

iris



Obiettivi

Il file "iris.csv" contiene i dati relativi al dataset Iris. Iris è un dataset multivariato introdotto da Ronald Fisher nel 1936. Consiste in 150 istanze di Iris misurate da Edgar Anderson e classificate secondo tre specie: Iris setosa, Iris virginica e Iris versicolor. Le quattro variabili considerate sono la lunghezza del sepal, la larghezza del sepal, la lunghezza del petalo e la larghezza del petalo.

Per ogni istanza di iris, nel file sono riportati, per riga, il nome della specie e le 4 variabili separate da virgole. Il nome della specie è una stringa, le variabili sono numeri decimali.

Si vuole scrivere un programma che legge il contenuto del file per eseguire delle interrogazioni sui vari campi (filtraggi). Ad esempio si potrebbe volere stampare la lista di tutte le variabili degli iris di tipo "Setosa".

Requisiti del progetto

1. Definire una classe o una struct '**iris**' che deve contenere tutte le informazioni sopra elencate. I dati membro devono avere questi nomi: variety, sepal_length, sepal_width, petal_length, petal_width.
 - a. Implementate i metodi fondamentali della classe
 - b. Implementare, tramite **operator<<**, la stampa del contenuto di un iris sullo stream di output
2. Definire una classe '**iris_dataset**' che deve essere una collezione di oggetti 'iris'. Potete implementarla da zero o appoggiarvi alle classi container della STL per la sua implementazione (es. std::vector, std::list,).
 - a. Implementate i metodi fondamentali della classe 'iris_dataset'.
 - b. Implementate i metodi accessori per poter aggiungere un iris alla collezione e per accedere ai singoli oggetti 'iris'.

- c. La classe deve essere dotata di iteratori (se usate classi container della libreria standard riutilizzate i loro iteratori)
- d. Implementare, tramite `operator<<`, la stampa del contenuto di un oggetto 'comuni' su uno stream di output.

3. Nel `main.cpp`

- a. Definire una funzione globale `load_dataset` che legge il file `iris.csv` per riempire un oggetto 'iris_dataset' passato in input.
- b. Usando gli algoritmi della libreria STL in `<algorithm>`, e gli iteratori della classe 'iris_dataset', provate a filtrare i dati e/o a manipolarli usando diverse strategie mediante funtori definiti da voi.

Esempi di filtraggi/manipolazioni potrebbero essere:

- Calcolare la media delle quattro variabili per ciascun tipo di iris (potete usare uno stesso funtore parametrizzato diversamente)
- Contare il numero di istanze che hanno `sepal_length` compreso in un certo range di valori
- Contare il numero di istanze di iris Versicolor che hanno `petal_length` maggiore di un certo valore
- Determinare l'istanza di un certo tipo di iris che è la più vicina a quella "media" per quella classe
- ...