

Caso di Studio : Data Mining

Lorenzo Levanto

Guida utente

Versione Base

Una volta avviati sia l'applicativo server sia l'applicativo client, viene richiesto all'utente in quale modo caricare il training set già predisposto:

- ◆ tramite **file di testo**;
- ◆ tramite **file binario**;
- ◆ tramite **database**.

Input: file di testo

Se vogliamo caricare un file di testo è sufficiente **digitare "1"** e **incollare il percorso del file** denominato "*provaC.dat*" (*presente nella cartella server\bin\provaC.dat*).

In caso di path del file errato, l'applicativo ci segnalerà un **errore** e ci chiederà di inserire nuovamente il percorso di un file di testo.

Il **training set** contenuto nel file di testo deve essere formattato nel seguente modo:

- ◆ nel primo rigo del file si indica il **numero di variabili** presenti nel training set (@schema);
- ◆ si procede con la **descrizione di ogni variabile**, specificando se si tratta di una variabile discreta o di una variabile numerica (@desc);
- ◆ si dichiara la **variabile target** (@target);
- ◆ si procede con l'**elenco degli esempi** presenti nel training set dopo aver indicato il **numero** di essi.

Se il training set non è formattato nel modo sopra citato possono generarsi alcuni errori che verranno segnalati a schermo ma se, invece, il training set contenuto nel file è formattato nel modo corretto, esso verrà caricato sul server e, a questo punto, verrà chiesto al client di inserire l'**esempio** di cui si vuole **calcolare il target**. Chiaramente, in caso di input sbagliati da parte dell'utente, saranno generati degli errori.

Una volta inseriti i parametri in modo corretto, il server procederà ad eseguire i calcoli necessari e invierà al client l'output contenente il **valore predetto**; chiederà, infine, all'utente se vuole eseguire un'altra predizione.

L'applicativo funziona allo stesso modo se l'utente decide di inserire in input un **file**

binario o un **database**.

Esempio: $X[0]=A$; $X[1]=5$; $k=8$

Input: file binario

Per inserire in input un file binario è necessario **digitare “2”** e **incollare il percorso** del file *“provaC.dmp”* (contenuto nella cartella *server\bin\provaC.dmp*).

Se viene inserito un path errato genererà un errore e richiederà di inserire il percorso corretto

Dopo aver inserito un file corretto, il training set verrà caricato sul server e verranno eseguiti gli stessi passaggi che si sono succeduti avendo un file di testo come metodo di caricamento dell’input.

Input: database

Se si vuole usare un database come metodo di caricamento è sufficiente **digitare “3”** e **indicare il nome della tabella** (*provaC* nel nostro caso) dalla quale prelevare i valori del training set ma ciò è possibile soltanto se prima è stato eseguito lo script SQL contenuto nella cartella *server\bin\script.sql*

Versione Estensione

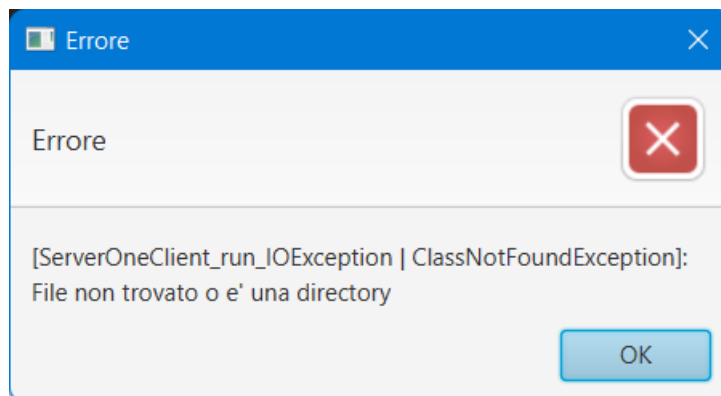
Nella versione estesa è stato deciso di mantenere lo stesso server, implementando invece **un’interfaccia grafica** per il client.

Alla prima apertura dell’applicativo client apparirà una schermata iniziale che richiederà all’utente di inserire i parametri da impostare per la connessione al server e inseriremo localhost e 8080.

