
PROGETTO K-MEANS

DOCUMENTAZIONE

ANNO ACCADEMICO: 2022-2023

CORSO: METODI AVANZATI DI PROGRAMMAZIONE

DOCENTE: ANNALISA APPICE

A CURA DI :

AGNESE D'ADDABBO - GAMAL ELMETKANANY - MATTEO CAPOCCHIANO

Indice

1 Introduzione

2 Funzionamento Algoritmo

2.1 Vantaggi e Svantaggi Algoritmo

3 Implementazione del sistema

3.1 Introduzione al sistema

3.2 Implementazione del Sistema Base

3.3 Implementazione del Sistema con Estensione

4 Diagrammi UML

4.1 Diagrammi dei package e delle classi parte Server versione base

4.2 Diagrammi dei package e delle classi parte Client versione base

4.3 Diagrammi dei package e delle classi parte Client estesa

5. Manuale Utente

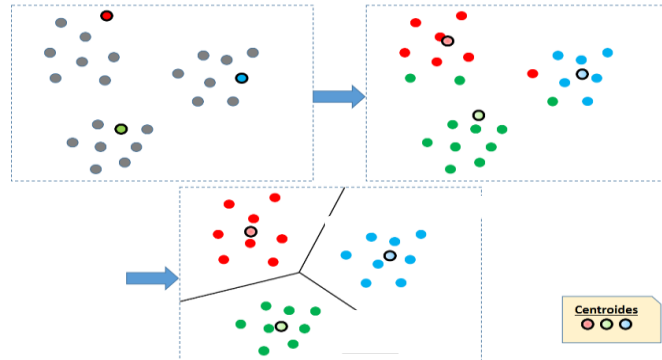
5.1 Test per versione base

5.2 Test per versione estesa

INTRODUZIONE AL PROGETTO

L'analisi dei dati è un elemento fondamentale nello studio e nell'interpretazione delle informazioni estratte da un database.

Uno degli strumenti chiave



in questo processo è la **clusterizzazione**, tecnica basata sulla capacità di identificare e categorizzare pattern simili nei dati, raggruppandoli in cluster o gruppi omogenei. Questo approccio consente di semplificare la comprensione di dati complessi e allo stesso tempo rivelare relazioni e tendenze nascoste.

Metodo ampiamente utilizzato per la clusterizzazione è ***l'algoritmo K-Means***. Quest'ultimo ha l'obiettivo di suddividere un insieme di dati in gruppi omogenei chiamati cluster. L'utente specifica a priori il numero di cluster desiderato, rappresentato da K , e l'algoritmo identifica in modo automatico tali gruppi. I cluster formati da questo algoritmo sono spesso di forma globulare e mantengono una densità costante, garantendo così che i dati all'interno di ciascun cluster siano simili tra loro, mentre i dati tra cluster diversi siano diversificati.

L'uso dell'algoritmo K-Means in combinazione con servizi di database come MySQL rappresenta un potente strumento per l'analisi dei dati, consentendo di estrarre informazioni significative da tabelle di dati complessi. Questo approccio unisce la potenza della clusterizzazione con la gestione efficiente dei dati tramite servizi di database, offrendo così una solida base per l'analisi e l'interpretazione delle informazioni.

FUNZIONAMENTO ALGORITMO

Come già annunciato in precedenza l'algoritmo K-Means è un metodo di clustering ampiamente utilizzato che si basa su un processo di raggruppamento dei dati in cluster.

Ma vediamo nel dettaglio il suo funzionamento:

inizialmente, vengono selezionati casualmente K punti come centroidi iniziali dei cluster, rappresentanti il centro di ciascun gruppo. Successivamente, ogni punto dei dati viene assegnato al cluster il cui centroide è più vicino, basandosi su una misura di distanza come la distanza euclidea. Una volta completata questa fase di assegnazione, i centroidi vengono aggiornati calcolando la media delle posizioni dei punti all'interno di ciascun cluster. Questo processo di assegnazione e aggiornamento viene ripetuto fino a quando i centroidi convergono a posizioni stabili o fino al raggiungimento di un numero massimo di iterazioni. Al termine delle iterazioni si otterranno k -cluster

È fondamentale, tuttavia, scegliere in anticipo il numero di cluster, denotato da K . Questa scelta influisce direttamente sulla forma e sulla significatività del raggruppamento dei dati. Una selezione imprecisa di K può portare a risultati meno informativi o addirittura fuorvianti.

VANTAGGI E SVANTAGGI ALGORITMO

K-means è un algoritmo, e come tutti gli algoritmi presenta criticità e pregi.

PREGI:

- Efficienza: dal punto di vista computazionale, richiede risorse limitate ed è in grado di convergere rapidamente;
- Adeguato per dataset di grandi dimensioni;
- Semplice da implementare: poiché si basa sul concetto di assegnazione dei punti ai cluster e sull'aggiornamento dei centroidi, risulta concettualmente agevole.

DIFETTI:

- K-Means è sensibile all'inizializzazione: la scelta iniziale dei centroidi può influenzare il risultato finale; pertanto, l'algoritmo potrebbe convergere in diversi cluster in base alla scelta iniziale.

L'algoritmo richiede la specifica del numero di Cluster: l'utente deve predefinire il numero di cluster K in anticipo, il che potrebbe non essere sempre ovvio.

UN SISTEMA DI CLUSTER K-MEANS BASATO SU JAVA: ARCHITETTURA CLIENT/SERVER

Nel presente documento, esploreremo un sistema di cluster K-Means basato su Java, che offre agli utenti la possibilità di analizzare e utilizzare questo algoritmo per clusterizzare dati da un database tramite un'architettura client/server.

INTRODUZIONE AL SISTEMA

Per comprendere appieno il funzionamento di questo sistema, è essenziale capire la sua struttura e le funzionalità offerte. Il sistema è composto da due componenti principali: il client e il server.

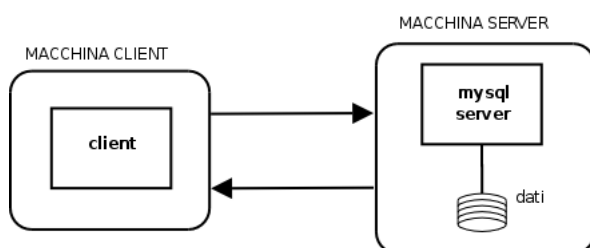
CLIENT:

Il client fornisce un'interfaccia intuitiva agli utenti che desiderano interagire con il sistema. Gli utenti possono comunicare con il server tramite il terminale e ottenere risultati relativi alle operazioni di clustering e al caricamento di file. Questo componente consente agli utenti di personalizzare le impostazioni, come la definizione del numero di cluster (K) e la scelta della tabella da cui estrarre i dati.

SERVER:

Il server rappresenta il cuore pulsante del sistema. È in grado di gestire diverse richieste provenienti da client diversi, garantendo un'esperienza efficiente e reattiva. Le principali funzionalità del server includono:

1. Estrazione dei dati da una tabella di un database.
2. Esecuzione del processo di clustering sui dati estratti.
3. Salvataggio automatico dei risultati del clustering.
4. Lettura dei centroidi dei cluster dai file di salvataggio.
5. Monitoraggio dello stato del server tramite il terminale.



IMPLEMENTAZIONE DEL SISTEMA BASE

Il gruppo ha lavorato diligentemente per sviluppare il sistema base del progetto focalizzato sull'implementazione dell'algoritmo K-Means Clustering. Questo sistema è stato realizzato in Java e si compone di due componenti essenziali: un client e un server. Questa parte della relazione esplorerà dettagliatamente l'implementazione di questa architettura client/server.

Il server, che svolge un ruolo centrale nell'intero sistema, richiede una macchina con un database MySQL in esecuzione. Esso è configurato per essere raggiungibile sulla porta 8080 e ha la capacità di comunicare con più client contemporaneamente. Questa caratteristica è particolarmente importante, poiché consente a diversi utenti di sfruttare le funzionalità offerte in modo simultaneo. Le funzionalità primarie disponibili tramite l'interfaccia del server mediante il client CLI sono le seguenti:

- 1. Scoperta dei Cluster:** Gli utilizzatori sono in grado di interagire con il server, fornendo il nome della tabella archiviata nel database e specificando il numero di cluster da identificare. Questa operazione è di rilievo per l'analisi e la segmentazione dei dati.
- 2. Archiviazione dei Cluster:** I cluster generati vengono automaticamente conservati nella macchina in cui il server è in esecuzione. Ciò garantisce l'integrità dei risultati e la capacità di accedervi in un secondo momento. Inoltre, il server mantiene un registro degli eventuali errori che potrebbero verificarsi durante il processo di salvataggio.
- 3. Estrazione dei Centroidi dei Cluster:** Gli utilizzatori sono in grado di richiedere al server i centroidi dei cluster, specificando il percorso del file in cui sono stati previamente memorizzati. Questa funzionalità è essenziale per l'esame dei dettagli dei cluster generati.

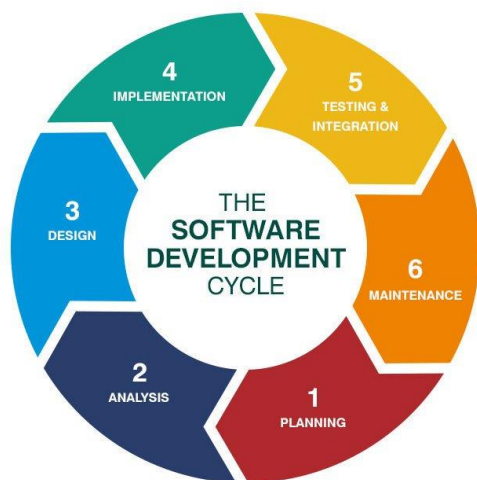
Queste caratteristiche, offerte attraverso il sistema, costituiscono una base solida per l'esecuzione dell'algoritmo K-Means Clustering, consentendo agli utenti di effettuare operazioni di cluster e di analisi dati in modo agevole. Simultaneamente, la possibilità di archiviare i risultati assicura la disponibilità di dati di riferimento futuri. Il server , inoltre, è in grado di gestire gli errori e le eccezioni, notificando l'utente tramite l'interfaccia CLI in caso di interruzioni nella comunicazione o problemi con i dati di input. Questo contribuisce a garantire una migliore esperienza utente e una maggiore affidabilità del sistema.

