Creo infine una tabella con:

* Tutti i generi
* Totale commenti positivi
* Totale commenti negativi
* Totale commenti positivi maschili
* Totale commenti negativi femminili
* Preferito in maggioranza da: (UOMINI/DONNE)

Ho un problema di classificazione:

* 1 se la recensione è positiva (>= 6 stelle)
* 0 se la recensione è negativa (<5 stelle)

Prendo tutti i commenti relativi al film

Li filtro tenendo solo quelli dove è possibile capire se chi ha scritto il commento è un uomo (M) o una donna (F)

Ottengo una lista di set {[commento, M/F, rating]}

Il rating ci servirà per capire l’efficienza del nostro modello.

500 film per ogni genere

Generi film

www.imdb.com