

# ExACT Explainable Clustering: Unravelling the Intricacies of Cluster Formation

*Federico Sabbatini*\*   Roberta Calegari<sup>†</sup>

\*Department of Pure and Applied Sciences (DiSPeA) – University of Urbino

<sup>†</sup>Department of Computer Science and Engineering (DISI) – University of Bologna

*f.sabbatini1@campus.uniurb.it*, *roberta.calegari@unibo.it*

KoDis'23: The 2<sup>nd</sup> International Workshop on Knowledge Diversity  
September 3, 2023, Rhodes, Greece



This work has been supported by European Union's Horizon Europe AEQUITAS research and innovation programme under grant number 101070363

# Next in Line...

- 1 Context & Motivation
- 2 Explainable Clustering with ExACT
- 3 Results & Conclusions

# Context

## Black-box predictors

- currently adopted in almost every field [Rocha et al. \(2012\)](#)
- detect patterns and relationships buried in data
- ✓ knowledge acquired and reused in similar applications
  - ✓ impressive predictive capabilities
- ✗ knowledge sub-symbolically represented (internal parameters)
  - ✗ not suitable for human comprehension (BB behaviour [Lipton \(2018\)](#))
  - ✗ no **explanations** provided for predictions
- unreliability in critical applications
  - e.g. healthcare, finance, etc.
- 👉 **interpretable predictors** [Rudin \(2019\)](#)
- 👉 symbolic knowledge-extraction techniques [Kenny et al. \(2021\)](#)

# Motivation I

## Traditional clustering

- several techniques available in the literature
- algorithm peculiarities related to clusters' properties
  - e.g. Gaussian mixture models (GMMs) [Murphy \(2012\)](#)
    - ✓ flexible, robust to outliers
    - ✗ sensitive to initialisation, tuning the component amount
  - e.g. DBSCAN [Ling \(1972\)](#)
    - ✓ robust to outliers, no assumptions on clusters' shape and amount
    - ✗ clusters with different densities, large data sets, tuning  $\epsilon$
  - e.g. k-means [Lloyd \(1982\)](#)
    - ✓ no assumptions on clusters' size and density, large data sets
    - ✗ tuning k, sensitive to outliers, large amount of dimensions
- no optimum technique outperforming all the others in any scenario
- ✗ common opaque behaviour

# Motivation II

## Explainable clustering

- growing relevance, especially in critical areas [Manduchi et al. \(2021\)](#)
- provide human-interpretable explanations about clusters' formation
  - e.g. tree-based [Basak and Krishnapuram \(2005\)](#); [Fraiman et al. \(2013\)](#); [Bertsimas et al. \(2018\)](#)
    - input space partitioning via perpendicular cutting hyperplanes
    - top-down, from root to leaves, data set iterative partitioning
    - bottom-up, grouping data set instances to compose clusters
  - e.g. IMM [Dasgupta et al. \(2020\)](#)
    - tree-based,  $k$  clusters =  $k$  leaves  $\implies$  max depth =  $k-1$
    - explanations for assignments based on at most  $k-1$  constraints
  - e.g. CLASSIX [Chen and Güttel \(2022\)](#)
    - based on data sorting, grouping and merging
    - global explanation in terms of starting points' coordinates
    - local explanation for instances belonging to specific clusters
    - local explanation for pairs belonging or not to the same cluster

# Contribution

## Explainable clustering with EXACT

Extends the current state of the art

- 👉 tree-based clustering
- 👉 exhaustive
- 👉 hierarchical rules
- 👉 disjoint rules
- ! shorter human-interpretable explanations
- ! higher predictive performance

## Next in Line...

- 1 Context & Motivation
- 2 Explainable Clustering with ExACT**
- 3 Results & Conclusions

# ExACT

## Basic idea

- ExACT = EXplainable Automated Clustering Technique
- recursive algorithm performing explainable clustering
- based on top-down binary decision tree induction

