Linkage Criteria for Agglomerative Clustering based on Counterfactual Distances

Georgios Vardakas, Antonia Karra, Evaggelia Pitoura, Aristidis Likas
Department of Computer Science and Engineering, University of Ioannina, Greece

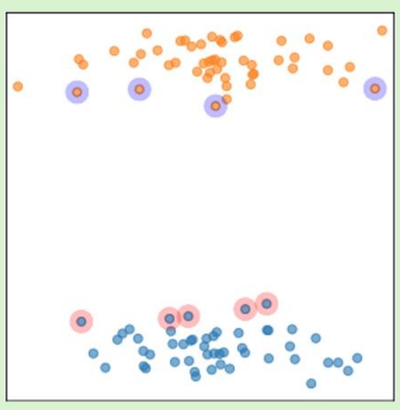
Abstract

• **Motivation:** Standard agglomerative clustering linkages may inadequately capture true cluster separation, particularly in the presence of noise or heterogeneous cluster densities.
• **Key Idea:** We introduce counterfactual distances, where the counterfactual of a point is its nearest point in another cluster, and show that mutual counterfactuals effectively characterize cluster borders and inter-cluster frontiers.
• **Contribution:** We propose the Mutual Counterfactual (MCF) and Iterative Mutual Counterfactual (IMCF) linkage criteria for agglomerative clustering, and demonstrate improved performance over conventional linkages in comparative experiments.

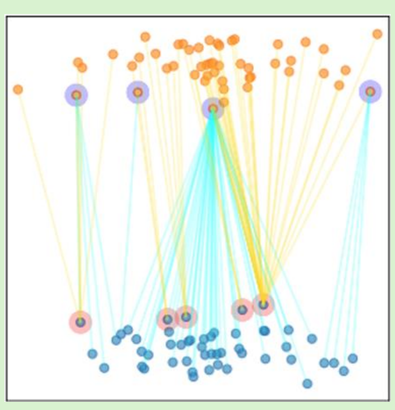
1. Introduction

• **Agglomerative clustering** is a fundamental unsupervised learning technique that builds a nested cluster structure (dendrogram) by iteratively merging similar clusters, enabling multi-resolution data exploration.
• **Linkage criterion** selection is critical in agglomerative clustering, as classical approaches (single, average, complete, centroid, Ward) balance noise sensitivity, cluster geometry, and compactness differently.
• **Existing linkage methods** often fail to reliably capture true inter-cluster separation, especially in the presence of noise, overlapping regions, or clusters with different densities and shapes.
• **Cluster separation** is primarily determined by border points, yet classical methods either ignore cluster frontiers or rely on oversimplified representations (e.g., a single pair of nearest points in single linkage).
• **In this work**, we introduce counterfactuals as a principled way to identify **cluster frontiers**, leveraging nearest cross-cluster points to better characterize inter-cluster boundaries.
• Building on this insight, we propose **counterfactual-based linkage criteria** that operate directly on pairwise distances and provide more reliable cluster separation measures for agglomerative clustering.

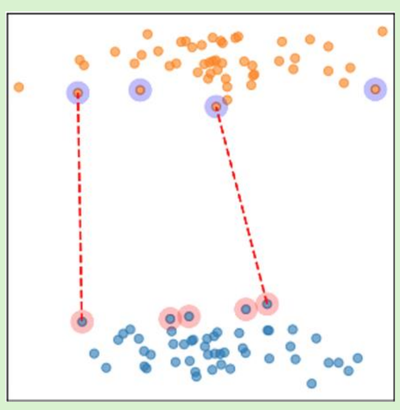
CFE solutions - Extensions



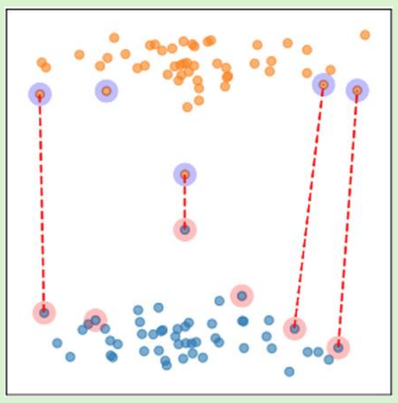
(a) Counterfactual points (marked circles) are located on the cluster borders.



