INDICAZIONI DEL PROF

Il precedente progetto sembrava una semplice raccolta dati, senza scopo e senza risultati. Obbiettivo sembra quindi quello di mostrare una qualche relazione o confronto tra fonti.

NUOVA PROPOSTA DI PROGETTO

Confronto tra opinione tra la community Marvel di twitter e quella di Reddit, valutando la frequenza e numerosità delle risposte, l’opinione circa ogni film (sentiment analysis).

Metodo acquisizione dati Twitter:

* Libreria tweepy account developer v2 (chiavi API)
* Ricerca post e commenti a questi sull’account migliore (Marvel entertaintment o la migliore community)
* Ricerca su migliore e peggiore film della marvel (dal 2010)

Metodo acquisizione dati Reddit:

* Chiavi api gratuite fino a post del 2010 (poi a pagamento)
* Libreria PRAW
* Ricerca post e commenti a questi sull’account migliore (Marvel entertaintment o la migliore community)
* Ricerca su migliore e peggiore film della marvel (dal 2010)

Scraping:

* Aggiungere informazioni per ogni post/film con info riguardanti incassi, data uscita, cast, spese, utili per contestualizzare la scelta del film e fare ulteriori analisi per cluster.

Problematiche da risolvere:

* Come ricercare post allo stesso modo tra twitter e reddit? Potrei usare degli hashtags o parole chiavi.
* È meglio ricercare in tutto il mondo social o solo su alcuni canali? Es: community.
* Se ricerco post via API che info posso aggiungere con lo scraping? Info sul film basta?

Ricerca per articoli google news per avere un contesto di riferimento e una chiave (evento) su cui confrontare i post twitter e reddit