**A study of clustering**

Eugen-Richard Ardelean1, Raluca Laura Portase1

1Department of Computer Science, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

**\* Corresponding authors:** [ardeleaneugenrichard@gmail.com](mailto:ardeleaneugenrichard@gmail.com), <placehoder-raluca>

**ORCID Author IDs:**

Eugen-Richard Ardelean: 0000-0002-0098-4228

Raluca Laura Portase: 0000-0002-8985-4728

**Abstract:** TBWritten

**Keywords:** clustering, TBAdded

# Introduction

# Materials and Methods

## Clustering algorithms

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | Year | Code | ? |  |  |  |

## Clustering validation scores

Internal vs external – discussion from edging distance

Score summary (from edging distance / ISBM)

# Results

## Simple clustering

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

## Overlap

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

## Imbalance

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

## Non-convexity

Internal metrics are not good, use external – discussion from edging distance

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

## Sample scalability

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

## Feature scalability

* Data table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Name | #Samples | #Features | #Clusters | Provenience |  |
| D1 |  |  |  |  |  |  |

* Data plot
* Plot result – hierarchical

# Discussion