
LEARNING ANALYTICS TO UNDERSTAND THE STUDENTS' SENTIMENTS

Fedela Feldia Loperfido [feldialop@gmail.com], Università degli Studi di Bari Aldo Moro, Anna Dipace [anna.dipace@unimore.it], Università degli Studi di Modena e Reggio Emilia, Alessia Scarinci [alessia.scarinci@uniba.it], Università degli Studi di Bari Aldo Moro, Italy

Abstract

What emotional experience can students live in digital mediated learning processes? In this paper we connect Learning analytics and Grounded theory to analyse the emotional presence of students in 11 courses within EduOpen (www.eduopen.org) MOOCs' platform. We analysed through a bottom up process and Nvivo 11 Plus software the forum dedicated to the students' self-presentation from all of the courses. By going ahead with the analysis, we defined a set of categories composed by a three-level system. At a more general level we have the macrodimensions "Sentiment about EduOpen" and "Emotions toward topics". Each of these dimensions is composed by some "child" categories and subcategories (nodes to Nvivo's language). After defining the entire set of categories and categorizing all the texts (circular process), we run some graphs on Nvivo showing the hierarchical structure of dimensions, the relations among dimensions and sources, and the clusters of dimensions by coding similarity. Results show how some courses are more composed by negative or positive sentiments and how the motivations dimension heavily characterizes the broad emotional dimension of students. In an evidence based action-research perspective, these results give interesting suggestions to personalize the learning activities proposed to students by EduOpen.

Abstract in Italian

Quale esperienza emotiva possono vivere gli studenti nei processi di apprendimento mediati dalle tecnologie? Nell'articolo colleghiamo Learning Analytics e Grounded theory per analizzare la presenza emotiva degli studenti in 11 corsi nella piattaforma MOOCs EduOpen (www.eduopen.org). Abbiamo analizzato attraverso un processo dal basso verso l'alto e il software Nvivo 11 Plus il forum dedicato all'autopresentazione degli studenti presente in tutti i corsi. Procedendo con l'analisi, abbiamo definito un insieme di categorie composte da un sistema a tre livelli. A un livello più generale abbiamo le macrodimensioni "Sentiment about EduOpen" e "Emozioni verso gli argomenti". Ognuna di queste dimensioni è composta da una serie di categorie e sottocategorie "secondarie" (nodi della lingua di Nvivo). Dopo aver definito la serie di categorie e classificato i testi, abbiamo eseguito alcuni grafici su Nvivo che mostrano la struttura gerarchica delle dimensioni, le relazioni tra dimensioni e fonti e i cluster di dimensioni codificando la somiglianza. I risultati mostrano come alcuni corsi siano maggiormente rappresentati da sentimenti negativi o positivi (sia verso l'argomento o la disposizione logistica del corso) e come la dimensione delle motivazioni caratterizza la vasta dimensione emotiva degli studenti. In una prospettiva di ricerca-azione basata sull'evidenza, i risultati offrono interessanti suggerimenti per personalizzare le attività di apprendimento.

Keywords: emotional experience, learning analytics, sentiment analysis, open learning, MOOC, Nvivo

Theoretical framework

This contribution connects three different fields: the area of learning analytics, the area of education specifically interested in digital mediated learning processes, and the approaches focused on the emotional dimension in learning. Namely, learning analytics is the measurement, collection, analysis and reporting of data about students and the contexts they learn through. The aim of learning analytics is to understand, personalize and optimize learning and the environments in which it occurs. Learning analytics are mainly used in learning

contexts mediated by the use of digital environments, since they can produce an amount of data about the traces each student or entire groups of learners leave online, successful activities, difficult experiences, and so on (Rienties & Rivers, 2014). In relation to the field of learning analytics, we stress the emotional dimension of learning as well. Speaking about feelings and emotions from a general and classical perspective, we can think that human beings can feel universal emotions, such as anger, disgust, fear, happiness, sadness, and surprise (Ekman, 1999) or joy-sadness, anger-fear, trust-distrust and surprise-anticipation (Plutchik, 2013). However, we can refer to emotion and, specifically, to emotions and learning, after answering the question “How can we define and understand emotions at a more specific level?”. According to Zembylas (2008), there is no agreement about what an emotion is and is characterized by. Indeed, emotions can be understood at least through three different perspectives: (a) Emotions as private and belonging to an intimate experience, as defined by psychodynamic approaches; (b) Emotions as sociocultural phenomena, as understood by social constructionist approaches; (c) Emotions as described by interactionist approaches, which transcend the dichotomies (e.g. mind/body, individual/social) established in the previous two and aims at bridging their differences. However, even if there is no a common definition of emotions, authors claim that they are not separated from the learning context (Lehman, 2006; Lipman, 1991). Coherently to this, for example, communities of inquiry (Garrison, Anderson, & Archer, 2000) are digital mediated learning experiences characterized by the cognitive presence, the social presence, the teaching presence and the emotional presence (Cleveland-Innes, & Campbell, 2012). This last is understood as the “emotional expression part of being socially present online” (ibid., p.272). If we still stay at this general layer, we can connect the interesting about the emotional dimension and the learning analytics by referring to Sentimental analysis, also known as Opinion mining looking for both negative and positive sentiments people have about the digital environment they use. However, this connection does not suggest how we can understand emotions at a more specific level. As for this point, Cleveland-Innes and Campbell (2012) approach the emotional experience of students through Grounded theory, that is by doing a content analysis of texts, looking for contents about emotions and defining a grid of categories through a bottom up process (from the text to the categories).

