Poster presentation: SAC'22, Machine Learning and its Applications Track

Authors: Sophie Steger ¹, Bernhard C. Geiger ², Marek Śmieja ³ Affiliations: Graz University of Technology, Austria ¹ Know-Center GmbH, Austria ² Jagiellonian University, Poland ³

on April 27, 2022



Experiments

Limitations 00 Conclusior

» Content

Introduction

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion

- * Unsupervised clustering
- Semi-supervised clustering

- * Unsupervised clustering
- Semi-supervised clustering

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Unsupervised clustering



Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Unsupervised clustering

Given:

* Set of unlabeled data points

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Unsupervised clustering

Given:

* Set of unlabeled data points

Aim:



Experiments

Limitations 00 Conclusior

» Unsupervised clustering

Given:

* Set of unlabeled data points

Aim:

* Group data points based on some "similarity" measure

- * Unsupervised clustering
- Semi-supervised clustering

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Semi-supervised clustering



Experiments

Limitations 00 Conclusior

» Semi-supervised clustering

Prior information:

* Partition-level side information: labeled data points



Experiments

Limitations 00 Conclusior 000

» Semi-supervised clustering

- * Partition-level side information: labeled data points
- * Instance-level side information: pairwise constraints



Experiments

Limitations 00 Conclusior

» Semi-supervised clustering

- * Partition-level side information: labeled data points
- * Instance-level side information: pairwise constraints
 - * must-link constraints



Experiments

Limitations 00 Conclusior

» Semi-supervised clustering

- * Partition-level side information: labeled data points
- * Instance-level side information: pairwise constraints
 - * must-link constraints
 - * cannot-link constraints

- * Markov Aggregation
- * Clustering via Markov Aggregation
- * Algorithm

- * Markov Aggregation
- * Clustering via Markov Aggregation
- * Algorithm

Semi-Supervised Clustering via Markov Aggregation

» Markov Aggregation

Experiments

Limitations

Conclusion

Introduction

Given:

page 4

Semi-Supervised Clustering via Markov Aggregation ○●●○○○○○○

» Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

Given:

* Markov chain $\mathbf{X} \sim \operatorname{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Semi-Supervised Clustering via Markov Aggregation

» Markov Aggregation

Experiments

Limitation: 00 Conclusior

Introduction

Given:

 $* \; \mathsf{Markov}$ chain $X \sim \mathrm{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

Semi-Supervised Clustering via Markov Aggregation

» Markov Aggregation

Experiments

Limitation: 00 Conclusior 000

Introduction

Given:

 $* \; \mathsf{Markov}$ chain $X \sim \mathrm{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

 $* \;$ Find a Markov chain $\mathbf{Y} \sim \operatorname{Mar}(\mathcal{Y}, \mathbb{Q})$ on the alphabet $\mathcal Y$ that

Semi-Supervised Clustering via Markov Aggregation

» Markov Aggregation

Experiments

Limitation: 00 Conclusior 000

Introduction

Given:

 $* \; \mathsf{Markov}$ chain $X \sim \mathrm{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

- $*\,$ Find a Markov chain $\mathbf{Y}\sim \operatorname{Mar}(\mathcal{Y},\mathbb{Q})$ on the alphabet $\mathcal Y$ that
 - $\ast~$ best approximates ${\bf X}$

Semi-Supervised Clustering via Markov Aggregation

» Markov Aggregation

Experiments

Limitation: 00 Conclusior 000

Introduction

Given:

 $* \; \mathsf{Markov}$ chain $X \sim \mathrm{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

- $* \;$ Find a Markov chain $Y \sim \operatorname{Mar}(\mathcal{Y}, \mathbb{Q})$ on the alphabet $\mathcal Y$ that
 - $\ast~$ best approximates ${\bf X}$
 - $\ast\,$ on a smaller alphabet ${\cal Y}$

>>

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

Introduction

Given:

* Markov chain $\mathbf{X} \sim \operatorname{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

- * Find a Markov chain $\mathbf{Y} \sim \operatorname{Mar}(\mathcal{Y}, \mathbb{Q})$ on the alphabet $\mathcal Y$ that
 - * best approximates \mathbf{X}

Markov Aggregation

* on a smaller alphabet ${\mathcal Y}$

How:

>>

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

Introduction

Given:

* Markov chain $\mathbf{X} \sim \operatorname{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