D'altra parte, ***il client*** da riga di comando consente agli utenti di connettersi a una macchina che ospita un'istanza del server. Nella sua semplice interfaccia a menu, gli utenti possono scegliere tra due opzioni principali: la scoperta di cluster e la lettura da file. Questa facilità d'uso rende il sistema accessibile anche a utenti con meno esperienza tecnica.

In sintesi, il progetto base rappresenta un solido punto di partenza per l'implementazione dell'algoritmo K-Means Clustering, offrendo un'architettura client/server efficiente e funzionale. Grazie a questa configurazione, gli utenti possono eseguire operazioni di cluster e analisi dei dati in modo agevole, mantenendo al contempo la possibilità di salvare i risultati per riferimenti futuri.

IMPLEMENTAZIONE DEL SISTEMA CON ESTENSIONE

E' stato deciso di espandere il progetto base sviluppando un'interfaccia grafica per il client, poiché questa scelta offre numerosi benefici. In primo luogo, abbiamo voluto migliorare l'usabilità complessiva del software. È evidente che un'interfaccia grafica rende l'applicazione molto più accessibile rispetto a un'interfaccia basata su testo, semplificando notevolmente l'interazione con il



programma, infatti un'interfaccia consente all'utilizzatore di vedere i dati e le informazioni presenti nel software in maniera più rapida, migliorando la comprensione e facilitando la manipolazione dei dati stessi.

Inoltre, anche se il software è destinato all'uso interno, non si può trascurare il fatto che l'aggiunta di un'interfaccia grafica consentendo spesso di personalizzare l'aspetto del software in base alle preferenze dell'utente, migliora l'esperienza complessiva. Questa modernizzazione è fondamentale per rimanere competitivi nel mondo delle applicazioni.

Altra ragione per la quale è stata implementata un'interfaccia grafica è stata sicuramente la facilità di manutenzione, in quanto le eventuali modifiche e aggiornamenti possono essere apportati in maniera più agevolata rispetto alla riscrittura di comandi testuali.

Oltre a ciò, l'implementazione di un'interfaccia grafica ha dato l'opportunità al gruppo di ampliare la propria conoscenza che fino a quel momento è stata limitata all'utilizzo del terminale.

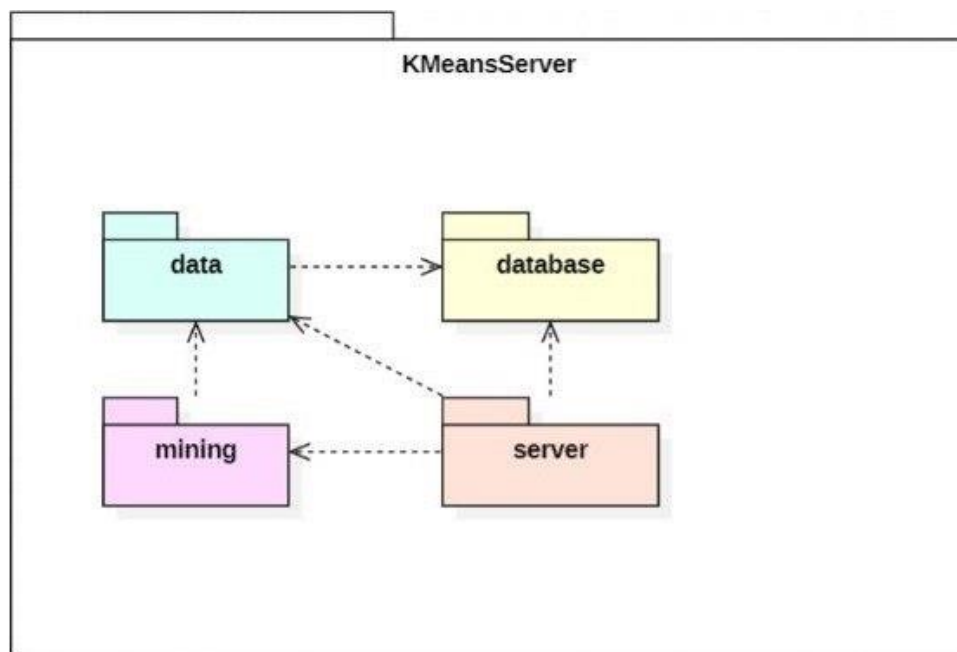
DIAGRAMMI UML

Attraverso l'utilizzo del software StarUML il gruppo ha realizzato i seguenti diagrammi delle classi e dei package.

