

Studio dell'engagement su social networks

Tesi di: Vittorio Maggio, 817034

Relatore: prof.ssa Elisabetta Fersini

L'evoluzione dei social networks



«Thefacebook is an online directory that connects people through social networks at colleges»

2004



64% Degli italiani si informano sui social



81% Dei consumatori si informano su un brand o prodotto, online

2019

Progetto di stage

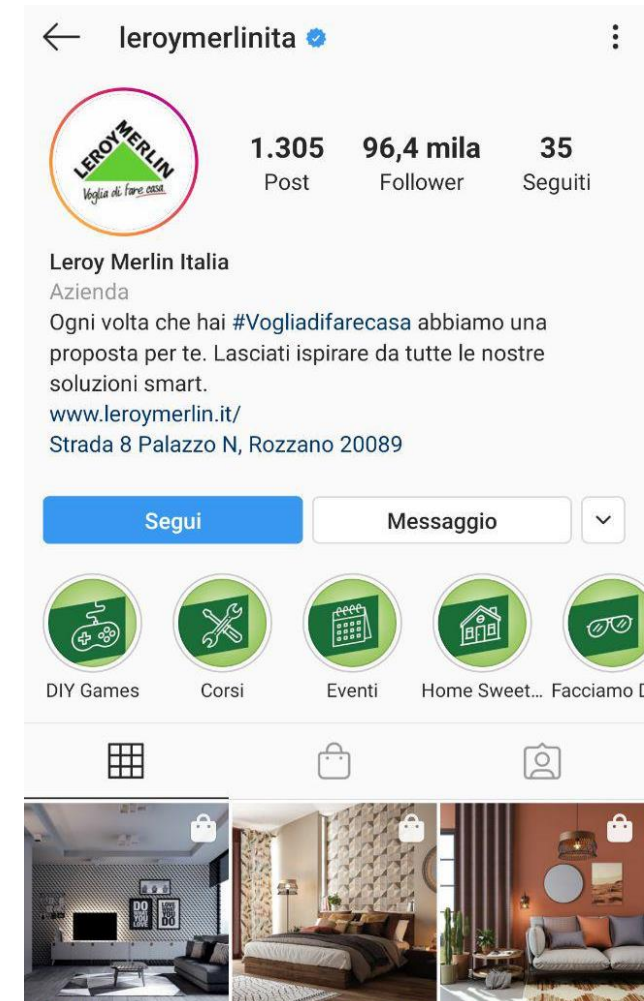
Stimare l'engagement di un nuovo post delle pagine social di Leroy Merlin.

Social : Instagram

Parametro per la stima dell' engagement : $\frac{\text{Num. di likes}}{\text{Num. di followers}}$

Gli steps seguiti sono stati:

- Instagram Data Scraping;
- Sentiment Analysis;
- Analisi dell' immagine;
- Addestramento del modello predittivo.



Instagram Data Scrapping

Sono stati considerati tutti i post pubblicati nell'anno 2018 e 2019 fino al mese di settembre, contenenti foto, delle pagine:

- Leroy Merlin Italia (290 post);
- Leroy Merlin Francia(625 post);
- Leroy Merlin Brasile(563 post).



← Descrizione

← Hastags

← Foto

← Likes

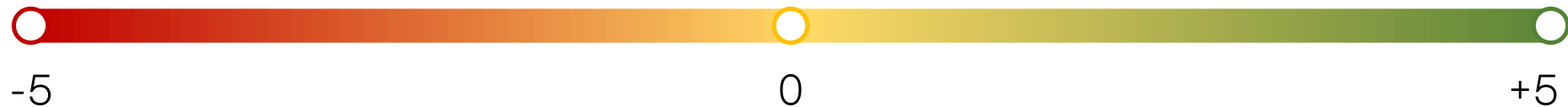
← Data

Sentiment Analysis

Per ogni **descrizione** dei post è stata effettuata un'analisi del sentiment per poterne stabilire la polarità (negativa, neutra o positiva).

Steps seguiti:

- Traduzione del testo della descrizione dalla rispettiva lingua all'inglese
- Normalizzazione del testo
- Applicazione del lessico AFINN

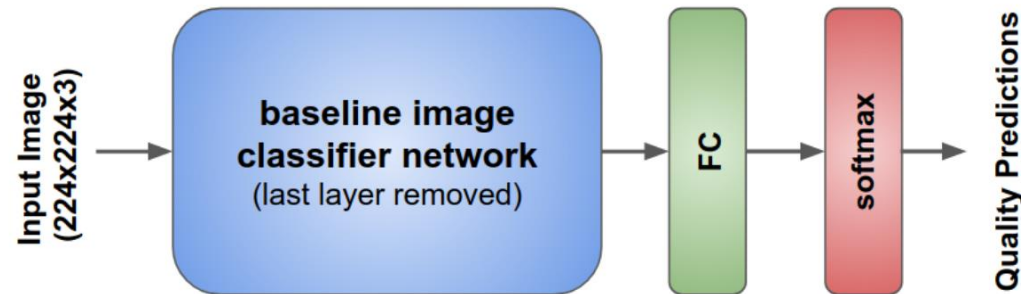


Lessico: AFINN

Analisi dell'immagine

Per quantificare l'estetica dell'immagine condivisa è stata usato un classificatore proposto da Google: «**NIMA: Neural Image Assessment**». In particolare, è stato usato un modello pre-addestrato a cura di *ideal.de*.

Il modello usato consiste in una rete neurale convoluzionale addestrata con il dataset TID2013, che prende in input un'immagine e restituisce un punteggio (da 0 a 9) riguardante la tecnica e l'estetica dell'immagine.