## Features

- recursive strategy and tree induction
- GMMs to identify clusters
- DBSCAN to remove outliers
- hypercubes to obtain human-interpretable clusters' approximation
- applicable to clustering, classification and regression tasks

# ExACT – User-Defined Parameters

## Required parameters

$\delta$  maximum recursion depth

- $\delta = d \implies d + 1$  clusters (upper-bound)
- trade-off: deep trees  $\rightarrow$  high performance  $\rightarrow$  small readability

$\theta$  error threshold calibrating predictive performance

- tree nodes with errors smaller than  $\theta$  are not further partitioned
- wrong assignments / misclassifications / mean absolute error

$\xi$  maximum amount of clusters identifiable via GMMs

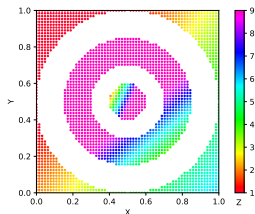
- upper-bound value used for automated BIC score evaluations
- provide  $\xi$  values  $\geq$  the expected number of clusters

# ExACT – The Algorithm

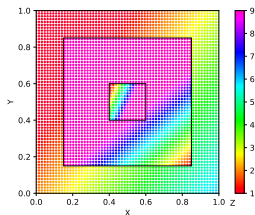
## Fundamental steps

1. associate the surrounding cube with the tree root
  - set it as the current node
2. apply GMM clustering to the current node and identify  $n$  clusters
3. apply DBSCAN to the  $n$  clusters for outlier removal
4. approximate each cluster with a minimal hypercube
5. select the best approximating cube (biggest volume)
6. create the difference cube
  - ! difference cube = current node - approximating cube
7. expand the current tree node
  - right child  $\rightarrow$  best approximating cube
  - left child  $\rightarrow$  corresponding difference cube
8. repeat from 2. for the right child node if  $predictiveError > \theta$

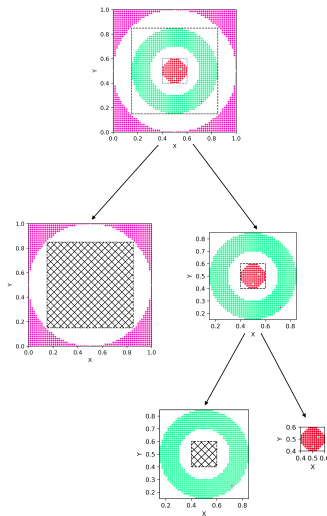
# ExACT – Intuitive Example



Data set



ExACT partitioning



Binary tree induced by ExACT

# ExACT – Tuning GMMs and DBSCAN

## Tuning the number of components in GMMs

Automated tuning via Bayes information criterion (BIC)

1. train several GMM instances with different component amount
2. calculate the BIC score for each GMM instance
3. select the GMM instance having lowest BIC score

## Tuning the $\varepsilon$ and *minPoints* parameters in DBSCAN

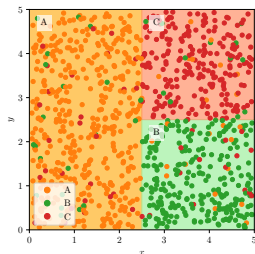
Automated tuning via knee method [Rahmah and Sitanggang \(2016\)](#)

1. compute the  $k$ -nearest neighbour distances for each training instance
2. plot the distances in ascending order
3. determine the knee in the plot
  - i.e. steepest slope amongst the slopes between pairs of adjacent points
4.  $\text{minPoints} = k$ ,  $\varepsilon$  = distance corresponding to the knee