In this contribution, we connect both learning analytics and grounded theory to analyse the emotional experience of students in an online learning context made by eleven courses. This integrated system allows us exploring sentimental and emotional dimensions at macro, meso and micro levels of the context. At the same time, we also created a three-levels set of categories for the emotional analysis, composed by general dimensions, more specific categories and further subcategories.

Aims

- To explore the emotional processes experienced by students during the participation in MOOCs proposed by EduOpen (www.eduopen.org);
- To personalize the learning activities, according to students’ emotional experience.

Context and data

This research is supported by Unifg Tutoring – UniTutor project and the context of analysis is EduOpen, an international Moodle platform lead by the University of Foggia (IT). We can better describe the context by referring to the macro, meso and micro levels composing it. At a macro level, EduOpen is realized by 17 Italian Universities and several foreign partnerships. It started in 2014 and is an action-research project periodically rearranged thanks to evidence-based methods. Until now, it involved more than 70,300 learners from all over the world and proposed 140 courses. Indeed, the activities of EduOpen are online courses loaded on the Moodle based platform www.eduopen.org. Through a micro perspective, we can describe that each course refers to a specific topic (e.g. math for beginners, animals, English, and so on), and is managed by a university teacher and an online tutor of the EduOpen team. Furthermore, at the end of a

course, students receive a participation certification, an open badge or ECTS. More specifically, each course spends three-five weeks and is composed by:

- A self-presentation forum where students usually write down a post about themselves, the place they live, the wishes and expectations they have about the course, and so on;
- A number of MOOCs videotaped by the teacher and related to the topic of the course;
- Another forum where students can ask further explanations to the teacher;
- An evaluation section, where students fill in online tests during or at the end of the course.

At a meso level, we can say that all of the courses are categorized in different fields (such as, Literature, Science, and so on), in several pathways (an ensemble of courses connected each other by a main theme) and/or in the catalogue that a specific University partner proposes. In this paper, data are characterized by the self-presentation forums of all the courses managed by the University of Foggia (IT). These are 11 courses and have involved 43,345 students in total (10,277 of them completed the course they were unrolled in). Therefore, we especially look at the micro level of each course and at the meso level of the group of courses proposed by the University of Foggia.

Method of analysis

According to both Grounded Theory and Sentiment analysis approach, we:

1. Created a first general grid of analysis, composed by the two general dimensions “Positive sentiments” and “Negative sentiments” referred to the learning experience in the digital context;
2. Categorization of the texts through qualitative content analysis (Mayring, 1997), by using Nvivo 11 Plus;
3. Generation of further dimensions and their specific categories, emerging from the interaction between grounded approach and theoretical concepts;
4. Team discussion about the building of the grid and the categorization;
5. Checking of the categorization according the team discussion;
6. Analysis of the nodes (the categories to the software) by using Nvivo 11 Plus.

Results

During the analysis, we realized that the first version of the grid needed to be much more enriched. Therefore, we created a double grid, able to grasp three levels of the students’ emotional experience in the University of Foggia EduOpen courses. In other words, we defined two general dimensions: (a) “Sentiment about Eduopen”, grasping what students felt about Eduopen, its services and the arrangement of the courses; (b) “Emotions toward topic”, observing the feelings about the topic of the specific course students participated in. That is, the first dimension is about the feelings toward the digital environment, the concept of EduOpen, the arrangement of the environment. The second one refers to the feelings about the topic of the specific course. Furthermore, as Figure 1 and Figure 2 show, the category “Sentiment analysis” is composed by two more specific categories: “Negative sentiments” and “Positive sentiments”. These, in turn, are composed by other two subcategories for each (moderately/very negative; moderately/very positive). The figure shows the hierarchical relation among “parents” categories and “child” ones too, as elaborated through Nvivo.