- * Find a Markov chain $\mathbf{Y} \sim \operatorname{Mar}(\mathcal{Y}, \mathbb{Q})$ on the alphabet $\mathcal Y$ that
 - $\ast~$ best approximates ${\bf X}$

Markov Aggregation

 $\ast~$ on a smaller alphabet ${\cal Y}$

How:

 $\ast~$ Mapping of ${\bf X}$ onto ${\bf Y}$ via a

>>

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

Introduction

Given:

* Markov chain $\mathbf{X} \sim \operatorname{Mar}(\mathcal{X}, \mathbb{P})$ on the alphabet \mathcal{X}

Aim:

- * Find a Markov chain $\mathbf{Y} \sim \operatorname{Mar}(\mathcal{Y}, \mathbb{Q})$ on the alphabet $\mathcal Y$ that
 - $\ast~$ best approximates ${\bf X}$

Markov Aggregation

 $\ast~$ on a smaller alphabet ${\cal Y}$

How:

- $\ast~$ Mapping of ${\bf X}$ onto ${\bf Y}$ via a
 - * deterministic mapping $\rightarrow g: \mathcal{X} \rightarrow \mathcal{Y}$

Semi-Supervised Clustering via Markov Aggregation

» Cost function

Experiments

Limitation:

Conclusior 000



The mapping g should achieve that . . .

Semi-Supervised Clustering via Markov Aggregation

» Cost function

Experiments

Limitation:

Conclusior 000



The mapping g should achieve that . . .

1. the hidden Markov process ${\bf Y}$ is Markov &

Semi-Supervised Clustering via Markov Aggregation

» Cost function

Experiments

Limitations

Conclusior 000

Rationale

The mapping g should achieve that . . .

- 1. the hidden Markov process ${\bf Y}$ is Markov &
- 2. preserves the *temporal dependence structure* of \mathbf{X} .

>>

Cost function

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion

$igcap_{eta}(\mathbf{X},\mathbb{W}):=(1-2eta)\left(H(Y_2|Y_1)-H(Y_2|X_1) ight)-eta I(Y_1;Y_2) ight)$

Rana Ali Amjad, Clemens Blochl, and Bernhard C. Geiger. "A Generalized Framework For Kullback–Leibler Markov Aggregation". In: *IEEE Transactions on Automatic Control* 65.7 (July 2020), pp. 3068–3075. ISSN: 2334-3303. DOI: 10.1109/tac.2019.2945891. URL: http://dx.doi.org/10.1109/TAC.2019.2945891



Semi-Supervised Clustering via Markov Aggregation $\circ 000 \bullet \circ \circ \circ \circ \circ$

Experiments

Limitations 00 Conclusion

» Cost function

$$\mathcal{C}_{\beta}(\mathbf{X}, \mathbb{W}) := (1 - 2\beta) \left(H(Y_2|Y_1) - H(Y_2|X_1) \right) - \beta I(Y_1;Y_2)$$

^{IR®} Special cases:

Amjad, Blochl, and Geiger, "A Generalized Framework For Kullback-Leibler Markov Aggregation"

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion

» Cost function

$$\mathcal{C}_{\beta}(\mathbf{X}, \mathbb{W}) := (1 - 2\beta) \left(H(Y_2|Y_1) - H(Y_2|X_1) \right) - \beta I(Y_1; Y_2)$$

^{IR™} Special cases:

* $\beta = 0.5$: information-theoretic pairwise clustering

Amjad, Blochl, and Geiger, "A Generalized Framework For Kullback-Leibler Markov Aggregation"

Introduction 00000

Experiments

Limitations

Conclusior 000

» Cost function

$$\mathcal{C}_{\beta}(\mathbf{X}, \mathbb{W}) := (1 - 2\beta) \left(H(Y_2|Y_1) - H(Y_2|X_1) \right) - \beta I(Y_1; Y_2)$$

^{II}S Special cases:

- * $\beta = 0.5$: information-theoretic pairwise clustering
- * $\beta = 1$: information bottleneck problem

Amjad, Blochl, and Geiger, "A Generalized Framework For Kullback-Leibler Markov Aggregation"

- Markov Aggregation
- * Clustering via Markov Aggregation
- * Algorithm

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion

» Clustering via Markov Aggregation

Introduction

Idea:

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

* Minimizing the cost function ightarrow grouping of states ${f X}$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function o grouping of states ${f X}$
- \ast Use for grouping of data points
Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function ightarrow grouping of states ${f X}$
- \ast Use for grouping of data points