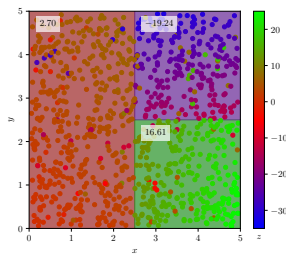
# ExACT – Outputs

## Outputs associated with the identified hypercubes

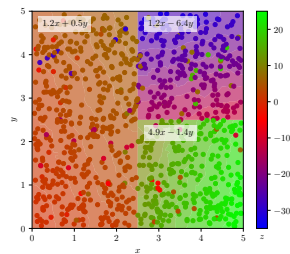
According to the generalisation proposed by Sabbatini et al. (2022a, 2023)



Clustering (cluster id)  
Classification (label)



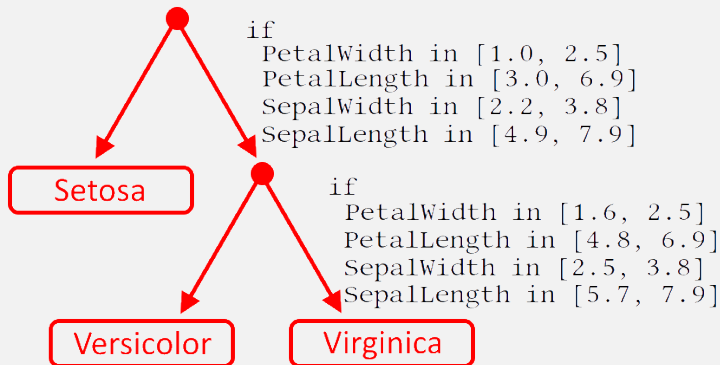
Regression  
(constant values)



Regression  
(linear equations)

# ExACT – Iris Data Set Example

## Example of ExACT clustering for the Iris data set



Feature values are expressed in cm

# ExACT – Implementation



<https://apice.unibo.it/xwiki/bin/view/PSyKE>  
<https://github.com/psykei/psyke-python>

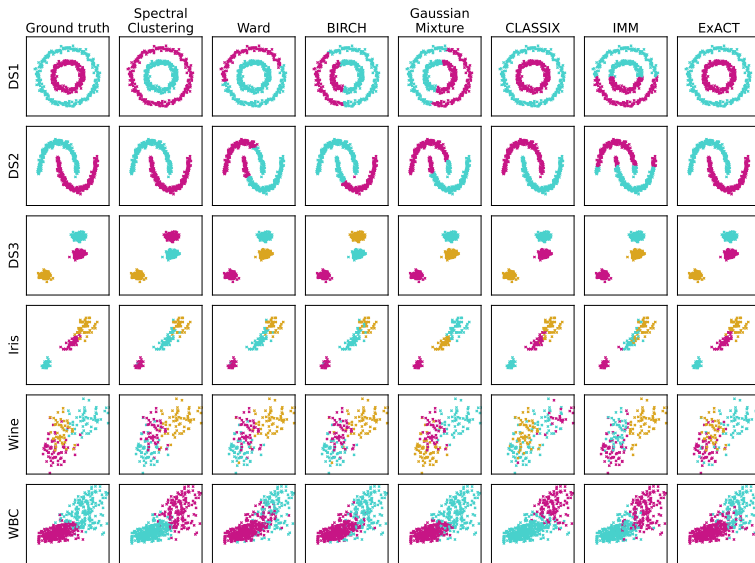
## The PSyKE framework Sabbatini et al. (2021, 2022b,c); Calegari and Sabbatini (2023)

- general-purpose SKE framework
- unified interface for knowledge extraction and explainable clustering
- tools for data manipulation and visualisation
- tools for automated parameter tuning
- main implementation in Python

