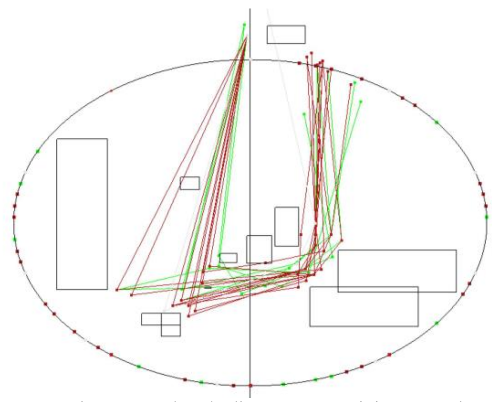


Fig. 9. Interactive "manual" rule discovery: remaining cases that do not satisfy discovered rules based on the shown rectangles with 96.63% of total coverage/recall with two times more rules.

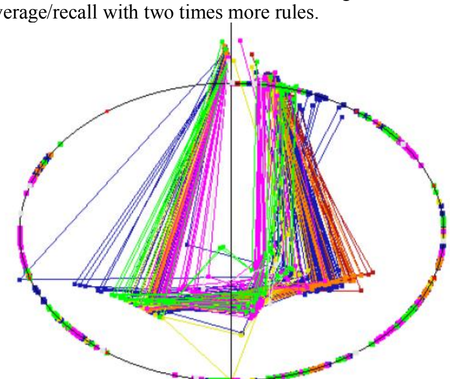


Fig.10. All Glass data of six classes in EPC

All other classes vs. class 7. Total all three rules **cover 87.57%** of cases with total **precision 97.31%** weighted by coverage (see Table 4). In addition, EllipseVis discovered rule $\mathbf{r}_1$ of **class 7 vs. all others** that covered 79.31% of all cases of class 7 with 91.30 % precision (see Table 5). It is intersect-based. As we pointed out before, these rules capture the non-linear relations between 4 attributes, which form a line that intersects the dominance rectangle), while point-based rules capture them between 2 attributes.

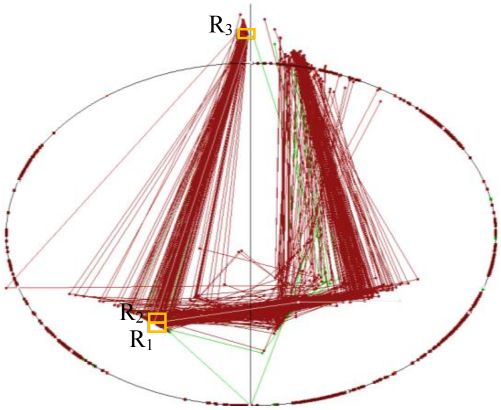


Fig. 11. Extracting rule for Glass: all other classes vs. class 6. Cases that satisfy extracted 3 rules $\mathbf{r}_1$-$\mathbf{r}_3$ based on the rectangles R_1- R_3.

TABLE 4. GLASS RULES FOR CLASS 7 VS. ALL OTHERS (POINT- BASED).

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	All but 7	14.05	96.15
$\mathbf{r}_2$	All but 7	54.59	100.00
$\mathbf{r}_3$	All but 7	18.92	91.43
All rules	All but 7	**87.57**	**97.53**

TABLE 5. GLASS RULES FOR CLASS 7 VS. ALL OTHERS (INTERSECT-BASED).

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	7	**79.31%**	**91.30**

E. Experiment with Car Data

Figs. 12, 13 and Table 6 show the results for Car data with total **coverage of 91.24% and 100% precision.**

TABLE 6. CAR DATA RULE EXPERIMENTATION RESULTS FOR CLASS UNACC (POINT-BASED)

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	UNACC	15.87	100
$\mathbf{r}_2$	UNACC	15.87	100
$\mathbf{r}_3$	UNACC	15.87	100
$\mathbf{r}_4$	UNACC	23.80	100
$\mathbf{r}_5$	UNACC	7.93	100
$\mathbf{r}_6$	UNACC	3.97	100
$\mathbf{r}_7$	UNACC	3.97	100
$\mathbf{r}_8$	UNACC	3.97	100
All rules	UNACC	91.24	100

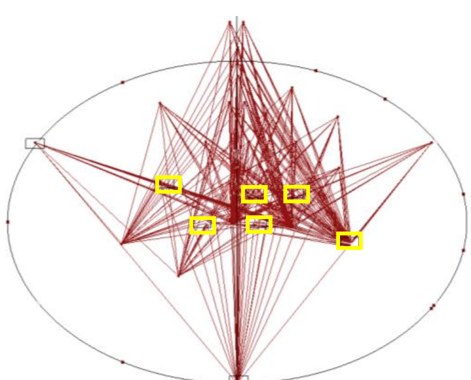


Fig. 12. Extracting rules for Car data: cases of class unacc (red) that satisfy extracted 8 rules $\mathbf{r}_1$-$\mathbf{r}_8$ based on the rectangles R_1 - R_8. (yellow)