How:

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function o grouping of states ${f X}$
- * Use for grouping of data points

How:

* Construct the transition probability matrix $\mathbb P$

$$P_{i \to j} \propto \mathrm{e}^{-rac{\|x_i - x_j\|_2^2}{\sigma_k}}$$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function o grouping of states ${f X}$
- \ast Use for grouping of data points

How:

* Construct the transition probability matrix $\mathbb P$

$$P_{i \to j} \propto \mathrm{e}^{-rac{\|x_i - x_j\|_2^2}{\sigma_k}}$$

 $* \, x_i, \, x_j$ coordinate of i-th and j-th data point

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function o grouping of states ${f X}$
- * Use for grouping of data points

How:

* Construct the transition probability matrix $\mathbb P$

$$P_{i \to j} \propto \mathrm{e}^{-rac{\|x_i - x_j\|_2^2}{\sigma_k}}$$

 $* x_i, x_j \ldots \ldots$ coordinate of i-th and j-th data point $* \|\cdot\|_2^2 \ldots$ squared Euclidean distance

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

Introduction

» Clustering via Markov Aggregation

Idea:

- * Minimizing the cost function o grouping of states ${f X}$
- * Use for grouping of data points

How:

 $\ast\,$ Construct the transition probability matrix $\mathbb P$

$$P_{i \to j} \propto \mathrm{e}^{-rac{\|x_i - x_j\|_2^2}{\sigma_k}}$$

* $x_i, x_j \ldots$ coordinate of *i*-th and *j*-th data point * $\|\cdot\|_2^2 \ldots$ squared Euclidean distance * $\sigma_k \ldots \ldots$ scaling parameter

Semi-Supervised Clustering via Markov Aggregation

- * Markov Aggregation
- * Clustering via Markov Aggregation
- * Algorithm

Semi-Supervised Clustering via Markov Aggregation $\circ\circ\circ\circ\circ\circ\circ\bullet\circ$

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm







.

.



Experiments

Limitations

Conclusior

» Semi-supervised sequential algorithm





- :
- -



Semi-Supervised Clustering via Markov Aggregation $\circ\circ\circ\circ\circ\circ\circ\bullet\circ$

Experiments

Limitations

Conclusior 000

page 8

» Semi-supervised sequential algorithm

Visualization



Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm





Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Semi-supervised sequential algorithm





Clusters: $y_1 \quad y_2 \quad y_3$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm





Clusters: $y_1 \quad y_2 \quad y_3$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm





Clusters: $y_1 \quad y_2 \quad y_3$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm





Clusters: $\begin{array}{c} y_1 & y_2 & y_3 \\ \bullet & \bullet & \bullet \end{array}$



Experiments

Limitations

Conclusior

page 9

» Semi-supervised sequential algorithm

Visualization





Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm





- :
- •



Semi-Supervised Clustering via Markov Aggregation $\circ \circ \circ \circ \circ \circ \circ \circ \bullet$

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm





- •
- •

 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

Semi-Supervised Clustering via Markov Aggregation ○○○○○○○● Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm





 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

 y_2 not allowed

Semi-Supervised Clustering via Markov Aggregation ○○○○○○○● Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm





 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

 y_2 not allowed

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Semi-supervised sequential algorithm

Visualization





 y_2 not allowed

 $C_{g_{y_3}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_3})$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior

» Semi-supervised sequential algorithm

Visualization





 y_2 not allowed

 $C_{g_{y_3}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_3})$

 $\Rightarrow g = \operatorname*{arg\,min}_{g_{y_i}} C_{g_{y_i}}$



Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm

Visualization



- :
- -

Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm

Visualization



 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Semi-supervised sequential algorithm

Visualization



- :
- •

 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

$$C_{g_{y_2}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_2})$$

Semi-Supervised Clustering via Markov Aggregation ○○○○○○○● Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm

Visualization





$$C_{g_{y_2}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_2})$$

 y_3 not allowed

Experiments

Limitations

Conclusior

» Semi-supervised sequential algorithm

Visualization





 $C_{g_{y_2}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_2})$

 y_3 not allowed

 $\Rightarrow g = \mathop{\rm arg\,min}_{g_{y_i}} C_{g_{y_i}}$



Experiments

Limitations 00 Conclusior

» Semi-supervised sequential algorithm

Visualization





 $C_{g_{y_2}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_2})$

 y_3 not allowed

 $\Rightarrow g = \mathop{\rm arg\,min}_{g_{y_i}} C_{g_{y_i}}$



Semi-supervised sequential algorithm



Experiments

 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm

Visualization



page 9

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm

Visualization





$$C_{g_{y_2}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_2})$$

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

» Semi-supervised sequential algorithm

Visualization



- $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$
- $C_{g_{y_2}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_2})$
- $C_{g_{y_3}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_3})$

