

Regression Tree Miner

Domenico Cascella, Patrizia Conte, Alessia Laquale

Novembre 2020

INTRODUZIONE

Un albero di regressione nel data mining consiste nel predire un valore numerico associato ad un esempio basandosi su valori osservati su altri attributi simili. L'albero di regressione è costruito utilizzando tecniche di apprendimento che prendono in input dei dati iniziali chiamati training set. RT-Miner è composto da due applicazioni, una client e una server. Lo scopo del progetto è quello di permettere la costruzione di alberi di regressione a partire da un database MySQL. Una volta acquisito il training set dal database, esso verrà salvato in formato file sul Server, nella quale sarà possibile acquisire dal file senza dover accedere nuovamente al database.

SERVER

Il processo Server si occupa di accogliere le richieste del client. In esso è implementato l'algoritmo RT che, in base alla scelta del client, acquisisce i dati dal database o dal file memorizzato nel Server. Il server funziona nel seguente modo:

1. Inizializza il Server con una porta prefissata(8080)
2. Il server è capace di gestire le richieste di più Client contemporaneamente
3. Il server è in grado di acquisire qualsiasi tabella da una base di dati indipendentemente dal numero di attributi in essa contenuti
4. Esegue l'algoritmo RT-Miner
5. Memorizza l'albero in un file
6. Restituisce al Client l'esito della computazione

CLIENT

Esso è il processo che avvia la comunicazione con il Server scambiandosi dei messaggi.

ESTENSIONE

L'estensione consiste in una interfaccia grafica implementata sull'applicazione lato Client con l'utilizzo della libreria JavaFX. Tale interfaccia permette all'utente di interfacciarsi con il Server in maniera interattiva. Altre differenze sono:

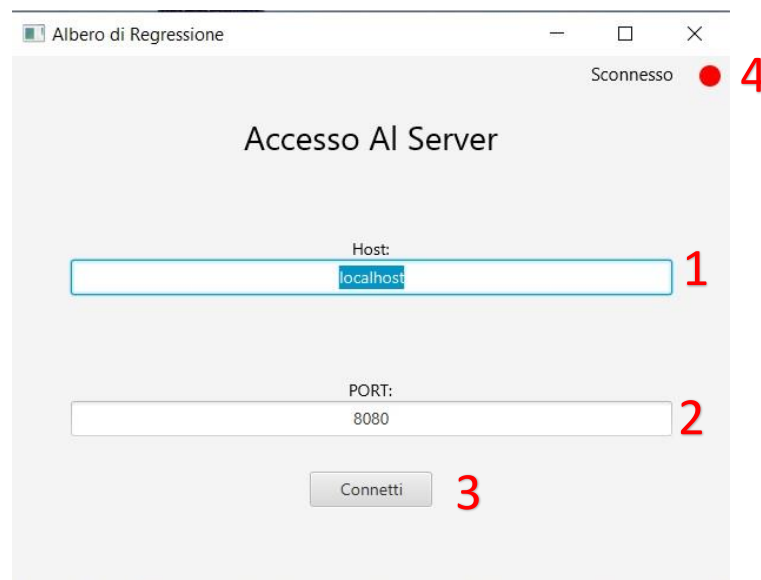
- Possibilità di inserire il nome dell'Host e la porta con cui comunicare;
- Reinserire il nome della tabella in caso di nome errato;
- In caso di predizione errata essa verrà ripetuta;
- Alla fine della predizione se l'utente decide di non continuare la predizione l'utente verrà riportato alla schermata di acquisizione.

MANUALE UTENTE

Client-Schermate:

1. Accesso al Server

L'utente dovrà inserire il nome dell'Host locale(1) e il numero di porta(2) con cui dovrà comunicare con il Server. Dopo aver inserito tali parametri confermiamo la connessione tramite il bottone “**Connetti**”(3). È possibile verificare lo stato della connessione tramite l'indicatore in alto a destra(4). In caso di inserimento host errato, porta errata o mancato avvio del Server verrà mostrato un messaggio di errore vicino le sezioni interessate.



