

Progetto A_Clus - Documentazione

A cura di: Vito Stefano
Lorenzo Gelao

2 aprile 2025

Indice

1	Introduzione	3
1.1	Agglomerative Clustering	3
1.1.1	Framework Analitico	3
2	Guida all'installazione	4
2.1	Installazione JDK	4
2.2	Variabili di ambiente	4
2.3	Installazione MySQL	5
3	Eseguire A-CLus Base	8

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.

1.1.1 Framework Analitico

La tecnica di clustering agglomerativo implementata nel sistema "A-CLUS" costituisce una metodologia avanzata per l'identificazione di correlazioni latenti nei dataset.

Diversamente da tecniche alternative come il k-means, l'approccio agglomerativo elimina la necessità di predefinire il numero di raggruppamenti. La procedura inizializza ciascun elemento come cluster individuale e procede con l'unificazione sequenziale dei cluster con maggiore affinità, applicando metodologie quali:

Con il procedere delle aggregazioni tra cluster, si sviluppa un dendrogramma che illustra la gerarchia delle aggregazioni. Il procedimento continua fino al raggiungimento di un cluster unificato o fino a una soglia di profondità stabilita dall'utilizzatore. La visualizzazione mediante dendrogramma rende la metodologia particolarmente comprensibile, facilitando l'esplorazione strutturale dei dati a diversi livelli di granularità. Inoltre, questa strategia non è influenzata dalla configurazione iniziale dei centroidi, riducendo così la probabilità di risultati inconsistenti e garantendo una rappresentazione più accurata dell'organizzazione interna del dataset. RiprovaClaude può commettere errori. Verifica sempre le risposte con attenzione.

2 Guida all'installazione

Prima di essere in grado di eseguire il programma, è necessario eseguire il file `risorse.bat` contenuto nella cartella Risorse.

Una volta eseguito il file, si aprirà una pagina di Powershell e seguirà un download.



Figura 1: Schermata di download

Al termine del download del file compresso delle risorse, verranno estratti i file necessari per l'esecuzione del programma.

Per il progetto in questione è necessario installare il Java Developer Kit (JDK) nella versione 22.0.1 e il software di gestione del database MySQL nella sua versione 8.0.39.

La prima scheda di installazione che apparirà è quella del JDK.

Verrà chiesto se si vuole eseguire l'installazione del JDK tramite permessi di amministratore, cliccare su Sì.

2.1 Installazione JDK

Una volta confermata l'esecuzione come amministratore, si procederà all'installazione del JDK.

Cliccare su **Avanti**, nuovamente **Avanti** e infine, nel caso in cui l'installazione sia andata a buon fine, si dovrà cliccare sul tasto **Chiudi**.

2.2 Variabili di ambiente

è consigliato creare una variabile di ambiente per il JDK. Per farlo, si dovrà cercare la voce **variabili di ambiente** nella barra di ricerca di Windows e cliccare su Modifica le variabili di ambiente per il tuo account.

Successivamente si dovrà cliccare sul pulsante **Variabili di ambiente** e si aprirà una schermata con le variabili di ambiente.

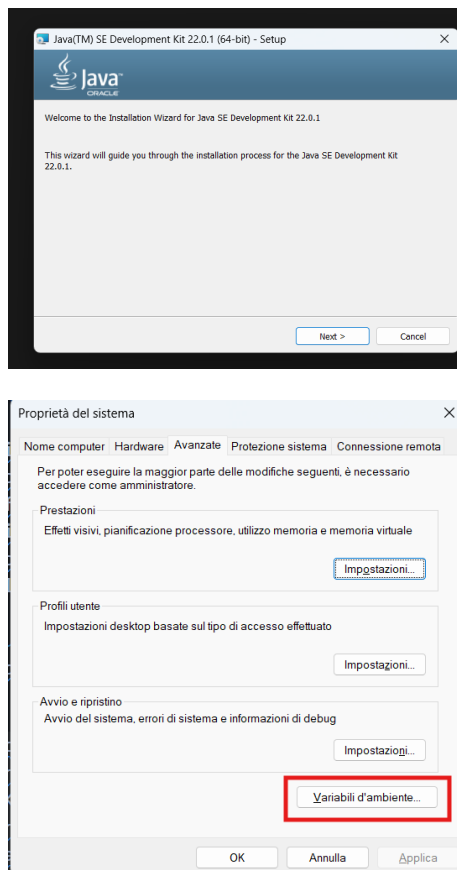


Figura 2: Schermata delle variabili di ambiente

Nella schermata, si dovrà cliccare sulla variabile **PATH** e modificarla in modo da aggiungere il percorso del JDK, inserendo il percorso in cui è stato installato il JDK.

Una volta completata l'installazione del JDK, si procederà automaticamente con l'installazione del MySQL.

2.3 Installazione MySQL

Anche qui sarà necessario fornire i permessi di amministratore, cliccare dunque su Sì.

