

# Progetto A\_Clus - Documentazione

A cura di: Vito Stefano  
Lorenzo Gelao

3 aprile 2025

## Indice

<b>1</b>	<b>Introduzione</b>	<b>3</b>
1.1	Agglomerative Clustering . . . . .	3
<b>2</b>	<b>Guida all'installazione</b>	<b>3</b>

## 1 Introduzione

Il progetto in questione verte sull'argomento dell'*Agglomerative Clustering*, una tecnica di clusterizzazione basata sui metodi Single distance e Average Distance.

Il progetto in questione è suddiviso in una parte client e una server, che comunicando tra loro, generano il dendrogramma, permettendo inoltre di visualizzare e memorizzare tali risultati o di caricarne dei precedenti.

È inoltre possibile visualizzare nuovamente dei file caricati in passato per visualizzare i cluster e i dendrogrammi associati.

### 1.1 Agglomerative Clustering

L'algoritmo di clustering utilizzato in tale progetto, come si può desumere dal nome, sfrutta il concetto di clustering agglomerativo. In pratica, tratta ciascun cluster in maniera separata dagli altri, unendo progressivamente quelli più vicini, in base a due criteri principali, nel nostro caso.

Il principale vantaggio rispetto ad altri algoritmi di clustering, come il k-means, è che in questo modo non è necessario specificare in anticipo la quantità di cluster da analizzare.

I criteri utilizzati nel progetto A-CLus sono i seguenti:

1. **Single-Link:** tale criterio determina la distanza minima tra i punti dei vari cluster

$$D(C1, C2) = \min_{(t1 \in C1, t2 \in C2)} (dist(t1, t2)) \quad (1)$$

Durante l'anno accademico 2024/2025, l'oggetto di ricerca è stato incentrato su "H-CLUS", una piattaforma con architettura client-server dedicata all'analisi dei dati mediante algoritmi di clustering gerarchico agglomerativo. La componente server esegue le operazioni di clustering impiegando le metodologie Single Link o Average Link per il calcolo delle distanze inter-cluster e la successiva costruzione del dendrogramma. L'applicativo client, implementato in linguaggio Java, offre agli utenti diverse funzionalità: il caricamento o la creazione di istanze HierarchicalClusterMiner, la rappresentazione grafica del dendrogramma e l'archiviazione dei risultati per analisi successive. È inoltre disponibile la funzione di importazione di file precedentemente salvati, consentendo agli utenti di riesaminare i cluster e i relativi dendrogrammi.

## 2 Guida all'installazione