Experiments

Limitations 00 Conclusior 000

» Semi-supervised sequential algorithm





 $egin{aligned} C_{g_{y_1}} &= \mathcal{C}_eta(\mathbf{X}, g_{y_1}) \ C_{g_{y_2}} &= \mathcal{C}_eta(\mathbf{X}, g_{y_2}) \ C_{g_{y_3}} &= \mathcal{C}_eta(\mathbf{X}, g_{y_3}) \ \Rightarrow g &= rgmin_{g_{y_i}} C_{g_{y_i}} \end{aligned}$

Experiments

Limitations

Conclusior

» Semi-supervised sequential algorithm





 $C_{g_{y_1}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_1})$

 $C_{g_{y_2}} = \mathcal{C}_{\beta}(\mathbf{X}, g_{y_2})$

 $C_{g_{y_3}} = \mathcal{C}_\beta(\mathbf{X}, g_{y_3})$

 $\Rightarrow g = \argmin_{g_{y_i}} C_{g_{y_i}}$

Experiments

- * Experimental setup
- * Influence of the hyperparameters
- * Side-information
- * Few labeled classes

Experiments

- * Experimental setup
- Influence of the hyperparameters
- Side-information
- * Few labeled classes

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior

» Clustering accuracy

Normalized Mutual Information:

$$\mathsf{NMI}(\mathcal{X},\mathcal{Y}) = rac{2I(\mathcal{Y};\mathcal{X})}{H(\mathcal{Y}) + H(\mathcal{X})}$$
Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Clustering accuracy

Normalized Mutual Information:

$$\mathsf{NMI}(\mathcal{X},\mathcal{Y}) = rac{2I(\mathcal{Y};\mathcal{X})}{H(\mathcal{Y}) + H(\mathcal{X})}$$

 $* \mathcal{X} \dots \dots$ true partition

>>

Semi-Supervised Clustering via Markov Aggregation

Clustering accuracy

Normalized Mutual Information:

$$\mathsf{NMI}(\mathcal{X},\mathcal{Y}) = rac{2I(\mathcal{Y};\mathcal{X})}{H(\mathcal{Y}) + H(\mathcal{X})}$$

* \mathcal{X} true partition * \mathcal{Y} estimated partition

page 10

Experiments

Limitations 00 Conclusion

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior 000

» Clustering accuracy

Normalized Mutual Information:

$$\mathsf{NMI}(\mathcal{X},\mathcal{Y}) = rac{2I(\mathcal{Y};\mathcal{X})}{H(\mathcal{Y}) + H(\mathcal{X})}$$

- $* \mathcal{X} \dots \dots$ true partition
- * ${\mathcal Y}$ estimated partition
- * NMI = 0: no mutual information

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusior 000

» Clustering accuracy

Normalized Mutual Information:

$$\mathsf{NMI}(\mathcal{X},\mathcal{Y}) = rac{2I(\mathcal{Y};\mathcal{X})}{H(\mathcal{Y}) + H(\mathcal{X})}$$

- * \mathcal{X} true partition
- * ${\mathcal Y}$ estimated partition
- * NMI = 0: no mutual information
- * NMI = 1: identical partitions

Semi-Supervised Clustering via Markov Aggregation

» Side-information

Experiments ○00●○○○○○○○○○○○○○ Limitations 00 Conclusior

>>

Semi-Supervised Clustering via Markov Aggregation

Side-information

Experiments

Limitations 00 Conclusior 000

Pairwise constraints:

* Select a fraction of data points:

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Side-information

- * Select a fraction of data points:
 - \Rightarrow Assign the true label

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Side-information

- * Select a fraction of data points:
 - \Rightarrow Assign the true label
 - \Rightarrow Generation of pairwise constraints

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusior

» Side-information

- * Select a fraction of data points:
 - \Rightarrow Assign the true label
 - \Rightarrow Generation of pairwise constraints
- $\ast\,$ Average resulting NMI over 10 runs

