

TOOL PER LA CLASSIFICAZIONE DEI SENTIMENTI DEGLI STUDENTI IMPLICATI IN MODULI DIDATTICI UNIVERSITARI IN MODALITÀ E-LEARNING

Giacomo Nalli¹, Daniela Amendola², Chiara Schettini³, Rossana Galassi⁴

¹Scuola di Scienze e Tecnologie, Sez. di Informatica, Università di Camerino
E-mail: giacomo.nalli@unicam.it

²Scuola di Bioscienze e Medicina Veterinaria, Università di Camerino
E-mail: daniela.amendola@unicam.it

³Advanced School of Studies, Università di Camerino
E-mail: chiara.schettini@unicam.it

⁴Scuola di Scienze e Tecnologie, Sez. di Chimica, Università di Camerino
E-mail: rossana.galassi@unicam.it

— **COMUNICAZIONE** —

ARGOMENTO: *Istruzione universitaria*

Abstract

Negli ultimi anni, le piattaforme e-learning risultano essere molto utilizzate, all'interno dei corsi universitari, come supporto alla didattica d'aula per garantire una formazione innovativa che veda lo studente attivo durante il processo di apprendimento al fine di migliorare le proprie conoscenze e competenze. Per raggiungere tali obiettivi è importante che oltre ad una buona progettazione didattica si tenga conto anche dell'aspetto emotivo degli studenti, poiché i sentimenti possono influenzare le loro motivazioni e quindi anche le performance finali. Il nostro lavoro consiste nella realizzazione di un software intelligente in grado di estrarre i sentimenti degli studenti dall'analisi del testo di un questionario a risposta aperta. L'esecuzione si conclude con l'invio automatico al docente di un feedback con le informazioni sui sentimenti estratti, che possono essere utili per verificare quanto il materiale fornito via e-learning sia funzionale per gli studenti. Conoscendo così le esigenze e le difficoltà degli studenti, il docente può modificare, ove e se necessario, la struttura del corso al fine di rendere positivi i sentimenti degli studenti e quindi aumentare la loro motivazione.

Keywords – Sentiment Analysis, Machine Learning, Moodle.

1 INTRODUZIONE

Negli ultimi anni, le piattaforme e-learning risultano essere molto utilizzate all'interno dei corsi universitari, in quanto forniscono tool utili anche come supporto alla didattica frontale.

Gli strumenti a disposizione all'interno di piattaforme e-learning, come ad esempio Moodle, permettono di creare dei percorsi online, dal quale possono trarne giovamento gli studenti ad esempio in termini di miglioramento del voto finale. [1]

Tuttavia, nonostante il docente abbia svolto un buon lavoro nella progettazione e costruzione del corso, in alcuni casi può accadere che gli studenti, pur seguendo il percorso in e-learning, non riscontrino un miglioramento nell'apprendimento o nella performance, non rispettando gli obiettivi formativi attesi dal docente. Questo può capitare ad esempio se lo studente prova sentimenti negativi durante la fruizione di un corso che lo portano a non essere motivato. Infatti, fattori affettivi ed emotivi, sembrano influenzare la motivazione degli studenti e, in generale il risultato del processo di apprendimento. [2]

Inoltre, le informazioni sulle emozioni, provate dagli studenti dopo la fruizione di un corso, possono restituire un importante feedback al docente, il quale può verificare se il corso risulta essere funzionale o se invece di difficile comprensione e fruizione. [3]

In letteratura per estrarre le emozioni degli studenti vengono utilizzate tecniche di Machine Learning che si basano sull'elaborazione e l'analisi dei testi, ad esempio post o tweet dei social network. [4]

Una delle tecniche di data mining più usate per l'estrazione dei sentimenti è la Sentiment Analysis (SA). SA è un'area di ricerca di Natural Language Processing (NLP) e Deep Learning in cui il linguaggio scritto è classificato usando una linguistica computazionale, secondo la polarità di opinione e identificazione delle emozioni. [5]

Il nostro progetto consiste nel creare un software intelligente in grado di estrarre i sentimenti degli studenti dall'analisi dei testi delle risposte aperte di un questionario erogato dopo la fruizione di un percorso on line.

Il software permetterà al docente di avere una panoramica dei sentimenti degli studenti, e se necessario intervenire modificando il percorso didattico. La possibilità di avere a disposizione un modello di classificazione può aiutare il docente ad effettuare delle predizioni in seguito a modifiche della struttura del corso. Può infatti verificare se a livello emotivo ci sia un miglioramento o un meno, e in caso negativo reiterare tale procedimento.

Per lo sviluppo dell'applicativo è stato utilizzato un percorso tutoriale on-line di Chimica di supporto alla didattica frontale per il primo anno del corso di laurea in lingua inglese in Biosciences and Biotechnology dell'Università degli Studi di Camerino erogato tramite la piattaforma Moodle. Hanno partecipato al corso 122 studenti internazionali.