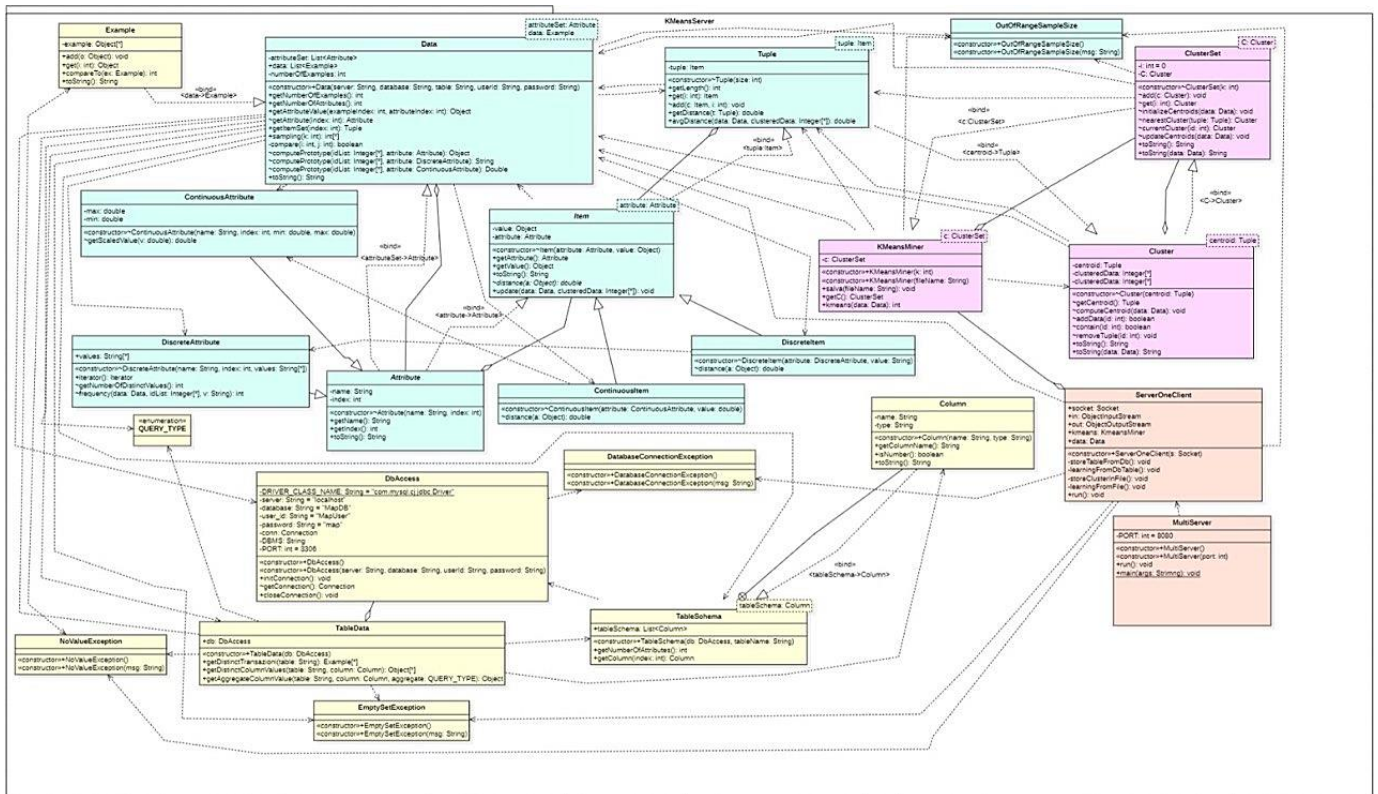
-VERSIONE BASE-

Parte Server

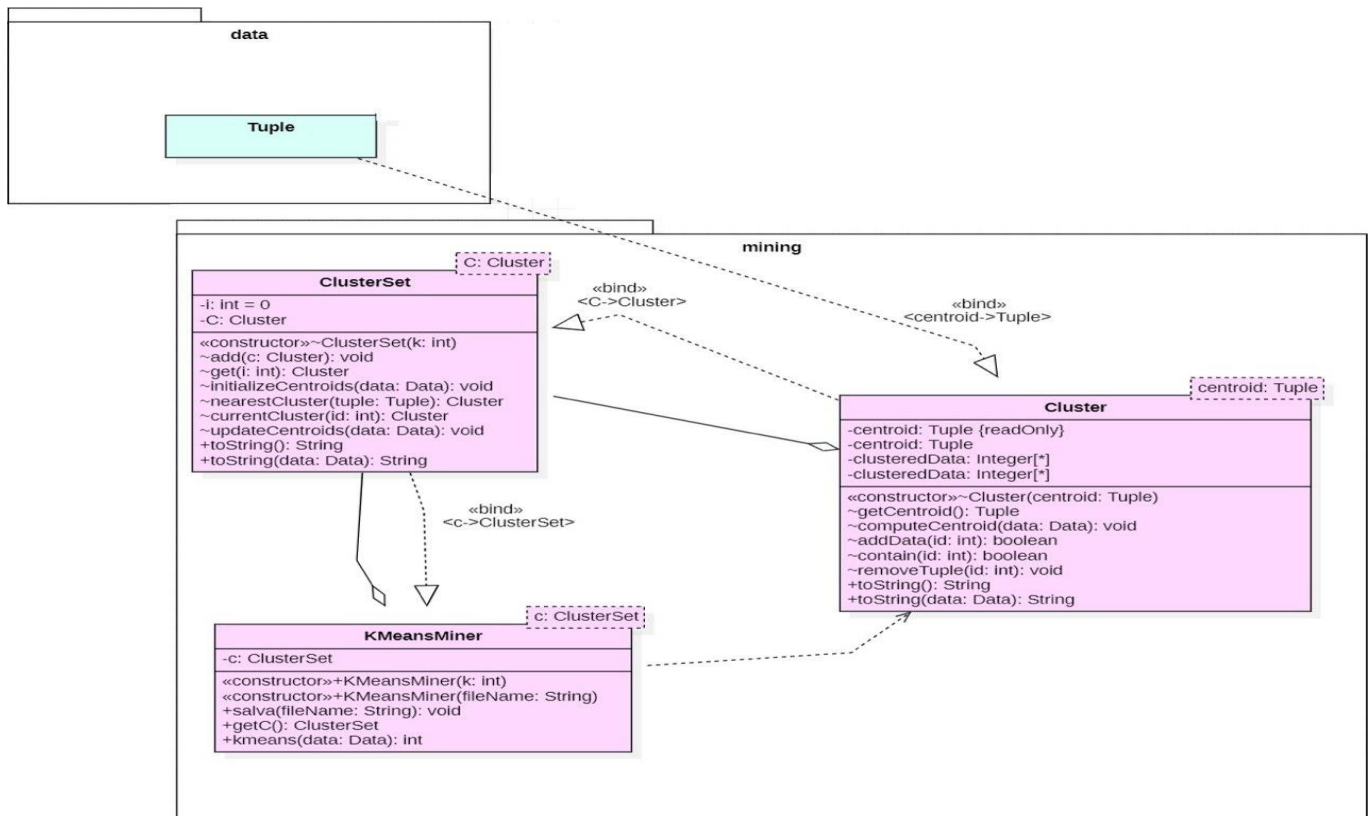
(DIAGRAMMI DEI PACKAGE)

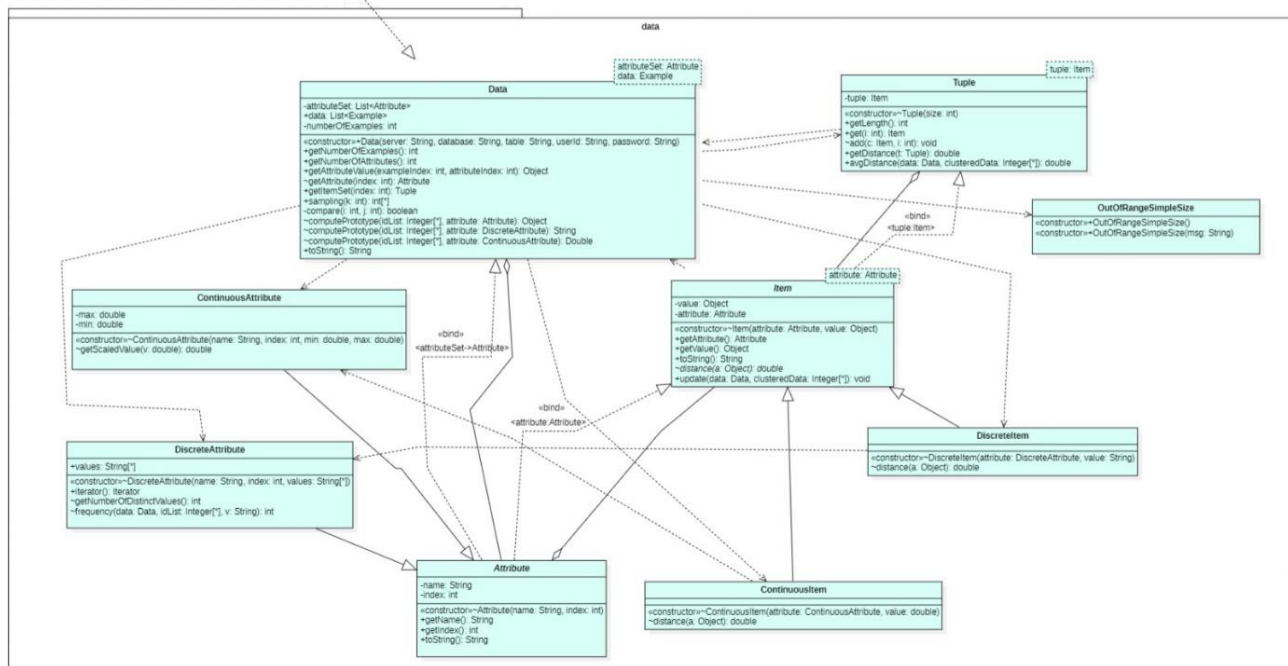
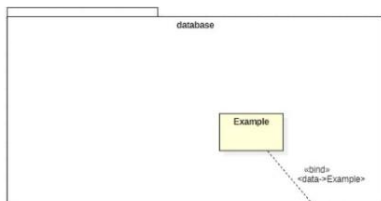


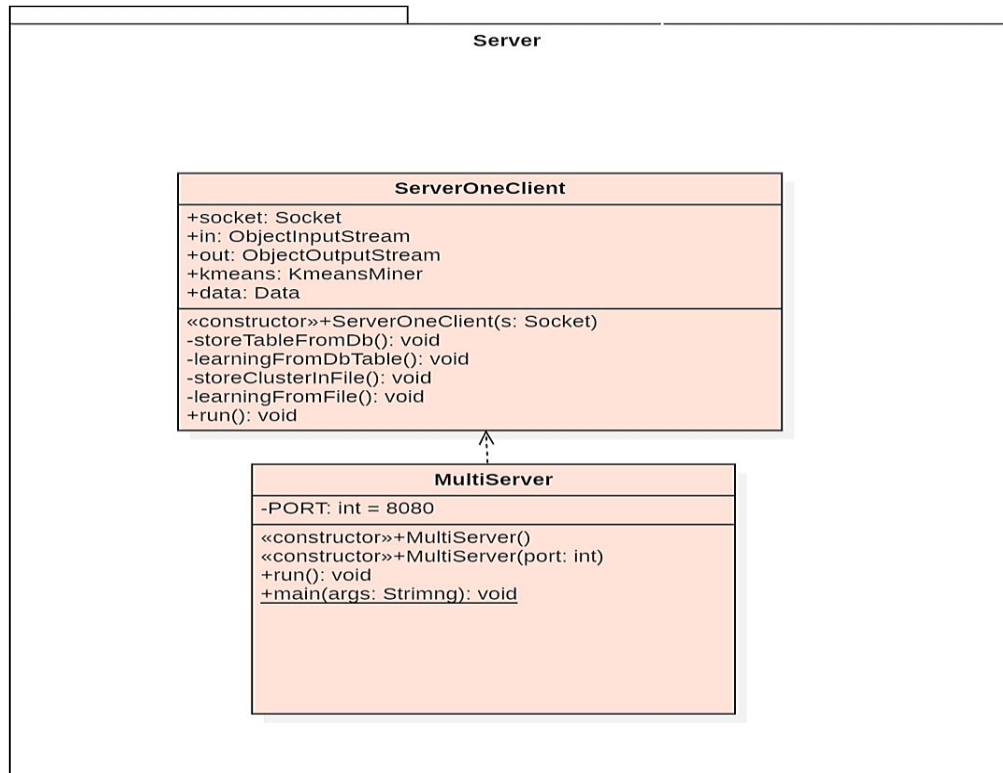
GENERALE:



NEL DETTAGLIO:

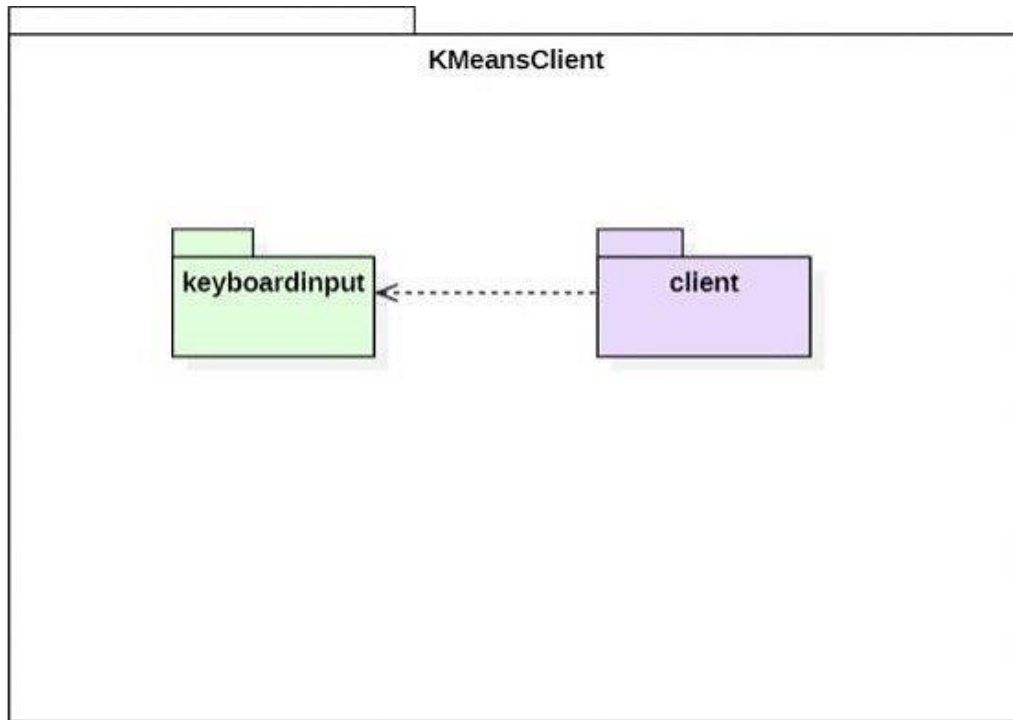




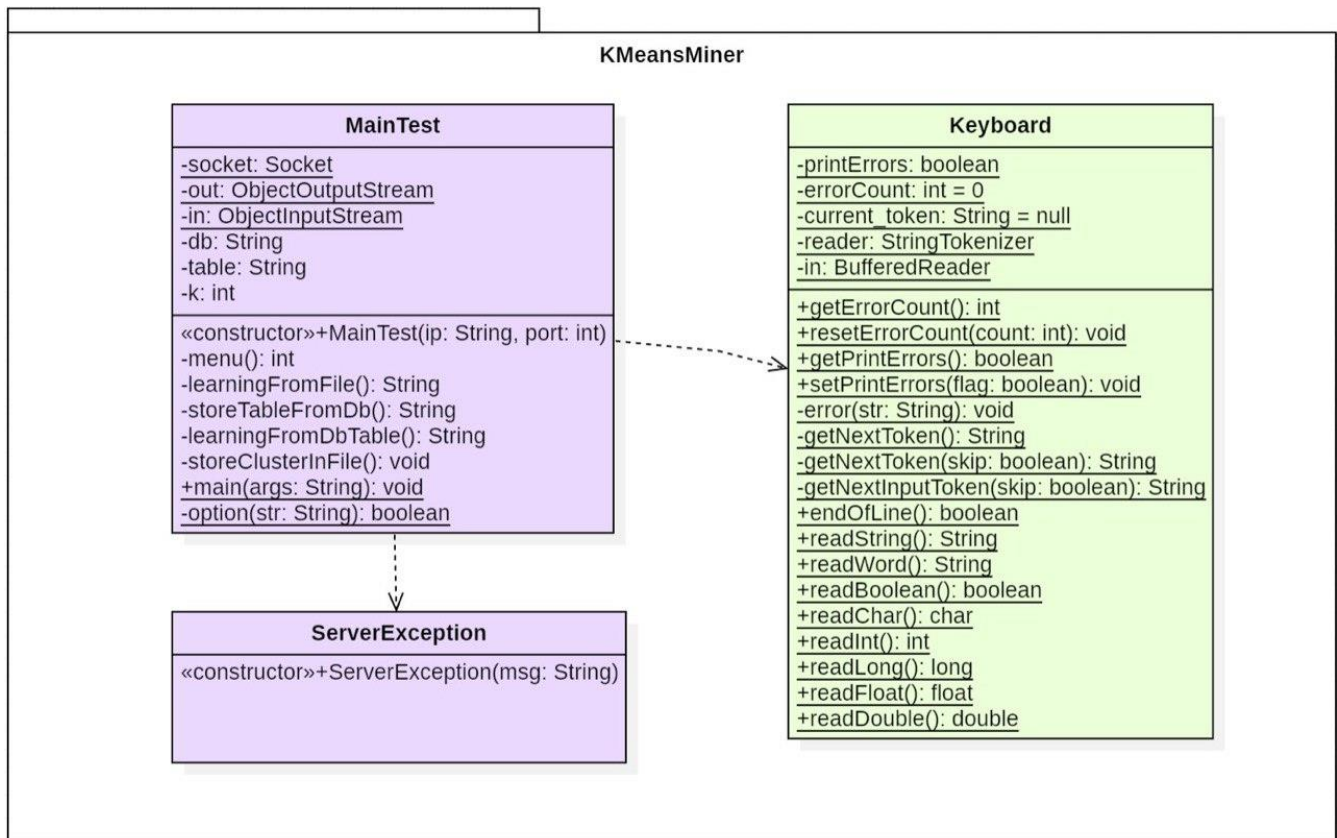


Parte Client

(DIAGRAMMI DEI PACKAGE)

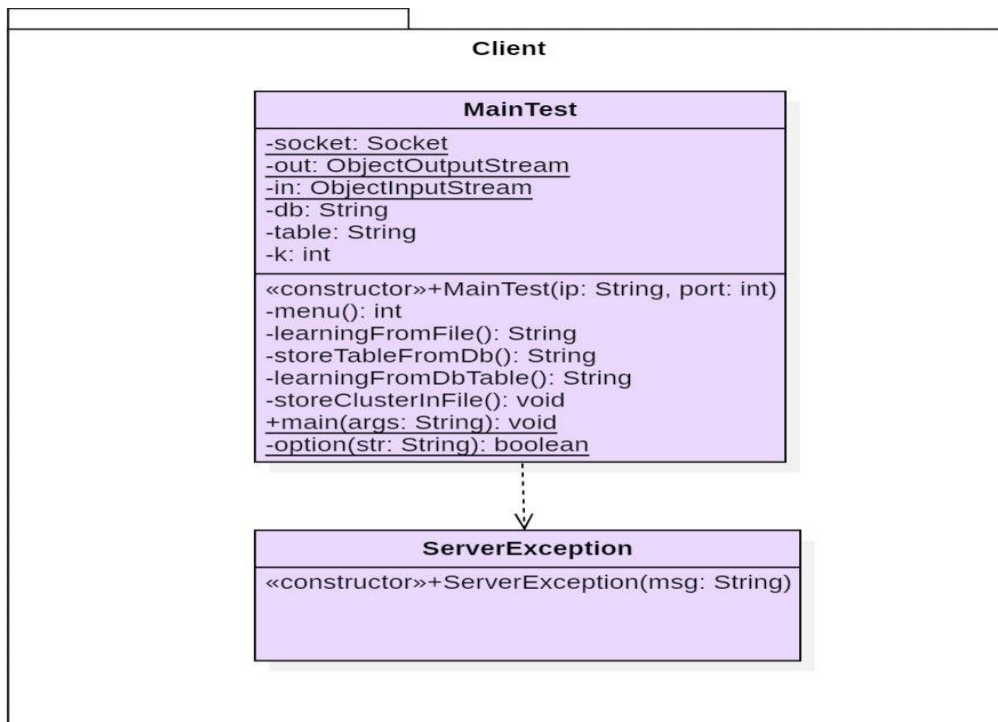


DIAGRAMMI DELLE CLASSI



GENERALE:

NELLO SPECIFICO:



keyboardMiner

Keyboard

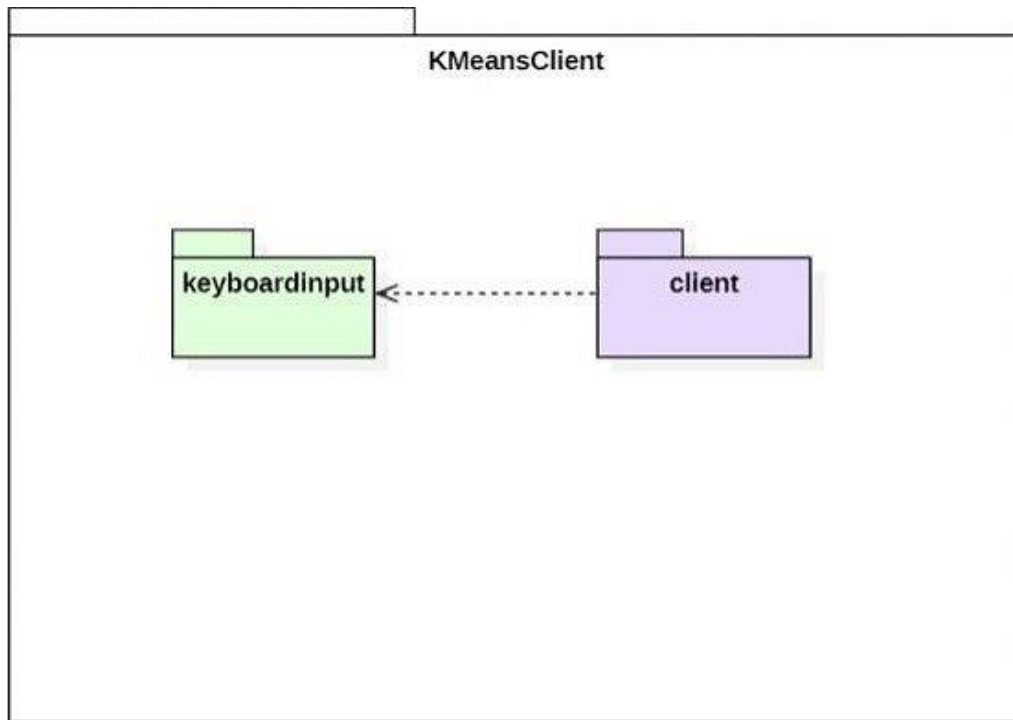
-printErrors: boolean
-errorCount: int = 0
-current_token: String = null
-reader: StringTokenizer
-in: BufferedReader

+getErrorCount(): int
+resetErrorCount(count: int): void
+getPrintErrors(): boolean
+setPrintErrors(flag: boolean): void
-error(str: String): void
-getNextToken(): String
-getNextToken(skip: boolean): String
-getNextInputToken(skip: boolean): String
+endOfLine(): boolean
+readString(): String
+readWord(): String
+readBoolean(): boolean
+readChar(): char
+readInt(): int
+readLong(): long
+readFloat(): float
+readDouble(): double