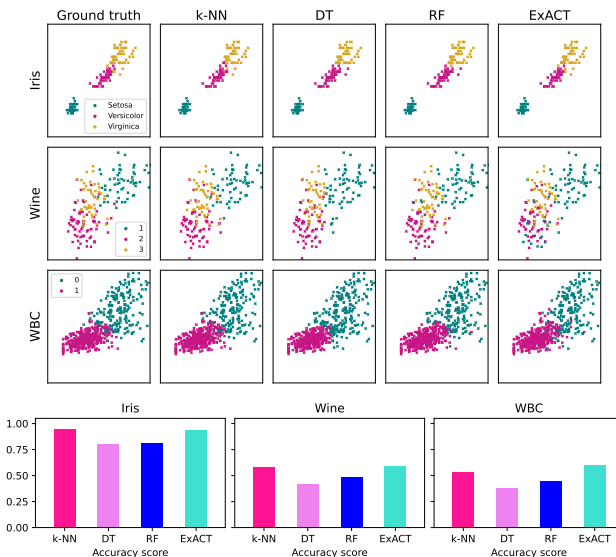
# Next in Line...

- 1 Context & Motivation
- 2 Explainable Clustering with ExACT
- 3 Results & Conclusions**

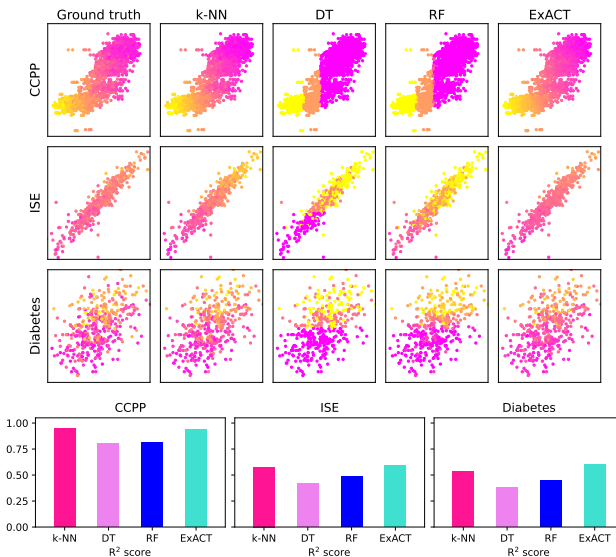
# Experimental Results – Clustering



# Experimental Results – Classification



# Experimental Results – Regression



# Conclusions

## ExACT

- new explainable clustering technique
- tree-based top-down strategy
- hypercube-based input space human-interpretable partitioning
- underlying GMMs and DBSCAN instances
- accepts continuous input features
- applicable for clustering, classification, regression tasks

## Future works

- find more efficient underlying algorithms
  - cluster detection/outlier removal
- enhance the splitting strategy of EXACT

# ExACT Explainable Clustering: Unravelling the Intricacies of Cluster Formation

*Federico Sabbatini*\*   Roberta Calegari<sup>†</sup>

\*Department of Pure and Applied Sciences (DiSPeA) – University of Urbino

<sup>†</sup>Department of Computer Science and Engineering (DISI) – University of Bologna

*f.sabbatini1@campus.uniurb.it*, *roberta.calegari@unibo.it*

KoDis'23: The 2<sup>nd</sup> International Workshop on Knowledge Diversity  
September 3, 2023, Rhodes, Greece



This work has been supported by European Union's Horizon Europe AEQUITAS research and innovation programme under grant number 101070363

# References I

- Basak, J. and Krishnapuram, R. (2005). Interpretable hierarchical clustering by constructing an unsupervised decision tree. *IEEE Trans. Knowl. Data Eng.*, 17(1):121–132.
- Bertsimas, D., Orfanoudaki, A., and Wiberg, H. M. (2018). Interpretable clustering via optimal trees. *CoRR*, abs/1812.00539.
- Calegari, R. and Sabbatini, F. (2023). The PSyKE technology for trustworthy artificial intelligence. 13796:3–16. XXI International Conference of the Italian Association for Artificial Intelligence, AlxIA 2022, Udine, Italy, November 28 – December 2, 2022, Proceedings.
- Chen, X. and Güttel, S. (2022). Fast and explainable clustering based on sorting. *CoRR*, abs/2202.01456.
- Dasgupta, S., Frost, N., Moshkovitz, M., and Rashtchian, C. (2020). Explainable k-means and k-medians clustering. *CoRR*, abs/2002.12538.
- Fraiman, R., Ghattas, B., and Svarc, M. (2013). Interpretable clustering using unsupervised binary trees. *Advances in Data Analysis and Classification*, 7(2):125–145.
- Kenny, E. M., Ford, C., Quinn, M., and Keane, M. T. (2021). Explaining black-box classifiers using post-hoc explanations-by-example: The effect of explanations and error-rates in XAI user studies. *Artificial Intelligence*, 294:103459.
- Ling, R. F. (1972). On the theory and construction of k-clusters. *The Computer Journal*, 15(4):326–332.