>>

Semi-Supervised Clustering via Markov Aggregation $_{\rm OOOOOOOOO}$

Side-information

Experiments

Limitations

Conclusion

Unsupervised Clustering





>>

Semi-Supervised Clustering via Markov Aggregation

Side-information

Experiments

Limitations 00 Conclusion

Unsupervised Clustering





»

Semi-Supervised Clustering via Markov Aggregation

Side-information

Experiments

Limitations

Conclusion

Semisupervised Clustering





»

Semi-Supervised Clustering via Markov Aggregation

Side-information

Experiments ○00●○○○○○○○○○○○○○ Limitations

Conclusion 000

Semisupervised Clustering



Experiments

- * Experimental setup
- * Influence of the hyperparameters
- * Side-information
- * Few labeled classes

Semi-Supervised Clustering via Markov Aggregation

» Influence of k

Experiments

Limitation 00 Conclusior



Setup:

page 12

Introduction 00000

» Influence of k

Experiments

Limitation: 00 Conclusior 000



Setup:

* Parameters: $\beta = \{0, 0.5, 1\}$, k varied

Introduction 00000

» Influence of k

Experiments

Limitation: 00 Conclusior 000



Setup:

- * Parameters: $\beta = \{0, 0.5, 1\}$, k varied
- \ast Fraction of labeled data: 0, 10, 20, 30%

Introduction 00000

» Influence of k

Experiments

Limitation: 00 Conclusior 000



Setup:

- * Parameters: $\beta = \{0, 0.5, 1\}$, k varied
- \ast Fraction of labeled data: 0, 10, 20, 30%
- * Datasets: circles, iris

Introduction 00000

»

Influence of k

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion



Circles . Sepal width 0.2 - 0.2 - 0.2 > 0-2 2.5 --15 -2.0 -• $\frac{1}{5}$ 10 6 8 Sepal length

Introduction 00000

» Influence of k

Experiments

Limitation: 00 Conclusior 000



Setup:

- * Parameters: $\beta = \{0, 0.5, 1\}$, k varied
- \ast Fraction of labeled data: 0, 10, 20, 30%
- * Datasets: circles, iris

Observations:

Introduction 00000

» Influence of k

Experiments

Limitations

Conclusion



Setup:

- * Parameters: $\beta = \{0, 0.5, 1\}$, k varied
- \ast Fraction of labeled data: 0, 10, 20, 30%
- * Datasets: circles, iris

Observations:

* Influence on accuracy for unsupervised and semi-supervised case

Semi-Supervised Clustering via Markov Aggregation

» Influence of k

Experiments

Limitations

Conclusion

Results: Circles dataset

Introduction 00000

» Influence of k

Experiments	
000000000000000000000000000000000000000	

Limitations

Conclusion

Results: Circles dataset



Introduction 00000

»

Influence of k

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion

Results: Circles dataset



Introduction 00000

Experiments

Limitation: 00 Conclusion

» Influence of k





Semi-Supervised Clustering via Markov Aggregation

» Influence of k

Experiments

Limitations 00 Conclusion

Results: Iris dataset

Introduction 00000

» Influence of k

Experiments

Limitations

Conclusior

Results: Iris dataset



Introduction 00000

»

Influence of k

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 00 Conclusion

Results: Iris dataset



Introduction 00000

Experiments

Limitations

Conclusion

» Influence of k





Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitation: 00 Conclusior 000



Setup:

Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitation: 00 Conclusior 000



Setup:

* Parameters: k = 20, β varied

Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitation:

Conclusion



Setup:

- * Parameters: k=20, β varied
- * Datasets: selected UCI datasets

Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitation:

Conclusion



Setup:

- * Parameters: k=20, β varied
- * Datasets: selected UCI datasets

Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitation: 00 Conclusior 000



Setup:

- * Parameters: k = 20, β varied
- * Datasets: selected UCI datasets

Observations:

Semi-Supervised Clustering via Markov Aggregation

» Influence of beta

Experiments

Limitations

Conclusion



Setup:

- * Parameters: k = 20, β varied
- * Datasets: selected UCI datasets

Observations:

* Influence on accuracy for unsupervised and semi-supervised case

Introduction 00000

Experiments

Limitations

Conclusior

» Influence of beta

Results: semi-supervised (20% labeled data)

→ CoMaC (0% labels) ····×··· CoMaC (20% labels)


Experiments

- * Experimental setup
- * Influence of the hyperparameters
- * Side-information
- * Few labeled classes