Il corso prevedeva 7 moduli ognuno dei quali ricopre uno specifico argomento della chimica. Ogni modulo presenta:

1. "video experiment", video esperimento relativo all'esercizio della durata massima di 5 minuti,
2. "submicroscopic view", (immagine o video massimo di 2 minuti),
3. "overview", panoramica delle fasi di risoluzione dell'esercizio,
4. "video tutorial", video dello svolgimento dell'esercizio, della durata massima di 20 minuti,
5. "multiple choice", quiz a risposta multipla,
6. "other materials", eventuali materiali di supporto (video, immagini, testi e file).

Al termine del corso è stato chiesto agli studenti di rispondere ad un questionario di gradimento a risposta chiusa/aperta. Le risposte aperte "Cosa è piaciuto di più o cosa è piaciuto di meno del corso e della piattaforma" sono state quelle utilizzate, nel presente progetto, per l'analisi dei sentimenti degli studenti.

2 CLASSIFICAZIONE DEI SENTIMENTI

In questa sezione descriviamo la metodologia utilizzata per la raccolta ed elaborazione dei dati al fine di estrarre i sentimenti degli studenti e creare un modello di classificazione. In Fig. 1 è rappresentato il flusso delle attività on-line, della ricerca e dello sviluppo del software.

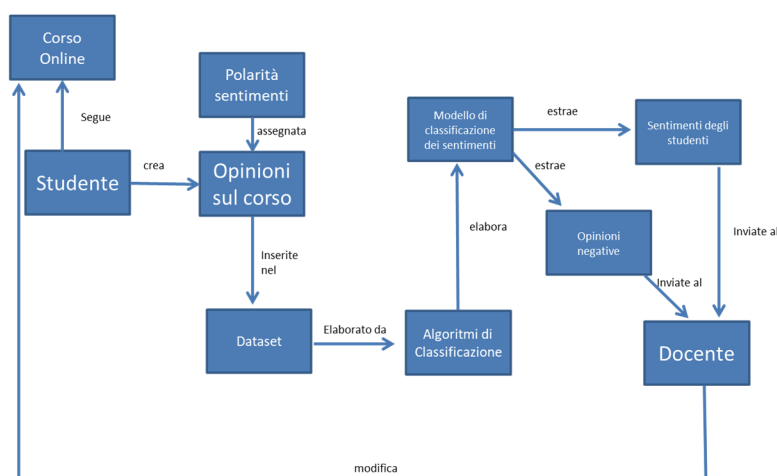


Figura 1 – Diagramma a blocchi del flusso delle attività online, ricerca e software

Per lo sviluppo del software sono state necessarie tre fasi:

1. selezione dei dati da analizzare,
2. scelta delle tecniche di Machine Learning,
3. visualizzazione dei dati.

Nella prima fase sono stati estratti tutti i risultati del questionario di gradimento effettuato dagli studenti, scaricando il relativo file Excel direttamente dall'attività Feedback di Moodle.

Tra tutte le risposte del questionario, sono state selezionate solo le risposte aperte che rappresentavano le opinioni degli studenti riguardo al corso, in particolare in relazione agli aspetti che sono piaciuti maggiormente e in misura minore. Tali dati sono stati organizzati, in un unico file Excel, in vettori di features (caratteristiche) [6] associando agli studenti le loro risposte. Il file Excel, chiamato dataset, rappresenta il file di input del nostro software di classificazione dei sentimenti.

Per poter sviluppare un modello predittivo in un algoritmo supervisionato come la Classificazione, non è sufficiente però avere solo il testo da analizzare. Infatti, è necessario che il nostro dataset sia completo di etichette (label) con l'indicazione della polarità del sentimento associata a quel testo, fornendo dei campioni che l'algoritmo di machine learning può studiare, estraendo una funzione che gli permetterà di effettuare previsioni future su nuovi testi non classificati.

Per completare il dataset è stato quindi necessario associare le polarità dei sentimenti relative ai testi, indicando se questi sono negativi, positivi o neutri.

Una volta ottenuto il dataset completo è stato possibile implementare il software, scegliendo di applicare la tecnica di Classificazione poiché permette di effettuare previsioni sui sentimenti in base al testo inserito. Tale tecnica prevede un input in valori numerici; per cui è stato necessario applicare un metodo che permettesse di convertire il testo in numero per estrarre delle funzionalità. Per fare questo è stato utilizzato l'algoritmo Term Frequency–Inverse Document Frequency (TF-IDF), un modello statistico che ha lo scopo di determinare quanto sia rilevante una parola in un determinato documento.[7]

In seguito, si è passati alla fase relativa alla creazione del modello predittivo di Classificazione. Per far ciò abbiamo diviso il dataset in due parti: training set, e test set.

Il training set è stato usato per l'addestramento dell'algoritmo mentre il test set per la valutazione dell'accuratezza del modello di machine learning. Successivamente è stato implementato l'algoritmo di Classificazione, che sulla base del training e test set, ha permesso di effettuare predizioni sui sentimenti provati dagli studenti.