Una volta eseguita la connessione con il server, l’applicativo client chiederà all’utente il **metodo di caricamento** del training set (che può essere nuovamente un file di testo, un file binario o un database)

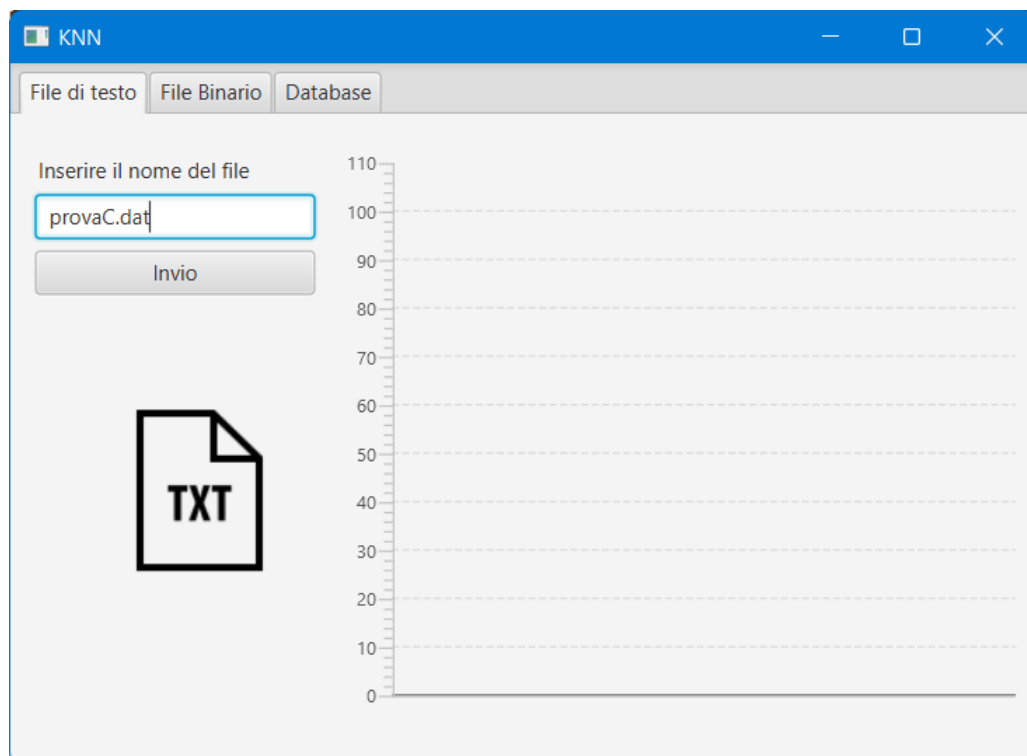
- ◆ se si vuole caricare il training set da un **file di testo**, basterà digitare *“provaC.dat”* nella sezione relativa al file di testo;
- ◆ se lo si vuole caricare mediante **file binario**, basterà digitare *“provaC.dmp”* nella sezione relativa al file binario;
- ◆ se, invece, lo si vuole caricare mediante **database**, basterà digitare *“provaC”* nella sezione relativa al database.

Anche in questo caso, se per il metodo di caricamento del training set scelto dall’utente si inserisce un parametro sbagliato, si avranno degli errori:



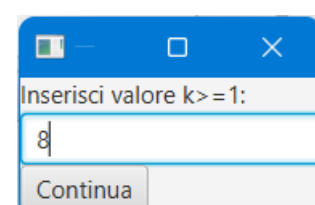
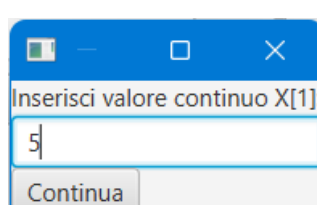
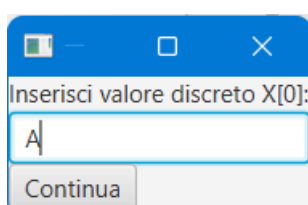
Una volta caricato il training set sul server, l'applicativo client procederà nello stesso modo previsto per la versione base del progetto, ovvero richiedendo in input all'utente **l'esempio** di cui si vuole **calcolare il target**.