# References II

- Lipton, Z. C. (2018). The mythos of model interpretability. *Queue*, 16(3):31–57.
- Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Trans. Inf. Theory*, 28(2):129–136.
- Manduchi, L., Hüser, M., Faltys, M., Vogt, J. E., Rätsch, G., and Fortuin, V. (2021). T-DPSOM: an interpretable clustering method for unsupervised learning of patient health states. In Ghassemi, M., Naumann, T., and Pierson, E., editors, *ACM CHIL '21: ACM Conference on Health, Inference, and Learning, Virtual Event, USA, April 8-9, 2021*, pages 236–245. ACM.
- Murphy, K. P. (2012). *Machine learning – A probabilistic perspective*. Adaptive computation and machine learning series. MIT Press.
- Rahmah, N. and Sitanggang, I. S. (2016). Determination of optimal epsilon (eps) value on DBSCAN algorithm to clustering data on peatland hotspots in Sumatra. In *IOP conference series: earth and environmental science*, volume 31, page 012012. IOP Publishing.
- Rocha, A., Papa, J. P., and Meira, L. A. A. (2012). How far do we get using machine learning black-boxes? *International Journal of Pattern Recognition and Artificial Intelligence*, 26(02):1261001–(1–23).
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215.

# References III

- Sabbatini, F., Ciatto, G., Calegari, R., and Omicini, A. (2021). On the design of PSyKE: A platform for symbolic knowledge extraction. In Calegari, R., Ciatto, G., Denti, E., Omicini, A., and Sartor, G., editors, *WOA 2021 – 22nd Workshop “From Objects to Agents”*, volume 2963 of *CEUR Workshop Proceedings*, pages 29–48. Sun SITE Central Europe, RWTH Aachen University. 22nd Workshop “From Objects to Agents” (WOA 2021), Bologna, Italy, 1–3 September 2021. Proceedings.
- Sabbatini, F., Ciatto, G., Calegari, R., and Omicini, A. (2022a). Hypercube-based methods for symbolic knowledge extraction: Towards a unified model. In Ferrando, A. and Mascardi, V., editors, *WOA 2022 – 23rd Workshop “From Objects to Agents”*, volume 3261 of *CEUR Workshop Proceedings*, pages 48–60. Sun SITE Central Europe, RWTH Aachen University.
- Sabbatini, F., Ciatto, G., Calegari, R., and Omicini, A. (2022b). Symbolic knowledge extraction from opaque ML predictors in PSyKE: Platform design & experiments. *Intelligenza Artificiale*, 16(1):27–48.
- Sabbatini, F., Ciatto, G., Calegari, R., and Omicini, A. (2023). Towards a unified model for symbolic knowledge extraction with hypercube-based methods. *Intelligenza Artificiale*, 17(1):63–75.
- Sabbatini, F., Ciatto, G., and Omicini, A. (2022c). Semantic Web-based interoperability for intelligent agents with PSyKE. In Calvaresi, D., Najjar, A., Winikoff, M., and Främling, K., editors, *Explainable and Transparent AI and Multi-Agent Systems*, volume 13283 of *Lecture Notes in Computer Science*, chapter 8, pages 124–142. Springer.