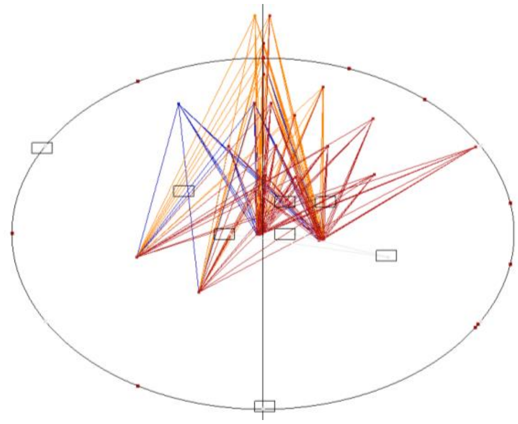


Fig. 13. Remaining Car cases of three other classes that do not satisfy rules $\mathbf{r}_1$ - $\mathbf{r}_8$ based on the rectangles R_1- R_8.

F. Experiment with Ionosphere Data

Figs. 14, 15, tables 7 and 8 show results with Ionosphere data [4]. The results of both intersect- and point-based rules cover both classes with similar coverage (**78.63, 71.51)** and precision (**91.37, 94.02**).

TABLE 7. 34-D IONOSPHERE DATA DOMINANCE RULE EXPERIMENTATION RESULTS (INTERSECT-BASED)

Rule	Class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	b	30.95	97.87
$\mathbf{r}_2$	b	34.13	90.70
$\mathbf{r}_1$ or $\mathbf{r}_2$	b	**65.08**	
$\mathbf{r}_3$	g	18.22	90.24
$\mathbf{r}_4$	g	27.11	90.16
$\mathbf{r}_5$	g	40.89	90.22
$\mathbf{r}_3$ or $\mathbf{r}_4$ or $\mathbf{r}_5$	g	**86.22**	
All rules	b,g	**78.63**	**91.37**

TABLE 8. 34-D IONOSPHERE DATA DOMINANCE RULE EXPERIMENTATION RESULTS (POINT BASED)

Rule	class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	b	29.37	100.00
$\mathbf{r}_2$	b	30.16	97.37
$\mathbf{r}_1$ or $\mathbf{r}_2$	b	**59.53**	
$\mathbf{r}_3$	g	54.67	90.24
$\mathbf{r}_4$	g	23.56	96.23
$\mathbf{r}_3$ or $\mathbf{r}_4$	g	**78.23**	
All rules	b,g	**71.51**	**94.02**

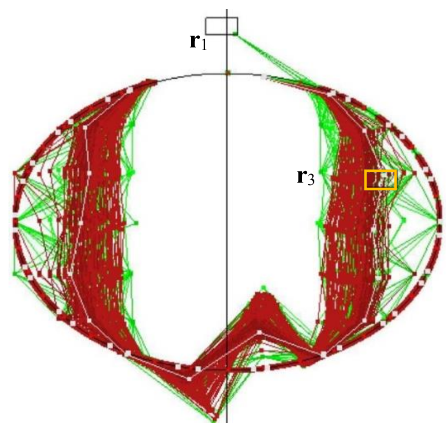


Fig 14. Ionosphere data with rules $\mathbf{r}_1$ (green) and $\mathbf{r}_3$ (red)

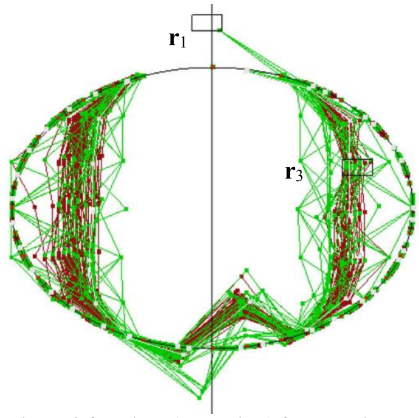


Fig 15. Data that satisfy Rule $\mathbf{r}_1$ (green class) for Ionosphere with data with cases from red class that satisfy Rule $\mathbf{r}_3$ from red class on the background.

G. Experiment with Abalone Data

Table 9, and Fig. 15 show the results for Abalone data [4] on full data and Table 10 shows in split 70%-30% to training and validation data.