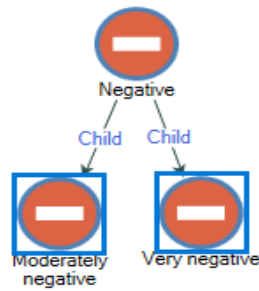


Figure 1. Negative sentiments to EduOpen child graph.
Negative sentiments has the two children nodes “Moderately negative” and “Very negative”

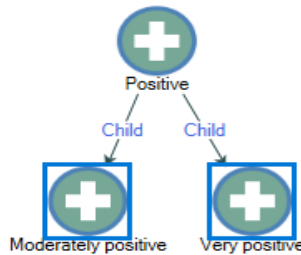


Figure 2. Positive sentiments to EduOpen child graph.
Positive sentiments has the two children nodes “Moderately positive” and “Very positive”

The dimension “Emotions to topic” was at the end shaped by a complex structure of categories. At a middle level, we grasped the three categories “Motivations”, “Negative sentiments” and “Positive sentiments (not to be confused with the two namesake categories “Positive” and “Negative sentiments” about the digital experience in EduOpen already described). “Motivations” refers to a category exploring a more cognitive dimension, even implying the students’ expectations about the contents of the course and the reason why they are going to attend the course. Indeed, it is composed by seven specific or “child” categories. “Negative sentiments” is about the feelings students have against the content of the course and is composed by five specific or “child” categories. “Positive sentiments” is about the good feelings students have toward the content proposed by the course and is shaped by five specific or “child” categories. In Table 1, we describe all the categories composing “Emotions to topic” (a graph like Figure 1 and 2 would be more impressive, but we think the table is more effective).

Table 1: Categories composing “Emotions to topic”

Macro level category	Meso level category	Micro level category (and eventual description)
Emotions to topic	Motivations	Deepen knowledge (to go in depth in the topic the course refers to)
		Home learning (participate because you can attend the course staying at home)
		Innovative methods (to be tried)
		Mind training
		Old knowledge renewal
		Practical effects (in daily job activities)
		Support to learning (of other contemporary learning experiences)
	Negative sentiments	Disorientation
		Fear
		Feeling in trouble
		Nostalgia (about past learning experiences on the same topic)
		Sense of unfinished
	Positive sentiments	Discovery and curiosity
		Enthusiasm

Feel interest
Hope (to better understand the contents in opposition to
past experiences)
Passion

After creating the final grid of analysis by making the categorization, we checked them (the grid and the first categorization) by a team discussion, until we reached a total agreement about both. At the end, we analysed the nodes and their relationships with the sources (the texts of the forum) by elaborating some graphs through Nvivo 11 Plus. The following graphs (Figure 3, 4, 5, 6) and their respective descriptions show the analysis we made, which we will go back to in the conclusions as well. Figure 3 suggests that, in the general dimension “Sentiment to EduOpen”, the category “Positive sentiments” is much more prominent than the one about negative sentiments. Furthermore, the moderately positive sentiments are more present in the texts than the high positive ones. Figure 4, instead, shows what the relations between nodes and sources are. As it is visible, in eight forums referring to the respecting courses (Biochemical pills, Math for absolute beginners, Law history and philosophy, Animals, Knowing History, History of Italian literature, Course of general mathematics, Tourism marketing through digital media) students express both positive and negative sentiments about the structure of the course and/or EduOpen as a learning experience. Furthermore, in the document of “Pedagogy and education, basic concept” course there are just positive sentiments’ references; whereas, in the course about Physics and Basic general pathology there are no sentiment expressions.

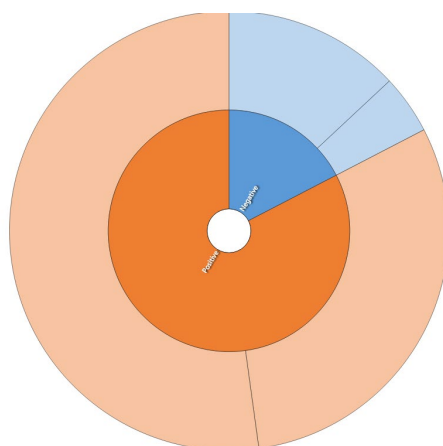


Figure 3. “Sentiment to EduOpen” hierarchical graph.
Dark orange section represents Positive sentiments in total, whereas the dark blue one represents Negative sentiments. The smallest light orange section is about the highly positive sentiments; the smallest light blue sections is about the highly negative sentiments.