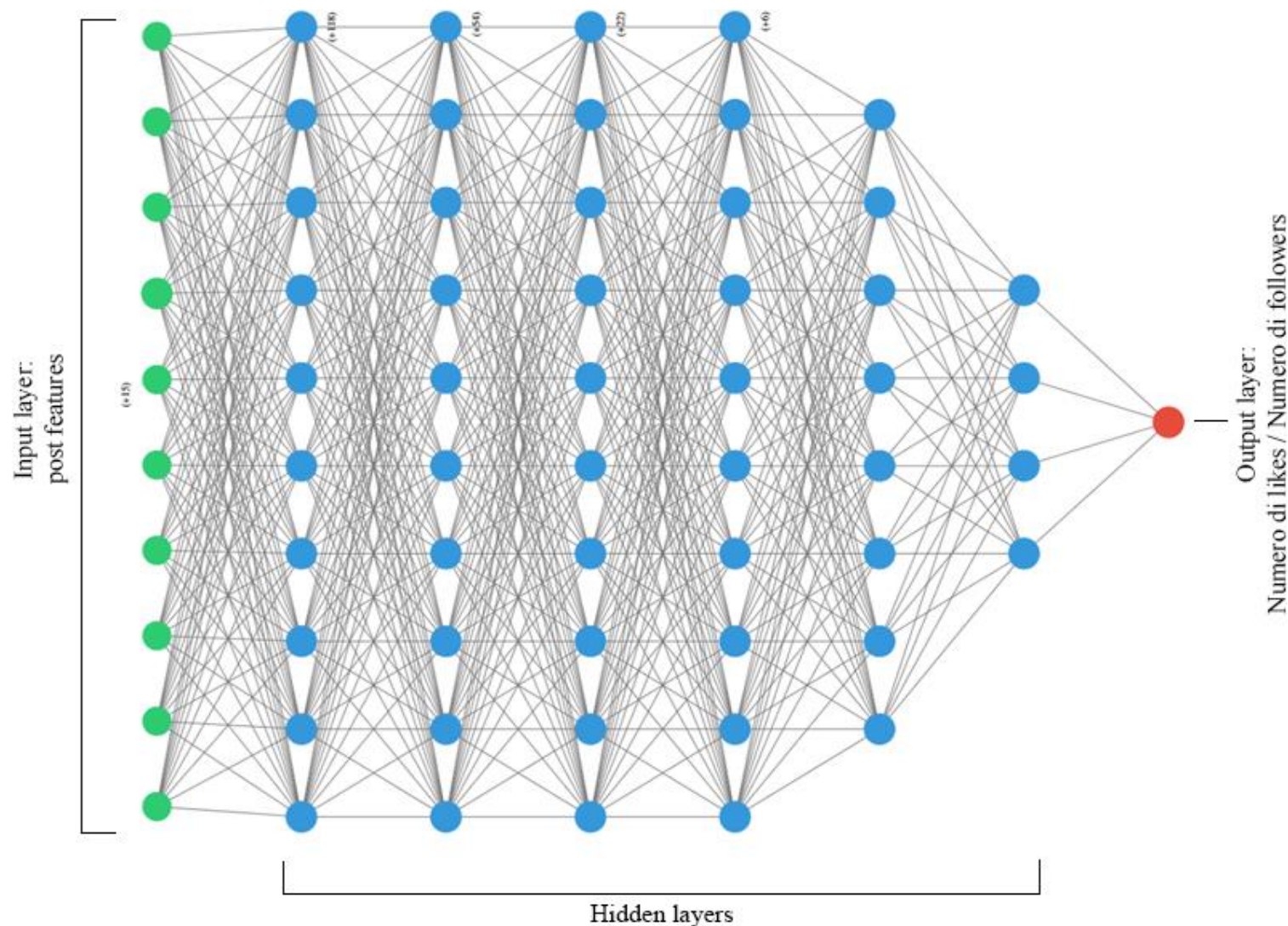
Dataset

Il dataset finale comprende 1,478 post, dei quali il 75% è usato come training set e il 25% come test set.

Per ogni post le features considerate sono:

- Numero caratteri della descrizione del post;
- Numero di hastags usati;
- Copertura totale degli hastags;
- Fascia oraria;
- Giorno della settimana;
- Tempo trascorso dalla pubblicazione del post precedente;
- Sentiment della descrizione del post;
- Punteggio sull'estetica dell'immagine.

Percettrone Multistrato



Funzione di attivazione strati nascosti:
Relu

Funzione di attivazione strato di output:
Lineare

Funzione di ottimizzazione:
Adam

Risultati

La correttezza del modello è stata misurata con il coefficiente di determinazione (R^2).

Test Set:

	LM IT	LM IT + LM FR	LM IT + LM FR + LM BR
$R^2 \approx$	- 0,045	0,018	0,129

Full set:

	LM IT	LM IT + LM FR	LM IT + LM FR + LM BR
$R^2 \approx$	0,071	0,263	0,318

Risultati

Margine di errore del 5%:

Vengono considerate corrette le predizioni che discostano al più del 5% dal valore reale.

Correttezza full set: 76%

Risultati come limite inferiore:

Viene sottratto al valore predetto il 5% e considerato il risultato come limite inferiore di like.

Correttezza full set: 85 %

Sviluppi futuri

- Applicazione del modello su altri social networks e su altri brand;
- Considerare i commenti di un post come ulteriore parametro di engagement;
- Analizzare e categorizzare le tipologie di utenti che interagiscono con la pagina.