-VERSIONE ESTESA-

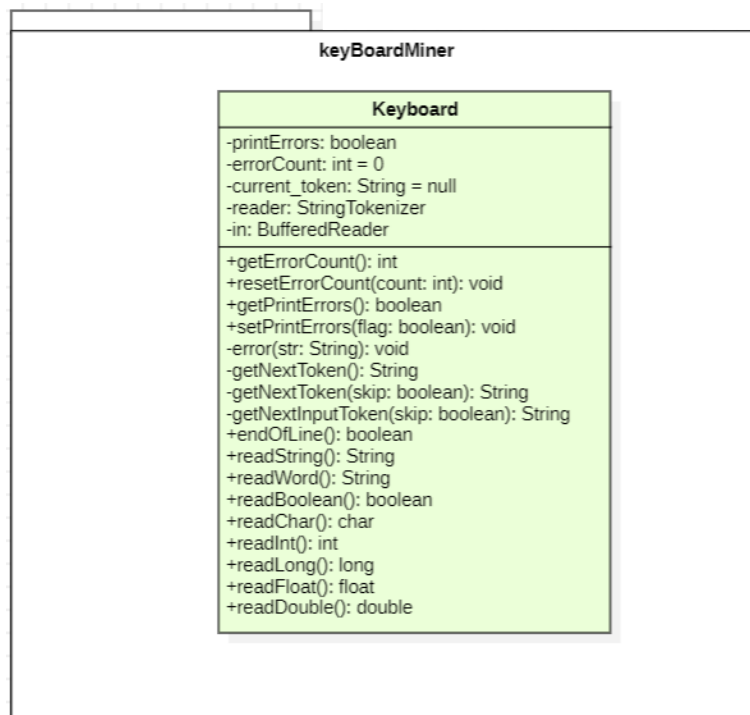
Parte Clinet

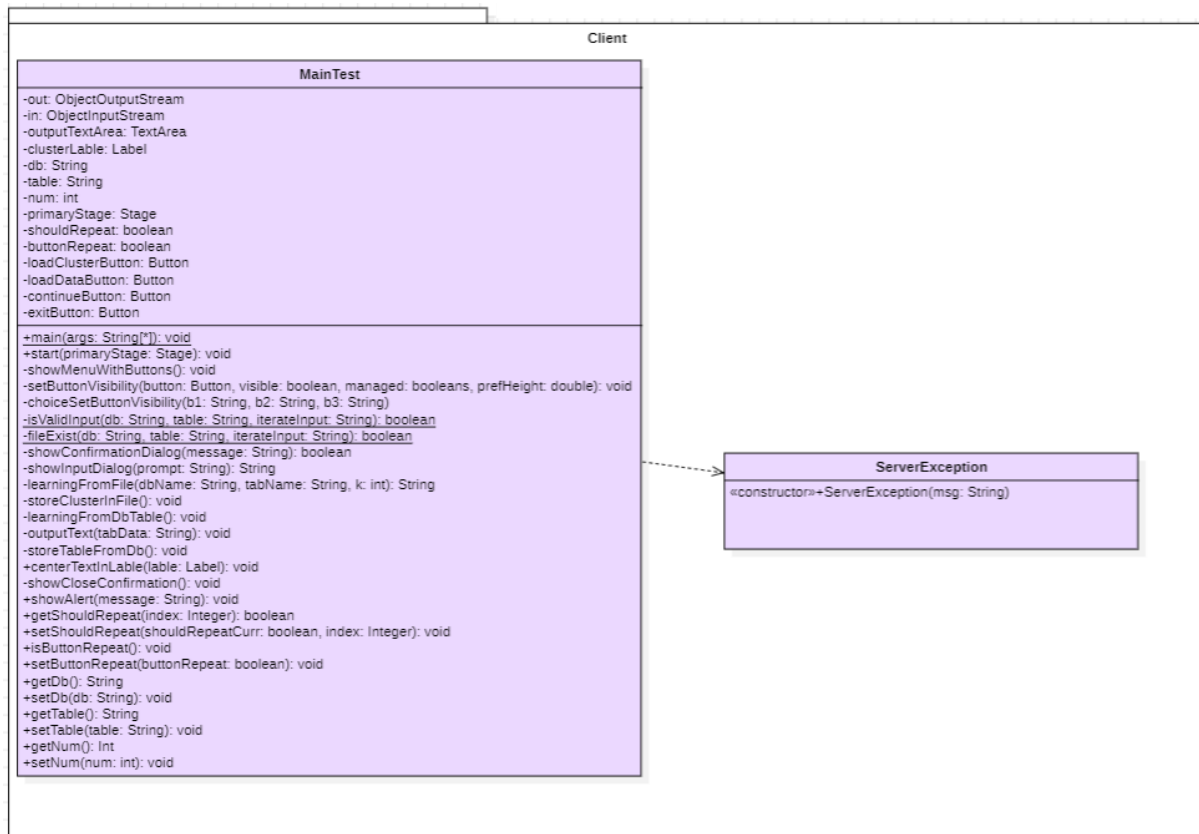
(DIAGRAMMI DEI PACKAGE)



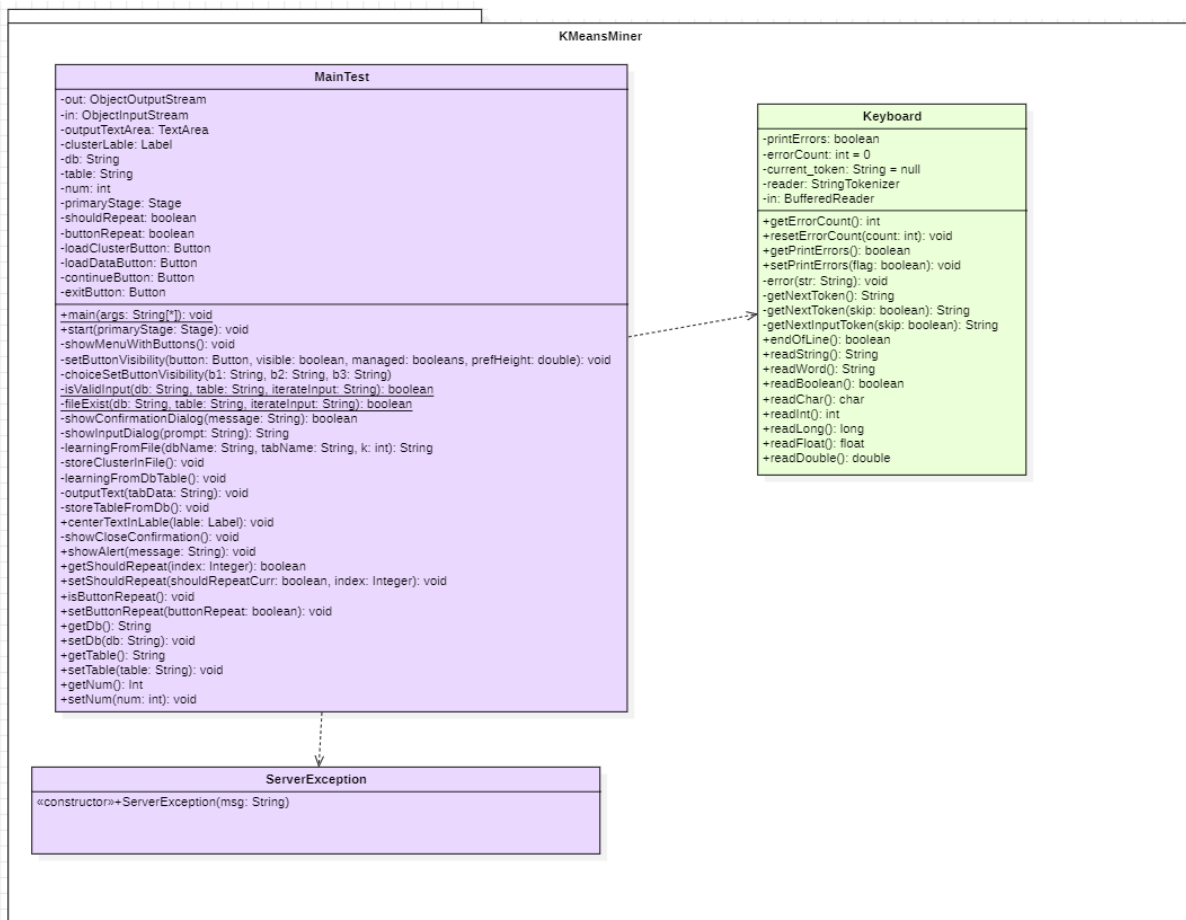
(DIAGRAMMI DELLE CLASSI)

GENERALE:





NEL DETTAGLIO:



TEST

Di seguito verranno riportati esempi di test del client console:

DESCRIZIONE	INPUT	OUTPUT
Avvio del client senza argomenti dopo aver avviato il server	Ø	Messaggio di errore: inserire ip e porta come argomenti

```
Inserire ip e porta come argomenti
```

DESCRIZIONE	INPUT	OUTPUT
Avvio del client senza aver avviato prima il server	Ø	Informazioni di connessione + errore

```
addr = localhost/127.0.0.1
java.net.ConnectException: Connection refused: connect
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento da menu di un valore accettato	1	Prossimo inserimento: Inserisci nome database

```
addr = localhost/127.0.0.1
Socket[addr=localhost/127.0.0.1,port=8080,localport=51803]
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta: 1
Inserisci nome database:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un DB inesistente	database	Prossimo inserimento: Inserisci nome tabella

```
addr = localhost/127.0.0.1
Socket[addr=localhost/127.0.0.1,port=8080,localport=51803]
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta: 1
Inserisci nome database: database
Inserisci nome tabella:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di una tabella inesistente	tabella	Prossimo inserimento: Inserisci numero cluster

```
addr = localhost/127.0.0.1
Socket[addr=localhost/127.0.0.1,port=8080,localport=51803]
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta: 1
Inserisci nome database: database
Inserisci nome tabella: tabella
Inserisci numero di cluster:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster inesistente	5	Errore I/O + scelta se si vuole ripetere l'esecuzione

```
addr = localhost/127.0.0.1
Socket[addr=localhost/127.0.0.1,port=8080,localport=51803]
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta: 1
Inserisci nome database: database
Inserisci nome tabella: tabella
Inserisci numero di cluster: 5
client.ServerException: Errore di I/O
Vuoi ripetere l'esecuzione? (y/n) |
```

DESCRIZIONE	INPUT	OUTPUT
Dopo avere accettato di ripetere l'esecuzione, inserimento di un database esistente	MapDB	Prossimo inserimento: Inserisci nome tabella

```
client.ServerException: Errore di I/O
Vuoi ripetere l'esecuzione? (y/n) y
Inserisci nome database: MapDB
Inserisci nome tabella:
```

DESCRIZIONE	INPUT	OUTPUT
inserimento di una tabella esistente	palytennis	Prossimo inserimento: Inserisci numero cluster

```
client.ServerException: Errore di I/O
Vuoi ripetere l'esecuzione? (y/n) y
Inserisci nome database: MapDB
Inserisci nome tabella: playtennis
Inserisci numero di cluster:
```