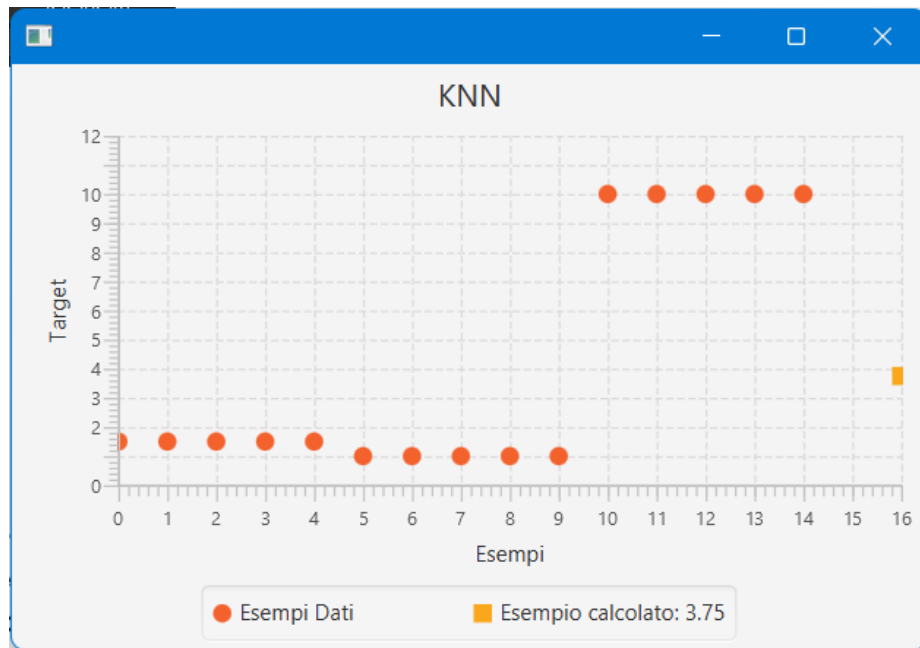
Input: file di testo



Dopo aver inserito l'esempio in input, il server esegue i calcoli e mostrerà l'output in un grafico contenente sia gli **esempi** presenti nel training set (rappresentati con un **cerchio rosso**) sia l'**esempio predetto** (rappresentato con un **quadrato giallo**).

L'esempio predetto viene anche segnalato nella legenda del grafico ("**esempio calcolato**").





Input: file binario

KNN

File di testo File Binario Database

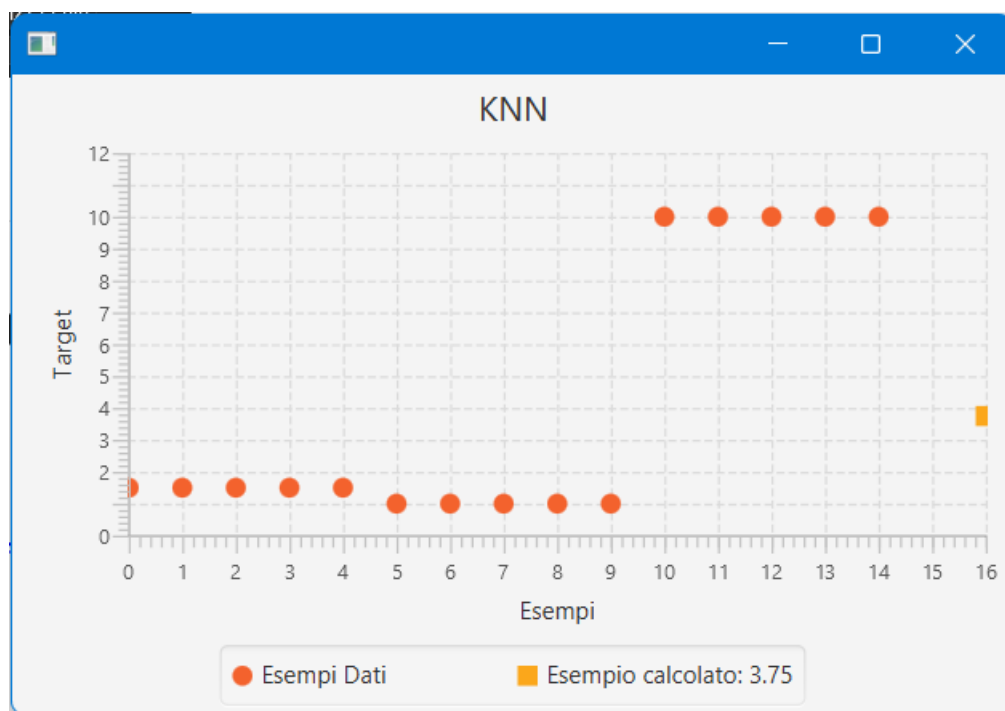
Inserire il nome del file binario

provaC.dmp

Invio

100
001

110
100
90
80
70
60
50
40
30
20
10
0



Input: database

The figure shows the "KNN" application window with the "Database" tab selected. The interface includes a text input field labeled "Inserire il nome della tabella" containing the text "provaC", and an "Invio" button below it. To the left of the vertical axis is a database icon. The vertical axis on the right ranges from 0 to 110.