Semi-Supervised Clustering via Markov Aggregation

» Side-information

Experiments

Limitation 00 Conclusior 000



Semi-Supervised Clustering via Markov Aggregation

» Side-information

Experiments

Limitation 00 Conclusior 000



Setup:

* Fraction of labeled data: 0, 10, 20, 30%

Semi-Supervised Clustering via Markov Aggregation

» Side-information

Experiments

Limitations

Conclusion



Setup:

 $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%

* $\beta=0.5,\,k=20$

¹Marek Smieja and Bernhard C. Geiger. "Semi-supervised cross-entropy clustering with information bottleneck constraint". In: *Information Sciences* 421 (Dec. 2017), pp. 254–271

» Side-information

Setup:

- $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%
- * $\beta=0.5,~k=20$
- \ast Comparison to state-of-the-art-algorithms ¹

 $^1 {\rm Smieja}$ and Geiger, "Semi-supervised cross-entropy clustering with information bottleneck constraint"

Experiments

Limitation: 00 Conclusior 000



» Side-information

Setup:

 $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%

* $\beta=0.5,~k=20$

st Comparison to state-of-the-art-algorithms 1

Observations:

Experiments

Limitation 00 Conclusior



 $^{^1 {\}rm Smieja}$ and Geiger, "Semi-supervised cross-entropy clustering with information bottleneck constraint"

» Side-information

Setup:

- $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%
- * $\beta=0.5,~k=20$
- st Comparison to state-of-the-art-algorithms 1

Observations:

 \ast Influence of labeled data on accuracy

¹Smieja and Geiger, "Semi-supervised cross-entropy clustering with information bottleneck constraint"

xperiments

Limitations

Conclusior 000



» Side-information

Setup:

- $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%
- * $\beta=0.5,~k=20$
- \ast Comparison to state-of-the-art-algorithms 1

Observations:

- * Influence of labeled data on accuracy
- * Performance compared to other methods

Experiments

Limitation: 00 Conclusion



¹Smieja and Geiger, "Semi-supervised cross-entropy clustering with information bottleneck constraint"

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations

Conclusion







Experiments

Limitation: 00 Conclusion

Results: Details





Experiments

Limitation:

Results: Details

Conclusior 000





Experiments

Limitation: 00

Results: Details

Conclusion



Experiments

- * Experimental setup
- * Influence of the hyperparameters
- * Side-information
- * Few labeled classes

Semi-Supervised Clustering via Markov Aggregation

» Few labeled classes

Experiments

Limitatior 00 Conclusior 000



Setup:

page 20



» Few labeled classes

Experiments

Limitation 00 Conclusior 000



Setup:

 $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%



» Few labeled classes

Experiments

Limitation 00 Conclusior 000



- $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%
- * Samples taken from 2 classes



» Few labeled classes

Experiments

Limitation 00 Conclusion



- $\ast\,$ Fraction of labeled data: 0, 10, 20, 30%
- * Samples taken from 2 classes
 - * Classes are selected randomly



» Few labeled classes

Experiments

Limitation 00 Conclusior 000



- * Fraction of labeled data: 0, 10, 20, 30%
- \ast Samples taken from 2 classes
 - * Classes are selected randomly
 - $\ast~$ Must cover > 30% of total data



» Few labeled classes

Experiments

Limitation 00 Conclusior 000



Setup:

- \ast Fraction of labeled data: 0, 10, 20, 30%
- $\ast\,$ Samples taken from 2 classes
 - * Classes are selected randomly
 - * Must cover > 30% of total data

Observations:

» Few labeled classes

Experiments

Limitation: 00 Conclusior 000



Setup:

- \ast Fraction of labeled data: 0, 10, 20, 30%
- \ast Samples taken from 2 classes
 - * Classes are selected randomly
 - * Must cover > 30% of total data

Observations:

 $\ast\,$ Influence of labeled data on accuracy



Experiments

Results: Few vs all classes

Conclusion

» Few labeled classes



»