DESCRIZIONE	INPUT	OUTPUT
inserimento di un numero di cluster esistente	3	Risultato lettura da file MapDBplaytennis3.dat + scelta se si vuole ripetere l'esecuzione

```
Inserisci nome database: MapDB
Inserisci nome tabella: playtennis
Inserisci numero di cluster: 3
0: Centroid=(rain 14.051666666666668 normal weak yes )
1: Centroid=(overcast 6.275 normal strong yes )
2: Centroid=(sunny 21.525 high strong no )

Vuoi ripetere l'esecuzione? (y/n)
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster non compreso nel range [1 - tuple playtennis}	15	Errore di I/O + scelta se si vuole ripetere l'esecuzione

```
Inserisci nome database: MapDB
Inserisci nome tabella: playtennis
Inserisci numero di cluster: 15
client.ServerException: Errore di I/O
Vuoi ripetere l'esecuzione? (y/n)
```

DESCRIZIONE	INPUT	OUTPUT
Scelta se non si vuole ripetere l'esecuzione	n	scelta se si vuole scegliere una nuova operazione da menu

```
Vuoi ripetere l'esecuzione? (y/n) n
Vuoi scegliere una nuova operazione da menu? (y/n)
```

DESCRIZIONE	INPUT	OUTPUT
Scelta se si vuole scegliere una nuova operazione da menu	y	scelta da menu

```
Vuoi scegliere una nuova operazione da menu? (y/n) y
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento nel menu di un valore accettato	2	scelta se si vogliono usare i valori di default

```
Scegli una opzione
(1) Carica Cluster da File
(2) Carica Dati
Risposta: 2
Vuoi usare i valori di default? (y/n)
```

DESCRIZIONE	INPUT	OUTPUT
Dopo aver scelto di usare i valori di default, scelta numero cluster	y	tabella: playtennis + inserimento: Numero di cluster

```
Vuoi usare i valori di default? (y/n) y
outlook, temperature, umidity, wind, play
1: sunny, 30.3, high, weak, no
2: sunny, 30.3, high, strong, no
3: overcast, 30.0, high, weak, yes
4: rain, 13.0, high, weak, yes
5: rain, 0.0, normal, weak, yes
6: rain, 0.0, normal, strong, no
7: overcast, 0.1, normal, strong, yes
8: sunny, 13.0, high, weak, no
9: sunny, 0.1, normal, weak, yes
10: rain, 12.0, normal, weak, yes
11: sunny, 12.5, normal, strong, yes
12: overcast, 12.5, high, strong, yes
13: overcast, 29.21, normal, weak, yes
14: rain, 12.5, high, strong, no

Numero di cluster:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster valido	3	numero iterazioni + risultato della clusterizzazione + scelta se si vuole ripetere l'esecuzione

```

Numero di cluster: 3
Numero di Iterazioni: 4
0: Centroid=(sunny 23.22 high weak no)
Examples:
[sunny 30.3 high weak no] dist=0.23366336633663376
[sunny 30.3 high strong no] dist=1.2336633663366339
[overcast 30.0 high weak yes] dist=2.223762376237624
[sunny 13.0 high weak no] dist=0.3372937293729372
[rain 12.5 high strong no] dist=2.353795379537954
AvgDistance=1.2764356435643565

1: Centroid=(rain 5.0200000000000005 normal weak yes)
Examples:
[rain 13.0 high weak yes] dist=1.2633663366336634
[rain 0.0 normal weak yes] dist=0.1656765676567657
[rain 0.0 normal strong no] dist=2.1656765676567655
[sunny 0.1 normal weak yes] dist=1.1623762376237623
[rain 12.0 normal weak yes] dist=0.2303630363036303
AvgDistance=0.9974917491749175

2: Centroid=(overcast 13.5775 normal strong yes)
Examples:
[overcast 0.1 normal strong yes] dist=0.4448019801980198
[sunny 12.5 normal strong yes] dist=1.0355610561056106
[overcast 12.5 high strong yes] dist=1.0355610561056106
[overcast 29.21 normal weak yes] dist=1.515924092409241
AvgDistance=1.0079620462046206

Vuoi ripetere l'esecuzione? (y/n)

```

DESCRIZIONE	INPUT	OUTPUT
Inserimento numero di cluster non compreso nel range [1 - tuple playtennis]	15	messaggio di errore: Inserire k compreso tra 1 e 14 + scelta se si vuole ripetere l'esecuzione

```
Numero di cluster: 15
client.ServerException: Inserire k compreso tra 1 e 14
Vuoi ripetere l'esecuzione? (y/n)
```

DESCRIZIONE	INPUT	OUTPUT
Dopo aver scelto di non utilizzare i valori di default, scelta del nome del server	n	Inserimento: Inserisci il nome del server (ad esempio localhost)

```
Vuoi usare i valori di default? (y/n) n
Inserisci il nome del server (ad esempio localhost):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un server inesistente	server	Inserimento: Inserisci il nome del database (ad esempio MapDB)

```
Inserisci il nome del server (ad esempio localhost): server
Inserisci il nome del database (ad esempio MapDB):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un database inesistente	'nome database'	Inserimento: Inserisci il nome della tabella (ad esempio playtennis)


```
Inserisci il nome del server (ad esempio localhost): server
Inserisci il nome del database (ad esempio MapDB): database
Inserisci il nome della tabella (ad esempio playtennis):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di una tabella inesistente	'nome tabella'	Inserimento: Inserisci il nome dell'utente (ad esempio MapUser)

```
Vuoi usare i valori di default? (y/n) n
Inserisci il nome del server (ad esempio localhost): server
Inserisci il nome del database (ad esempio MapDB): database
Inserisci il nome della tabella (ad esempio playtennis): tabella
Inserisci il nome dell'utente (ad esempio MapUser):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un nome utente inesistente	'nome utente'	Inserimento: Inserisci la password dell'utente (ad esempio map)

```
Inserisci il nome del server (ad esempio localhost): server
Inserisci il nome del database (ad esempio MapDB): database
Inserisci il nome della tabella (ad esempio playtennis): tabella
Inserisci il nome dell'utente (ad esempio MapUser): utente
Inserisci la password dell'utente (ad esempio map):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di una password inesistente	map	messaggio di errore del database inesistente + Reinserisci i dati + scelta se si vuole usare i valori di default

```

Inserisci il nome del server (ad esempio localhost): server
Inserisci il nome del database (ad esempio MapDB): database
Inserisci il nome della tabella (ad esempio playtennis): tabella
Inserisci il nome dell'utente (ad esempio MapUser): utente
Inserisci la password dell'utente (ad esempio map): map
Client.ServerException: [!] SQLException: Communications link failure

The last packet sent successfully to the server was 0 milliseconds ago. The driver has not received any packets from the server.
[!] SQLState: 08S01
[!] VendorError: 0

Reinserisci i dati

Vuoi usare i valori di default? (y/n)

```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un server esistente	localhost	Inserimento: Inserisci il nome del database (ad esempio MapDB)

```

Inserisci il nome del server (ad esempio localhost): localhost
Inserisci il nome del database (ad esempio MapDB):

```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un database esistente	MapDB	Inserimento: Inserisci il nome della tabella (ad esempio playtennis)

```

Inserisci il nome del server (ad esempio localhost): localhost
Inserisci il nome del database (ad esempio MapDB): MapDB
Inserisci il nome della tabella (ad esempio playtennis):

```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di una tabella esistente	playtennis	Inserisci il nome dell'utente (ad esempio MapUser)

```

Inserisci il nome del server (ad esempio localhost): localhost
Inserisci il nome del database (ad esempio MapDB): MapDB
Inserisci il nome della tabella (ad esempio playtennis): playtennis
Inserisci il nome dell'utente (ad esempio MapUser):

```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un nome utente esistente	MapUser	Inserimento: Inserisci la password dell'utente (ad esempio map)

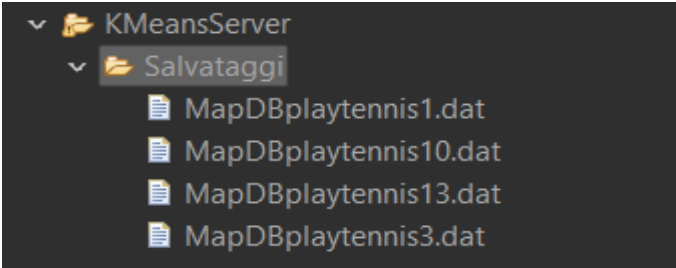
```
Inserisci il nome del server (ad esempio localhost): localhost
Inserisci il nome del database (ad esempio MapDB): MapDB
Inserisci il nome della tabella (ad esempio playtennis): playtennis
Inserisci il nome dell'utente (ad esempio MapUser): MapUser
Inserisci la password dell'utente (ad esempio map):
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento di una password esistente	map	tabella: playtennis + Inserimenro: Numero di cluster

```
outlook, temperature, umidity, wind, play
1: sunny, 30.3, high, weak, no
2: sunny, 30.3, high, strong, no
3: overcast, 30.0, high, weak, yes
4: rain, 13.0, high, weak, yes
5: rain, 0.0, normal, weak, yes
6: rain, 0.0, normal, strong, no
7: overcast, 0.1, normal, strong, yes
8: sunny, 13.0, high, weak, no
9: sunny, 0.1, normal, weak, yes
10: rain, 12.0, normal, weak, yes
11: sunny, 12.5, normal, strong, yes
12: overcast, 12.5, high, strong, yes
13: overcast, 29.21, normal, weak, yes
14: rain, 12.5, high, strong, no