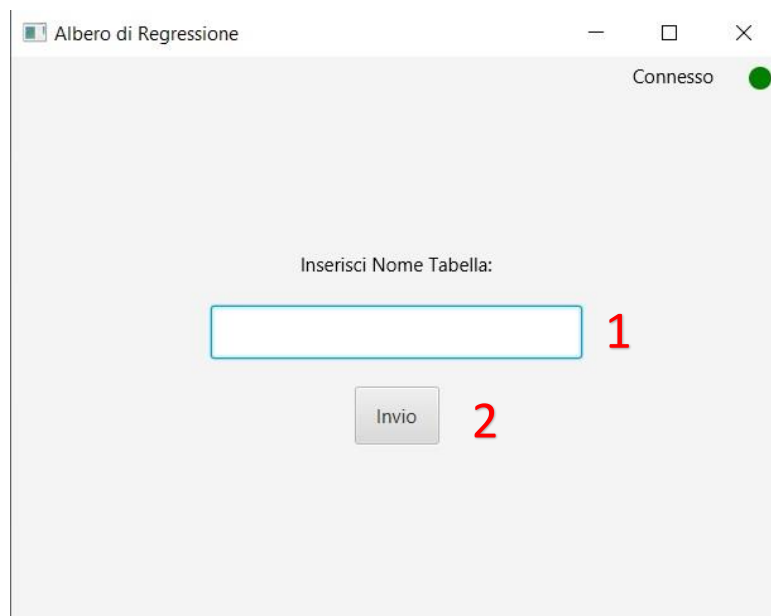
2. Acquisizione

Una volta effettuata la connessione con il Server si dovrà scegliere di acquisire il Training Set da un archivio(1) o da un database(2). Una volta fatta la scelta si effettuerà l'acquisizione tramite il bottone “**Acquisisci**”(3).



3. Inserimento nome file/tabella

In questa schermata l'utente inserirà il nome della tabella o del file(1) da cui acquisire il Training Set. Dopo l'inserimento del nome bisogna confermare la scelta tramite il bottone "Invio"(2). In caso di inserimento di nome errato, verrà mostrato un messaggio di errore nella sezione interessata.



4. Predizione

Nella sezione(1) troviamo le informazioni inerenti all'host, alla porta, all'acquisizione e il nome della tabella o del file scelto.

Nella sezione "Scelte"(2) verranno presentate le possibili opzioni che l'utente dovrà scegliere per scandire l'albero. Le scelte potranno essere fatte tramite la TextField(3) riportata sotto

inserendo il numero corrispondente alle diverse opzioni e confermare tramite il bottone “**Invia**” (4). In caso di inserimento di una lettera verrà mostrato un messaggio di errore, nel caso di inserimento di un valore fuori range la predizione sarà ripetuta nuovamente.

Nella sezione “Valore predetto”(5) verrà visualizzato il valore finale predetto.

The screenshot shows the 'Albero di Regressione' application window. At the top right, it says 'Connesso' with a green dot. The 'Info' section displays: Host: localhost, PORT: 8080, Acquisizione da: Data (labeled 1), and Nome tabella: provac. The 'Scelte' section has two radio buttons: '0:X=A' (labeled 2) and '1:X=B'. A blue arrow points from the 'Scelte' section to the 'Valore Predetto' section, which is empty (labeled 5). Below the 'Scelte' section is a text input field (labeled 3) and an 'Invia' button (labeled 4).

Una volta terminata la predizione verrà chiesto di ripeterla(1) oppure di non ripeterla(2) e in questo caso si ritornerà alla schermata “Acquisizione”.

The screenshot shows the 'Albero di Regressione' application window after a prediction. The 'Info' section remains the same. The 'Scelte' section now has a radio button labeled 'Fine Predizione!' (labeled 1). A blue arrow points from the 'Scelte' section to the 'Valore Predetto' section, which now displays '10.0' (labeled 2). Above the 'Valore Predetto' section, there are two buttons: 'Si' (labeled 1) and 'No' (labeled 2). Below the 'Scelte' section is an empty text input field and an 'Invia' button.

FUNZIONAMENTO

Per eseguire la sorgente del sistema bisogna avviare il file mapServer.bat presente nella cartella (/Distribuzione/Sorgente/ServerBase) e successivamente bisogna avviare il file mapClient.bat presente nella cartella (/Distribuzione/Sorgente/ClientBase).

Per eseguire l'estensione del progetto bisogna avviare il file mapServer.bat presente nella cartella (/Distribuzione/Estensione/ServerJavaFX) e successivamente bisogna avviare il file ClientJFX.bat presente nella cartella (/Distribuzione/Estensione/ClientJavaFX).

GUIDA PER L'INSTALLAZIONE

Per eseguire i file .bat è necessario aver installato la **JDK 8** e **MySQL Command Line 5.7** sul proprio sistema operativo. Inoltre bisogna accertarsi che il servizio di MySQL sia in esecuzione. Bisogna poi inizializzare il database MySQL eseguendo lo script mapdbProvac.sql e mapdbServo.sql presenti nella cartella *distribuzione*. Lo script SQL contiene le istruzioni per creare un database MapDB contenente una tabella di esempio chiamata provaC oppure una tabella di esempio chiamata Servo. Inoltre esso crea un nuovo utente (se non presente) chiamato MapUser con diritti di accesso per il database MapDB.