Semi-Supervised Clustering via Markov Aggregation 000000000

Experiments

Limitations

Results: Details

Conclusior 000

Few labeled classes





Experiments

Limitations

Results: Details

Conclusior

» Few labeled classes





Experiments

Limitations

Results: Details

Conclusior

» Few labeled classes



Limitations

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 0• Conclusior

» Limitations

A Erroneous partition-level side information

* Sensitive to noisy constraints

Semi-Supervised Clustering via Markov Aggregation

Experiments

Limitations 0• Conclusior

» Limitations

Erroneous partition-level side information

 Sensitive to noisy constraints

A Non exhaustive pairwise constraints

Experiments

Limitations 0• Conclusion

» Limitations

A Erroneous partition-level side information

* Sensitive to noisy constraints

A Non exhaustive pairwise constraints

 $* \hspace{0.1 in} ext{if} \hspace{0.1 in} (x,x') \in \mathcal{M} \hspace{0.1 in} ext{and} \hspace{0.1 in} (x,x'') \in \mathcal{N} ext{, then also } x' \hspace{0.1 in} ext{and} \hspace{0.1 in} x'' \hspace{0.1 in} ext{should not link}$

Experiments

Limitations 0• Conclusion

» Limitations

A Erroneous partition-level side information

* Sensitive to noisy constraints

A Non exhaustive pairwise constraints

- $* \hspace{0.1 in}$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{N}$, then also x' and x'' should not link
- * if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{M}$, then also x' and x'' should link

» Limitations

A Erroneous partition-level side information

 \ast Sensitive to noisy constraints

A Non exhaustive pairwise constraints

- $* \;$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{N},$ then also x' and x'' should not link

A Negative influence of cannot-link constraints.

Limitations 0• Conclusion

» Limitations

▲ Erroneous partition-level side information

 \ast Sensitive to noisy constraints

A Non exhaustive pairwise constraints

- $* \hspace{0.1 in}$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{N}$, then also x' and x'' should not link

A Negative influence of cannot-link constraints.

* Moderate numbers of cannot-link constraints

Limitations ○● Conclusior 000

» Limitations

▲ Erroneous partition-level side information

 \ast Sensitive to noisy constraints

A Non exhaustive pairwise constraints

- $* \;$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{N},$ then also x' and x'' should not link
- $* \hspace{0.1 in}$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{M}$, then also x' and x'' should link

Megative influence of cannot-link constraints.

- * Moderate numbers of cannot-link constraints
- * Initial coloring: cannot-link constraints ightarrow assign to the first cluster available

» Limitations

▲ Erroneous partition-level side information

 \ast Sensitive to noisy constraints

A Non exhaustive pairwise constraints

- $* \;$ if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{N},$ then also x' and x'' should not link
- * if $(x,x') \in \mathcal{M}$ and $(x,x'') \in \mathcal{M}$, then also x' and x'' should link

Megative influence of cannot-link constraints.

- * Moderate numbers of cannot-link constraints
- * Initial coloring: cannot-link constraints ightarrow assign to the first cluster available
- * Problem for few classes: constrained data pairs are forced to stay in place

Conclusion





Experiments

Limitations 00 Conclusior ○●○

$\$ Unsupervised clustering via Markov aggregation \longrightarrow accept pairwise constraints



» Conclusion

Experiments

Limitations 00 Conclusior ○●○

♀ Unsupervised clustering via Markov aggregation → accept pairwise constraints
▲ Lowers sensitivity to hyperparameters


Experiments

Limitations 00 Conclusion 000

» Conclusion

 $\$ Unsupervised clustering via Markov aggregation \longrightarrow accept pairwise constraints

- Lowers sensitivity to hyperparameters
- i Learns non-linear decision boundaries



Limitations 00 Conclusion

» Conclusion

$\$ Unsupervised clustering via Markov aggregation \longrightarrow accept pairwise constraints

- Lowers sensitivity to hyperparameters
- i Learns non-linear decision boundaries
- Competes with state-of-the-art semi-supervised clustering techniques



Limitations 00 Conclusion

» Conclusion

$\$ Unsupervised clustering via Markov aggregation \longrightarrow accept pairwise constraints

- Lowers sensitivity to hyperparameters
- **d** Learns non-linear decision boundaries
- Competes with state-of-the-art semi-supervised clustering techniques

Full-text: https://arxiv.org/abs/2112.09397



Limitations 00 Conclusion 000

» Conclusion

$\$ Unsupervised clustering via Markov aggregation \longrightarrow accept pairwise constraints

- Lowers sensitivity to hyperparameters
- Learns non-linear decision boundaries
- Competes with state-of-the-art semi-supervised clustering techniques
- Full-text: https://arxiv.org/abs/2112.09397

</> https://github.com/stegsoph/Constrained-Markov-Clustering.git

Thank you for your attention!