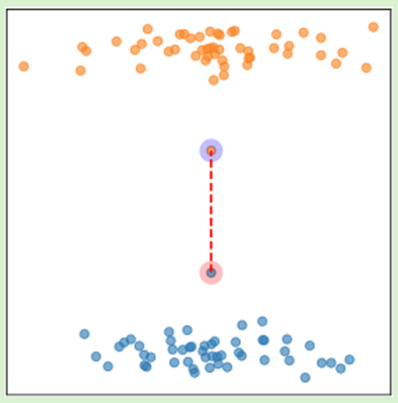
(b) Counterfactual links connecting each point with its corresponding counterfactual point.



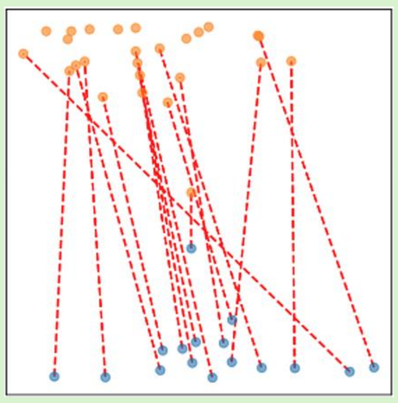
(c) Mutual counterfactual links.



(d) Mutual counterfactual links for a synthetic dataset with noise.



(e) Mutual counterfactual links for another synthetic dataset with noise, but with higher cluster separation.



(f) Iterative mutual counterfactual links.

2. Counterfactual Explanations

• **Counterfactual explanations** provide local, example-based insights by identifying minimal changes to a data point that alter a model's outcome, and are widely used in classification tasks.
• **In classification**, counterfactuals are computed by optimizing a trade-off between proximity to the original instance and achieving a desired target class, using distance and loss functions.
• **The counterfactual framework extends naturally to clustering**, where a counterfactual of a data point is defined as the closest instance that would be assigned to a different cluster.
• When counterfactuals are restricted to dataset points, they correspond to nearest cross-cluster neighbors and consistently lie on cluster borders, effectively defining cluster frontiers.
• Distances between points and their counterfactuals provide informative measures of **inter-cluster separation**, forming the **basis for the counterfactual-based linkage criteria**.

4. Experiments

TABLE II: Clustering performance (NMI score) for real datasets.

Dataset	MCF	single	IMCF	average
Digits	0.77	0.77	**0.82**	0.80
Iris	0.72	0.72	**0.78**	0.75
Olivetti Faces	**0.68**	0.67	0.70	**0.74**
Pendigits	**0.65**	0.58	**0.72**	0.69
Seeds	0.02	0.02	**0.70**	0.55
Waveform-v1	1.00	1.00	**1.00**	0.75
Wine	0.02	0.02	**0.82**	0.02

TABLE III: Clustering performance (NMI score) for synthetic datasets.

Dataset	MCF	single	IMCF	average
2d-20c-no0	0.95	**0.97**	**0.99**	0.98
2d-4c-no4	0.67	0.67	**0.99**	0.67
3-spiral	1.00	1.00	**0.07**	0.01
DS-850	**0.85**	0.84	**0.98**	0.93
complex8	**0.86**	0.85	**0.73**	0.67
compound	0.80	0.80	**0.84**	0.81
cure-t0-2000n-2D	1.00	1.00	**1.00**	0.60
donut2	**0.90**	0.02	0.17	**0.27**
ds2c2sc13	**0.96**	0.94	**0.90**	0.82
flame	0.02	0.02	**0.94**	0.62
jain	**0.25**	0.09	0.69	0.69
longsquare	0.91	0.91	**0.98**	0.91
triangle2	0.85	0.85	**0.97**	0.89

5. Conclusions

• We introduced counterfactual distance–based measures for quantifying cluster separation and showed that counterfactuals effectively characterize cluster frontiers.
• Mutual counterfactual distances provide a representative and robust estimate of inter-cluster separation.
• We proposed the Mutual Counterfactual and Iterative Mutual Counterfactual linkage criteria.
• The proposed linkages are distance-agnostic, require only the pairwise distance matrix, and do not assume access to the original data vectors.
• Experimental results within a standard agglomerative clustering framework demonstrate improvements over classical single and average linkage methods.

Counterfactuals for
Clustering and Deep Clustering:
Explainability, Fairness and Quality
Project Number: 15940
Hellenic Foundation for Research & Innovation (H.F.R.I)
"Basic Research Funding (Horizontal Support for all Sciences)"
National Recovery and Resilience