

# RECOMMENDER CONTENT BASED

## BASATO SU CATEGORIZZATORE NAIVE BAYES

Progettato e sviluppato da: Mele Costantino

### SVOLGIMENTO

Scopo del raccomandatore è l'estrazione delle informazioni contenute nel dataset (si è scelto ua.base), delle descrizioni degli item contenute all'interno del file content (che formeranno l'insieme di features da fornire in input al classificatore) e successiva predizione degli item che potrebbero essere di gradimento all'utente, in base ai training sets forniti al classificatore.

Per lo sviluppo del raccomandatore è stato scelto il seguente classificatore:

<https://github.com/ptnplanet/Java-Naive-Bayes-Classififer>

Sono state previste le categorie Like e Not Like per categorizzare gli item, ognuna di tali classi avrà il proprio training set costituito dall'insieme di feature estratte dagli item che sono stati votati dall'utente (la predizione è basata sull'estrazione dei seguenti attributi: **director** , **writer** , **star** , **genre** ) ; in base al rating dell'utente (Like 4-5 e Not Like 1-2) , l'insieme delle features che descrive ogni item contribuirà al completamento del training set delle due categorie , questo consentirà di generare il profilo dell'utente.

Il test set è formato dall'insieme di feature che descrivono gli item che non sono stati votati dall'utente.

Verranno effettuate k predizioni con k = numero degli item che non sono stati votati dall'utente.

Ogni predizione è formata da una tripla

<ID item , Categoria , Probabilità>

Le predizioni saranno ordinate in ordine decrescente in base a  $P(\text{Like} \mid \text{Item})$  e verranno fornite in output sono le prime, ovvero quelle che dovrebbero essere di maggiore gradimento per l'utente, insieme ai training sets delle due categorie; il classificatore userà l'argmax per generare le varie probabilità.

## DESCRIZIONE DELLE CLASSI

**Package Bayes:** Contiene l'insieme di classi che implementano il classificatore + la classe Item\_Category\_Probability, utile per la memorizzazione delle triple per le predizioni.

### **Package Dataset:**

Functions: classe che contiene una suite di metodi per il funzionamento del recommender, tra questi troviamo metodi per l'estrazione delle informazioni dal dataset e dal file content, metodi per la generazione dei training sets e del test set, infine un metodo che effettua il processo di raccomandazione degli item per l'utente.

Movie: classe che consente la memorizzazione delle informazioni relative agli item.

User\_Item\_Rating: classe che consente la memorizzazione delle informazioni estratte dal dataset.

Test: classe che comprende un esempio del funzionamento del recommender per un determinato utente.