Numero di cluster:
```

DESCRIZIONE	INPUT	OUTPUT
Inserimento numero cluster valido, creazione del file che contiene i centroidi	13	nuovo file: MapDBplaytennis13.dat contentente i centroidi



TEST

Di seguito verranno riportati esempi di test del client console:

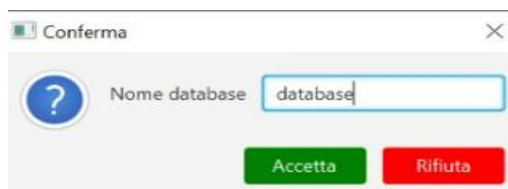
DESCRIZIONE	INPUT	OUTPUT
Avvio del client senza aver avviato prima il server	∅	Errore di connessione



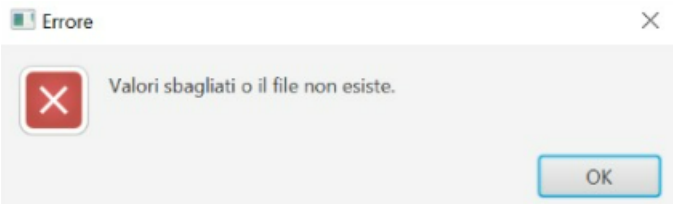
DESCRIZIONE	INPUT	OUTPUT
Scelta dell'opzione 'Caricamento cluster da file'	'Caricamento cluster da file'	Viene richiesto l'inserimento del nome del database



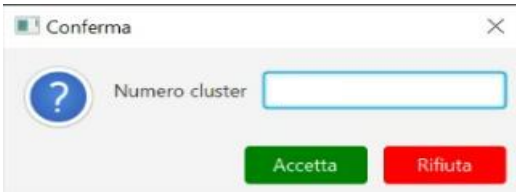
DESCRIZIONE	INPUT	OUTPUT
Inserimento di un DB inesistente	database	Viene richiesto l'inserimento del nome di una tabella



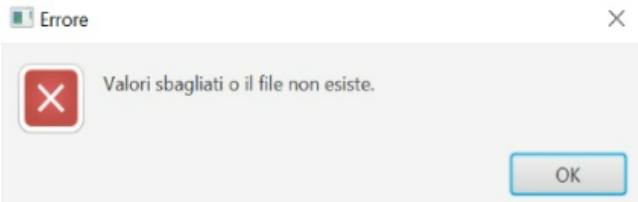
Al termine della compilazione dei campi, supponendo che siano stati inseriti valori corretti nei campi nome tabella e numero cluster, se verrà inserito il nome di un database non esistente come sopra riportato, apparirà il sgeuente messaggio d’errore:



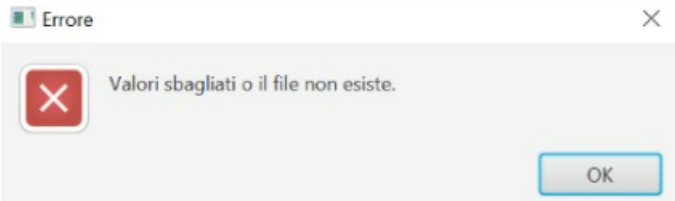
DESCRIZIONE	INPUT	OUTPUT
Inserimento di una tabella inesistente	tabella	Prossimo inserimento: In-serisci numero cluster



DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster inesistente	5	Errore I/O



Al termine della compilazione dei campi, supponendo che siano stati inseriti valori corretti nei campi nome database e numero cluster, se verrà inserito il nome di una tabella non esistente come sopra riportato, apparirà il seguente messaggio d’errore:



DESCRIZIONE	INPUT	OUTPUT
inserimento di un db esistente	MapDB	Inserisci nome tabella

Conferma

Nome database

Accetta Rifiuta

Conferma

Nome tabella

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
inserimento di una tabella esistente	palytennis	Inserimento del numero di cluster

Conferma

Nome tabella

Accetta Rifiuta

Conferma

Numero cluster

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
inserimento di un numero di cluster esistente	3	Risultato lettura da file

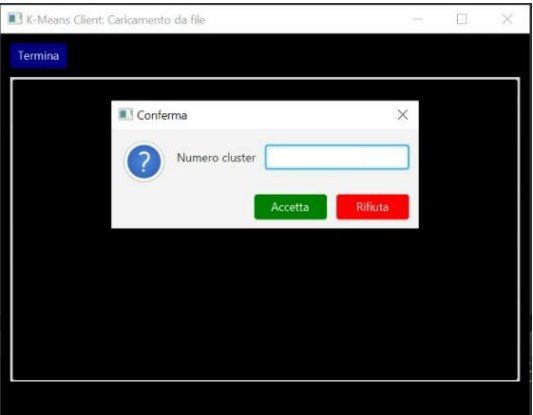
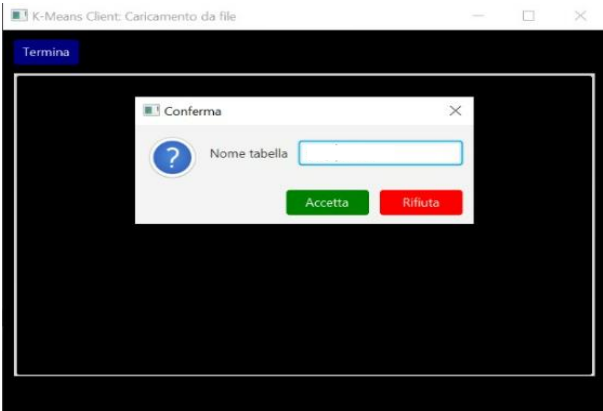
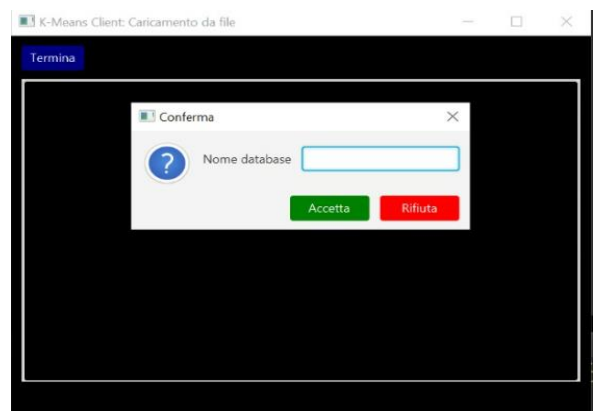
K-Means Client: Caricamento da file

Continua

Termina

0: Centroid=(sunny 16.301428571428573 normal weak yes)
1: Centroid=(overcast 6.3 normal strong yes)
2: Centroid=(rain 13.76 high strong no)

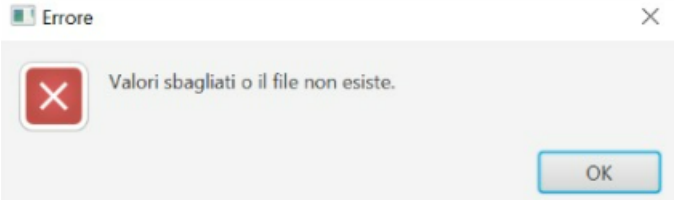
DESCRIZIONE	INPUT	OUTPUT
Scelta se si vuole ripetere l'operazione	si	Vengono richiesti il nome del data-base, nome della tabella e numero di cluster



DESCRIZIONE	INPUT	OUTPUT
Scelta se non si vuole ripetere l'esecuzione	no	scelta se si vuole scegliere una nuova operazione da menu



DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster non compreso nel range [1 - tuple playtennis}	15	Errore di I/O



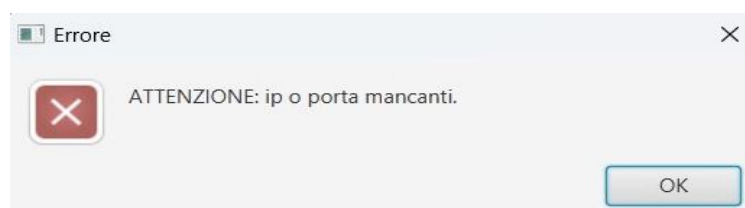
DESCRIZIONE	INPUT	OUTPUT
Terminare il programma	Termina	Conferma sul voler abbandonare il programma



DESCRIZIONE	INPUT	OUTPUT
Scelta dell'opzione 'Rifiuta' durante l'inserimento	Rifiuta	Messaggio di chiusura del programma



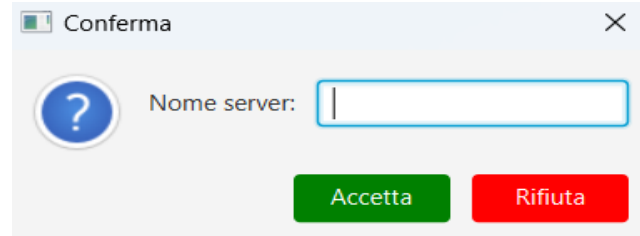
DESCRIZIONE	INPUT	OUTPUT
Se non vengono inseriti ip e/o numero di porta	Ø	Messaggio di errore: ip o porta mancanti



DESCRIZIONE	INPUT	OUTPUT
Scelta dell'opzione 'Caricamento dati da DB '	'Caricamento dati da DB '	Viene richiesto se si vogliono utilizzare valori di default



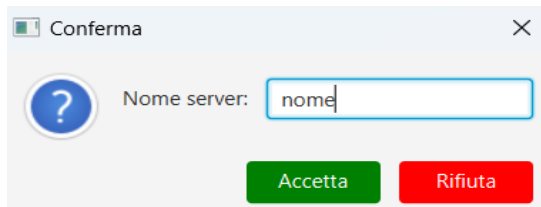
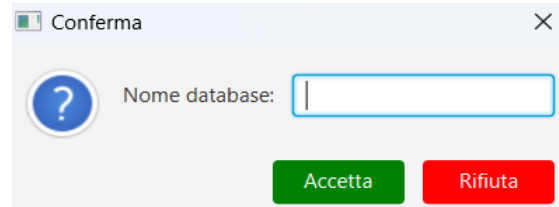
DESCRIZIONE	INPUT	OUTPUT
Scelta dell'opzione 'Rifiuta' nell'utilizzo dei valori di default	Rifiuta	Viene richiesto l'inserimento del nome del server



DESCRIZIONE	INPUT	OUTPUT
Scelta dell'opzione 'Rifiuta' nell'inserimento di un dato	Rifiuta	Messaggio: ''Hai scelto di non continuare''



DESCRIZIONE	INPUT	OUTPUT
Inserimento del nome del server	nome	Inserimento del nome del database

DESCRIZIONE	INPUT	OUTPUT
Inserimento del nome del database	database	Inserimento del nome della tabella

Conferma

Nome database:

Accetta Rifiuta

Conferma

Nome tabella:

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
Inserimento del nome della tabella	tabella	Inserimento del nome dell'utente

Conferma

Nome tabella:

Accetta Rifiuta

Conferma

Nome utente:

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
Inserimento del nome utente	utente	Inserimento del nome password

Conferma

Nome utente:

Accetta Rifiuta

Conferma

Password:

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
Inserimento del nome password	password	A seguito dell'esistenza del database con i dati sbagliati precedentemente inseriti, avremo un messaggio di errore

Conferma

Password:

Accetta Rifiuta

Errore

Il database non esiste o non trovato!

OK

DESCRIZIONE	INPUT	OUTPUT
Inserimento dei dati esistenti di un database	'localhost'-'MapDB'-'playtennis'-'MapUser'-'map'	Inserimento del numero di cluster(numerico)

Conferma

Nome server: localhost

Accetta Rifiuta

Conferma

Nome database: MapDB

Accetta Rifiuta

Conferma

Nome tabella: playtennis

Accetta Rifiuta

Conferma

Nome utente: MapUser

Accetta Rifiuta

Conferma

Password: map

Accetta Rifiuta

Conferma

Numero di cluster(numerico):

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster non numerico	k	Inserimento del numero di cluster(numerico)

Conferma

Numero di cluster(numerico): k

Accetta Rifiuta

Conferma

Numero di cluster(numerico):

Accetta Rifiuta

DESCRIZIONE	INPUT	OUTPUT
Inserimento di un numero di cluster numerico	11	Inserimento Tabella scelta con il numero di iterazioni e i relativi centroidi

Conferma

?

Numero di cluster(numerico):

11

Accetta

Rifiuta

<div><div>Continua</div><div>Termina</div></div> <div><div>outlook, temperature, umidity, wind, play 1: sunny, 30.3, high, strong, no 2: sunny, 30.3, high, strong, no 3: overcast, 30.0, high, weak, yes 4: rain, 13.0, high, weak, yes 5: rain, 0.0, normal, weak, yes 6: rain, 0.0, normal, strong, no 7: overcast, 0.1, normal, strong, yes 8: sunny, 13.0, high, weak, no 9: sunny, 0.1, normal, weak, yes 10: rain, 12.0, normal, weak, yes 11: sunny, 12.5, normal, strong, yes 12: overcast, 12.5, high, strong, yes 13: overcast, 29.21, normal, weak, yes 14: rain, 12.5, high, strong, no Numero di iterazioni: 2 0: Centroid=(sunny 30.3 high strong no) Examples: [sunny 30.3 high strong no] dist=0.0 AvgDistance=0.0 1: Centroid=(sunny 13.0 high weak no) Examples: [sunny 13.0 high weak no] dist=0.0 AvgDistance=0.0 2: Centroid=(overcast 6.3 normal strong yes) Examples: [overcast 0.1 normal strong yes] dist=0.2046204620462046 [sunny 12.5 normal strong yes] dist=1.2046204620462047 AvgDistance=0.7046204620462047 3: Centroid=(sunny 30.3 high weak no) Examples: [sunny 30.3 high weak no] dist=0.0 AvgDistance=0.0</div></div>

DESCRIZIONE	INPUT	OUTPUT
Scelta di utilizzare i valori di default	accetta	Inserimento numero di cluster(numerico)

Conferma

?

Nome server:

Accetta

Rifiuta

Conferma

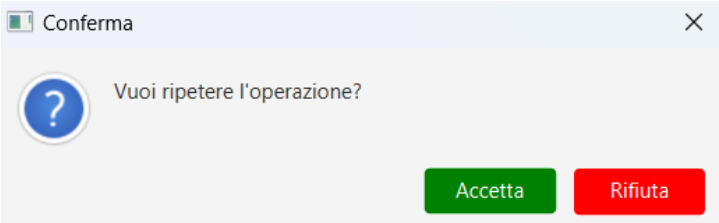
?

Numero di cluster(numerico):

Accetta

Rifiuta

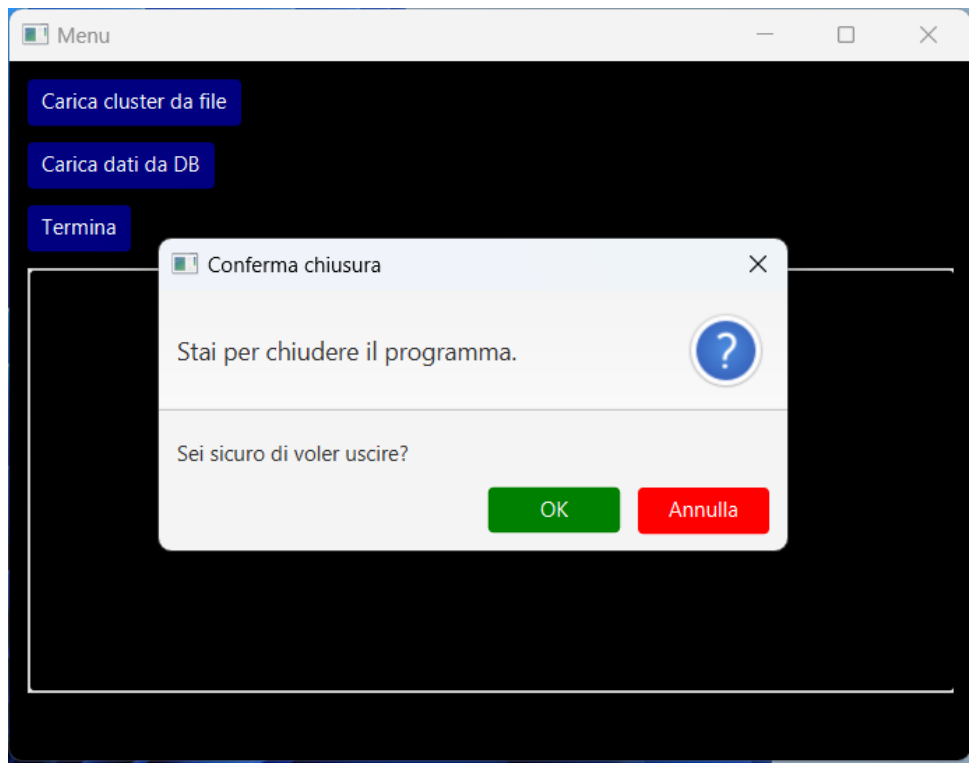
DESCRIZIONE	INPUT	OUTPUT
Scelta del tasto continua dopo aver visualizzato il precedente output	Continua	Messaggio di conferma per la ripetizione dell'operazione



DESCRIZIONE	INPUT	OUTPUT
Scelta di non ripetere l'operazione	Rifiuta	Menu del programma



DESCRIZIONE	INPUT	OUTPUT
Scelta di chiudere il programma attraverso la X della finestra	X	Scelta sulla chiusura del programma



DESCRIZIONE	INPUT	OUTPUT
Scelta di non continuare durante l'inserimento dei dati	Rifiuta	Messaggio sulla scelta di non voler continuare

Conferma

?

Nome server:

AccettaRifiuta

Errore

✖

Hai scelto di non continuare.

OK

ERRORI COMUNI

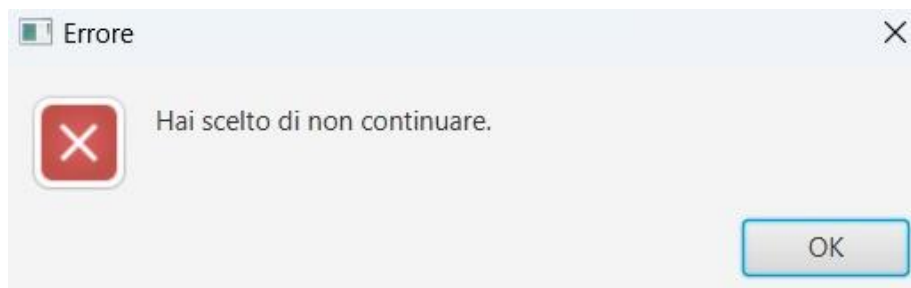
Durante l'esecuzione in entrambe le opzioni, sia il caricamento da file che da Db, è possibile incorrere in errori.

Di seguito verranno riportati quelli più comuni.

- Quando viene inserito un numero di Cluster che supera il range [1-tuple tabella]



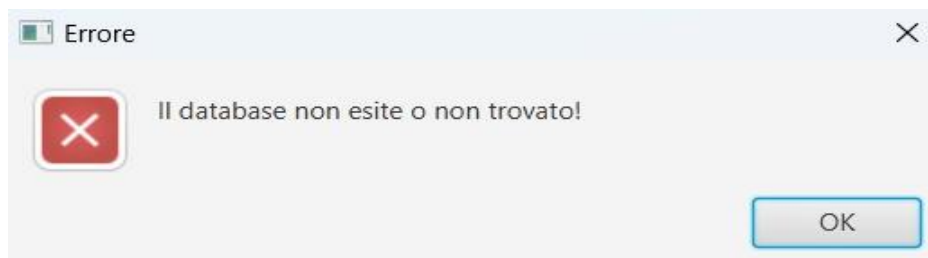
- Quando si decide di terminare il programma prima di finire l'esecuzione



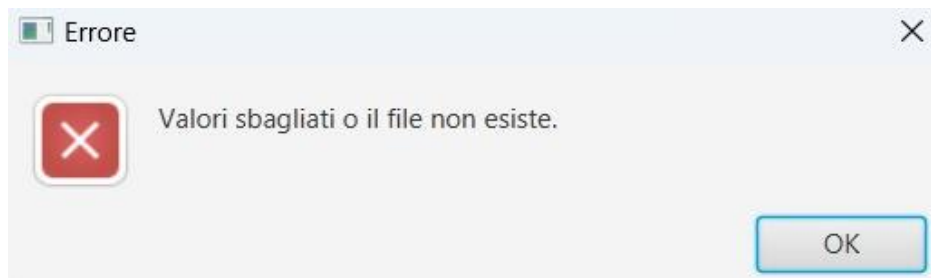
- Quando viene inserito un valore nullo o si viene chiuso l'inserimento tramite la 'X'



- Quando viene inserito il nome di un database non esistente



- Quando vengono inseriti valori di tabelle, database, numero cluster, nome utente o password non esistenti



- Quando viene avviato il Client prima del Server

