

Ho scelto questi dataset strutturati:

1- <https://archive.ics.uci.edu/ml/datasets/Iris>

gli attributi sono:

lunghezza petalo

larghezza petalo

lunghezza setolo

gli identificativi sono:

iris virginica

iris setosa

iris versicolor

2- <https://archive.ics.uci.edu/ml/datasets/Car+Evaluation>

gli attributi sono:

Classe Valore Auto: low, med, good, vhigh.

Prezzo di acquisto: vhigh, high, med, low.

Manutenzione: vhigh, high, med, low.

Porte: 2, 3, 4, 5 più.

Persone: 2, 4, più.

lug_boot: piccolo, medio, grande.

sicurezza: bassa, media, alta.

gli identificativi sono:

Modello

Prezzo

Accessori

3- <https://archive.ics.uci.edu/ml/datasets/Wine+Quality>

gli attributi sono:

1 - acidità fissa

2 - acidità volatile

3 - acido citrico

4 - zuccheri residui

5 - cloruri

6 - anidride solforosa libera

7 - anidride solforosa totale

8 - densità

9 - pH

10 - solfati

11 – alcool

gli identificativi sono:

qualità del vino (punteggio compreso tra 0 e 10)

ho scelto questi dataset non-strutturati:

1- <https://www.youtube.com/watch?v=Pj4hTYMhbNI>

questo è un video sul sito youtube, quindi unstructured data, bisognerebbe per analizzare i dati trovare le informazioni rilevanti ed elaborarli in questo caso, i sistemi di gestione di dati utilizzabili sono quelli basati sul modello dell'information retrieveal.

2- <https://www.instagram.com/ferrari/>

questa è la page instagram della ferrari, quindi unstructured data, bisognerebbe per analizzare i dati trovare le informazioni rilevanti ed elaborarli, è necessario prima estrarre tutto in una struttura per rilevare i dati contenuti anche in questo caso, i sistemi di gestione di dati utilizzabili sono quelli basati sul modello dell'information retrieval.