Schermata iniziale

Dopodiché si aprirà la schermata di installazione di MySQL.

Selezionare il tipo di setup **Full** e cliccare su **Next**. Tale scelta permetterà di installare tutti i componenti necessari per l'esecuzione del programma.

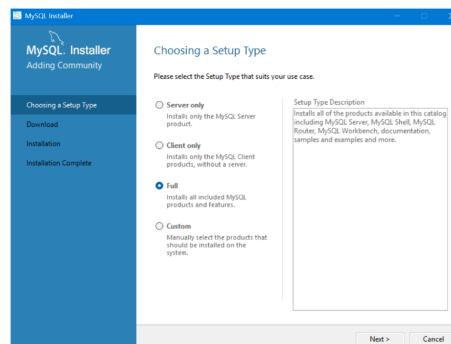


Figura 3: Scelta del setup

Download dei file

Si passerà alla schermata di download dei file necessari per l'installazione, cliccare su **Execute**.

AVVERTENZA

Tale schermata potrebbe richiedere un po' di tempo per il download dei file, a seconda della velocità della connessione internet.

Dopo aver scaricato i file, si dovrà cliccare nuovamente su **Execute**.

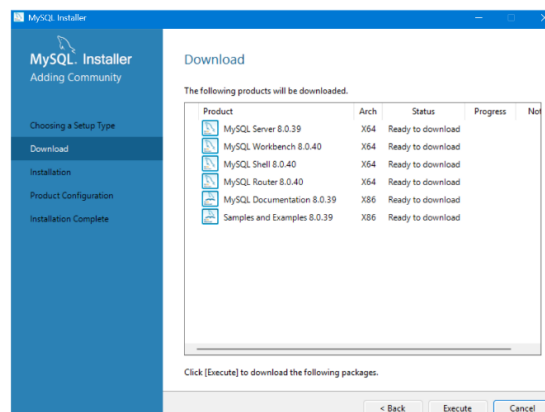


Figura 4: Download MySQL

Configurazione del account MySQL

Una volta terminato il download, si procederà con la configurazione dell'account MySQL. Si dovrà quindi scegliere una password, necessaria per accedere al database.

Si giungerà infine alla schermata di applicazione della configurazione del database. Cliccare su **Execute** per applicare la configurazione.

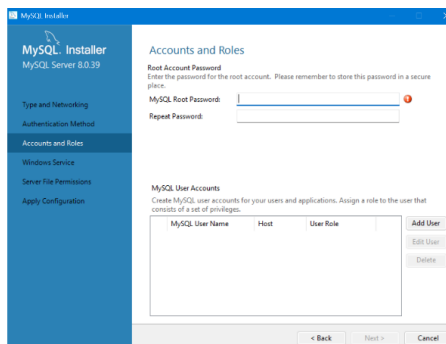


Figura 5: Configurazione account MySQL

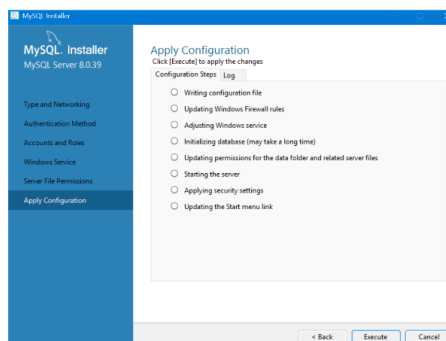


Figura 6: Completamento installazione

Una volta terminata anche l'installazione del MySQL, si dovrà cliccare su **Finish** per chiudere l'installler.

AVVERTENZA

Per un corretto funzionamento sia del JDK che del MySQL, è consigliato riavviare il computer.

Schermata finale

Al termine dell'installazione il programma `resources.bat` mostrerà la seguente schermata.



Figura 7: Completamento installazione

3 Eseguire A-CLus Base

Avvio dell'applicazione

Per inizializzare correttamente la versione standard di H-CLUS, attenersi alla seguente sequenza operativa:

1. Accedere alla directory `A-CLus_Base/Bat`
2. Eseguire il file `Start_Server.bat` facendo un doppio click su di esso, assicurando che il server sia in esecuzione
3. Dopo aver verificato il corretto funzionamento del server, eseguire il file `Start_Client.bat` facendo un doppio click su di esso

Avvertenza

È necessario mantenere attivo il terminale del server per garantire la comunicazione tra le componenti client e server.

Avvio dell'ambiente di sviluppo

Per aprire correttamente il codice sorgente di A-CLus, è necessario seguire le seguenti operazioni nell'ordine in cui sono presentate:

1. Aprire il progetto `Server`
2. Incorporare nelle librerie del progetto `Server` il file `mysql.connector.jar` presente nella cartella `A-CLus/Risorse`
3. Avviare `MultiServer.java`
4. Controllare che il server sia in esecuzione correttamente
5. Aprire il progetto `MainTest.java`