Il software, conclusa l'esecuzione, prevede la stampa dell'indicatore di precisione del modello per valutare quanto l'addestramento dell'algoritmo sia valido e quanto il modello sia affidabile per previsioni riguardo dati futuri non classificati.

L'esecuzione del software si conclude con l'invio automatico al docente di una e-mail con indicato il numero delle opinioni che hanno prodotto sentimenti negativi, positivi e neutrali, insieme all'invio di una tabella excel che racchiude i commenti degli studenti che hanno riscontrato sentimenti negativi.

3 RISULTATI E DISCUSSIONI

Per testare il software è stato necessario creare il dataset, il file da fornire in input al software, inserendo le opinioni che gli studenti hanno inserito nel questionario. Le risposte disponibili sono state 132, su un totale di 66 studenti che hanno svolto il questionario.

In questo lavoro si è deciso di elaborare solamente due tipologie di risposta aperta del questionario di gradimento, relative agli aspetti del corso e della piattaforma che sono stati graditi o meno dagli studenti. Tuttavia, il docente è libero di selezionare i testi da elaborare da tutte le attività presenti in piattaforma come ad esempio forum, chat, consegne, etc.

Sulla base del dataset fornito, dopo aver associato le polarità dei sentimenti ai testi, il software ha convertito le opinioni degli studenti dal formato testo a vettori numerici tramite l'algoritmo TF-IDF.

Successivamente è stato diviso il dataset in training set e test set per l'applicazione della tecnica di classificazione. Una volta eseguito l'algoritmo, il software ha elaborato i dati e restituito i risultati.

Su 132 risposte date dagli studenti, 86 sono risultate positive, 28 neutre e 18 negative.

I risultati ottenuti, insieme all'elenco delle opinioni che hanno prodotto sentimenti negativi, verranno successivamente inviati dal software al docente via mail. Questo passaggio serve per garantire al docente di avere a disposizione una visione di insieme dei risultati ottenuti ed una panoramica di tutti gli aspetti negativi messi in evidenza dagli studenti al fine di poter migliorare gli aspetti del corso che hanno ricevuto un giudizio negativo.

Il modello predittivo ha ottenuto un'accuratezza pari all'80%. Questo permette al docente di poter utilizzare il modello per effettuare nuove predizioni su risposte non classificate, estraendo i sentimenti degli studenti con un buon livello di precisione.

4 CONCLUSIONI E SVILUPPI FUTURI

In questo lavoro abbiamo realizzato un software finalizzato alla classificazione dei sentimenti provati dagli studenti dopo la fruizione di un corso in modalità blended learning. La possibilità di utilizzare tecniche di Classificazione permette di poter realizzare un modello predittivo, con una buona precisione, per classificare nuovi testi non processati, senza la necessità di eseguire nuovamente l'addestramento del dataset con tecniche di Machine Learning. Per il docente la possibilità di individuare i commenti etichettati come negativi, può essere utile ad analizzare eventuali problematiche sul corso e cercare di migliorare alcuni aspetti che possano aiutare gli studenti nella fruizione del corso e nell'aspetto motivazionale. Il modello predittivo può essere utilizzato in seguito ad una modifica del corso da parte del docente. In questo modo il docente può verificare se il cambiamento apportato porta anche a un miglioramento del sentimento negli studenti.

Il progetto di ricerca proseguirà con la realizzazione di un Plugin avanzato per Moodle che permetterà al docente di creare il modello predittivo in maniera semplice, selezionando automaticamente il testo da analizzare dalle attività di Moodle e con la possibilità di visualizzare i risultati direttamente in piattaforma. Essere in grado di rilevare e gestire informazioni sui sentimenti provati dagli studenti, possono contribuire inoltre a conoscere le loro esigenze in quel determinato momento e agire modificando opportunamente i materiali o guidarli in percorsi personalizzati.

Riferimenti bibliografici

- [1] Amendola D., Miceli C. *Online Physics Laboratory For University Courses*. Je-LKS, 2016.
- [2] Shen L., Wang M., Shen R. *Affective e-learning: Using emotional data to improve learning in pervasive learning environment*. Journal of Educational Technology & Society, 2009, pp. 176–189.
- [3] Ortigosa A., Martín J.M., Carro R. M. *Sentiment analysis in Facebook and its application to e-learning*. Computers in Human Behaviour, 2014, pp. 527–541.
- [4] Hemalatha I., Saradhi Varma G.P., Govardhan A. *Sentiment Analysis Tool using Machine Learning Algorithms*. International Journal of Emerging Trends & Technology in Computer Science, 2013.
- [5] Gkontzis A.F., Karachristos C.V., Panagiotakopoulos C.T., Stavropoulos E.C., Verykios V.S. *Sentiment Analysis to Track Emotion and Polarity in Student Fora*. Proceedings of the 21st Pan-Hellenic Conference on Informatics, 2017.
- [6] Bishop C. M. *Pattern recognition and machine learning*. Berlin, Springer, 2006.
- [7] Rajaraman A., Ullman J.D. *Data Mining*. Mining of Massive Datasets, 2010, pp. 1-17.