TABLE 9. 8-D ABALONE DATA DOMINANCE RULE EXPERIMENTATION RESULTS (POINT BASED)

Rule	class	Coverage/recall in class, %	Precision, %
$\mathbf{r}_1$	1	45.12	92.83

TABLE 10. 8-D ABALONE DATA 70%:30% SPLIT (POINT BASED)

Rule	Class	Training		Validation	
		Coverage, %	Precision, %	Coverage, %	Validation, %
$\mathbf{r}_1$	1	12.12	91.18	12.59	94.12
$\mathbf{r}_2$	1	11.28	90.08	12.59	88.24

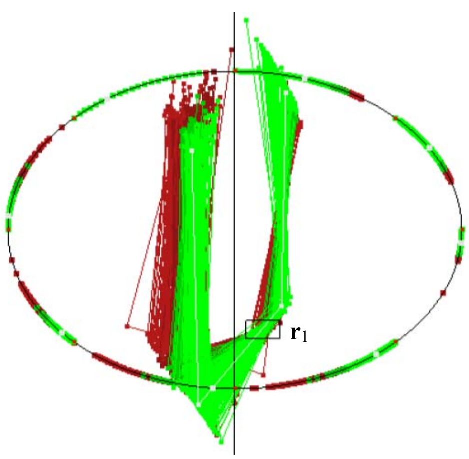


Fig 16. Abalone data with rule $\mathbf{r}_1$.

H. Experiment with Synthetic Data

Fig. 17 shows the same data with non-linear dependencies in parallel coordinates and EPC. A pattern is more complex pattern in parallel coordinates than in EPC.

V. SUMMARY AND CONCLUSION

Table 11 summarizes experiments presented in this paper showing precision of rules from 91% to 100% with recall from 45% to 99% and 1-8 rules per experiment.

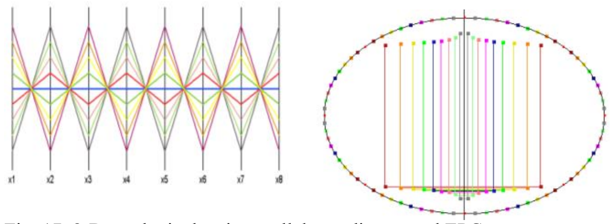


Fig. 17. 8-D synthetic data in parallel coordinates and EPC: (0.9, 0.1, 0.9, 0.1,…), (0.8, 0.2, 0.8, 0.2,…), ….,((0.1, 0.9, 0.1,0.9,…).

These results with EPC using the EllipseVis system appear promising to produce a small number of simple, visual and explainable rules which show cases clearly where a given class dominates.

TABLE 11. DOMINANCE RULE EXPERIMENTATION RESULTS

Experiment	n-D	Classes	Rules	Recall, %	Precision, %
Iris	4	3	3	100	98.66
Cancer	9	2	5	96.33	95.13
Glass 1	10	2[a]	3	87.06	98.29
Glass 2	10	2[a]	3	99.51	95.59
Glass 3	10	2[a]	3	87.57	97.53
Glass 4	10	2[a]	1	79.31	91.30
Car	6	4	8	91.24	100.00
Ionosphere 1	34	2	5	78.63	91.37
Ionosphere 2	34	2	4	71.51	94.02
Abalone	8	2	1	45.12	92.83

a. One class vs. all other classes.

While our results are competitive with presented in [1,5] and exceed some of the other state of the art results, the goal is not competing with the published result in accuracy for these datasets, but showing that the end users and domain experts can discover understandable rules themselves as *self-service* without programming and studying mathematical intricacies of ML algorithms. The next important advantage is simplified rule verification as we discussed using the Iris example above. Prospective work to further develop and evaluate EllipseVis and EPCs includes (1) evaluating EPC on non-benchmark, real-world data, (2) conducting a usability study on levels of human acceptance of manual and automatic rule discovery, and (3) comparing manual and automatic rule discovery results in lossless EPC to lossy data visualiziations.

REFERENCES

[1] Kovalerchuk B. Visual Knowledge Discovery and Machine Learning, Springer, 2018.

[2] Inselberg A. Parallel coordinates: visual multidimensional geometry and its applications. Springer Science & Business Media; 2009 Aug 15.

[3] McDonald, R., Ellipse Data Visualization, https://github.com/McDonaldRo/sure-epc.

[4] Dua, D. , Graff, C. UCI Machine Learning Repository Irvine, CA: University of California, 2019.

[5] Kovalerchuk B, Gharawi A. Decreasing Occlusion and Increasing Explanation in Interactive Visual Knowledge Discovery. HIMI 2018, LNCS 10904, pp. 505–526, 2018.

[6] van der Maaten, L.J.P.; Hinton, G.E. Visualizing Data Using t-SNE. Journal of Machine Learning Research. 9: 2579–2605., 2008.

[7] Ming Y, Qu H, Bertini E. Rulematrix: Visualizing and understanding classifiers with rules. IEEE transactions on visualization and computer graphics. 2018 Aug 20;25(1):342-52