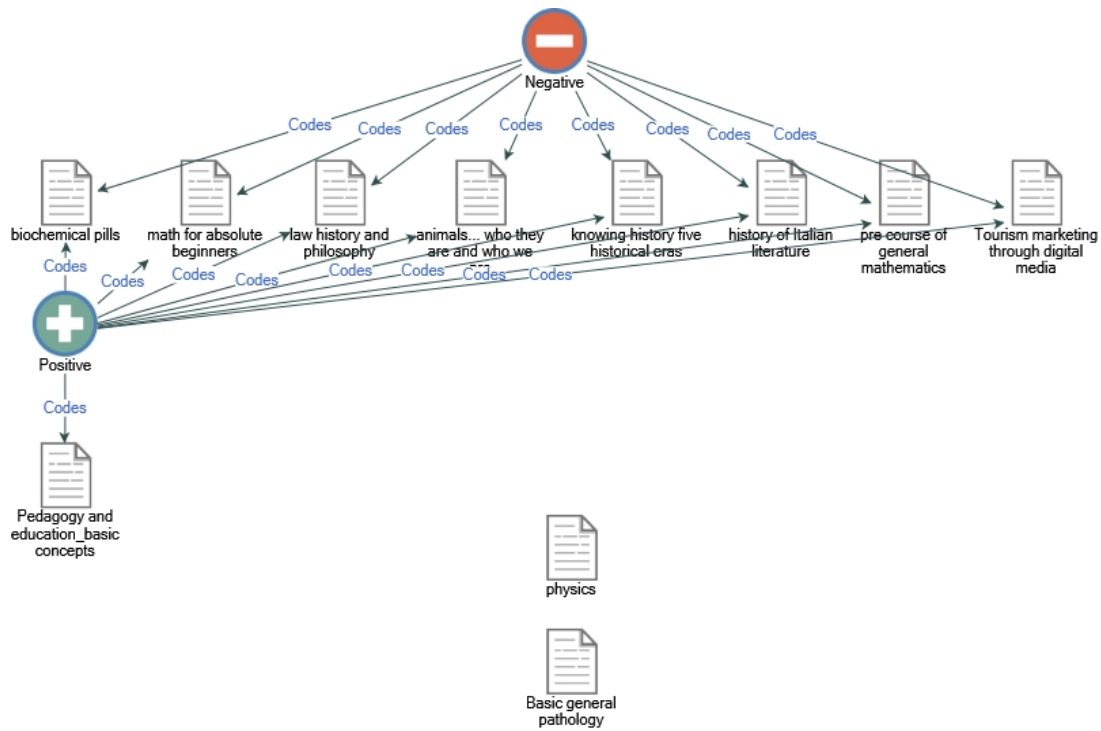


Figure 4. "Sentiment to EduOpen – sources" project map.

The red circle represents Negative sentiments; the green circle represents Positive sentiments. Arrows show the relation between each dimension and the forum of the specific course, that is if there are coded units of the text by using the dimensions.

What about the macrodimension "Emotions to topic"? Figure 5 shows that the "Motivation" mesocategory is the richest one, followed by "Positive sentiments" and then by "Negative sentiments", suggesting that the more cognitive aspects have a higher incidence in the texts.

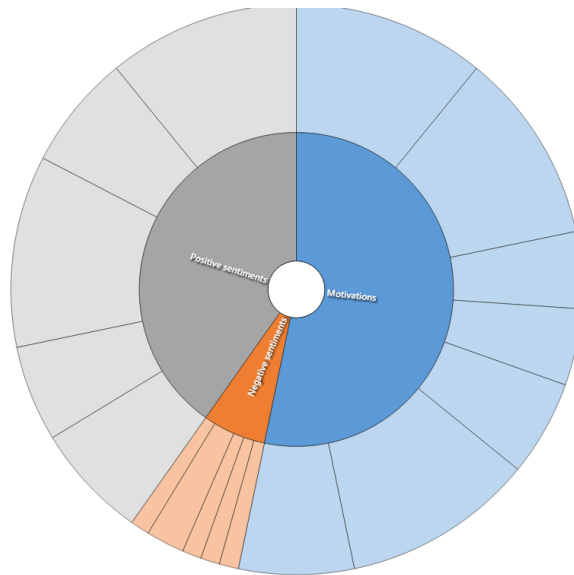


Figure 5. "Emotions to topic" hierarchical graph.

The blue section is about Motivations, the grey section is about Positive sentiments and the orange section is about negative sentiments toward the topic.

Figure 6, instead, describes the connections between codes and sources. As it can be seen, the category "Motivation" is related to all of the sources, whereas the category "Positive sentiments" is used on all of the courses' texts except than in "Physics". Negative sentiments are involved in just three sources (Math for absolute beginners, Law History, Pedagogy and Education. Basic concepts).

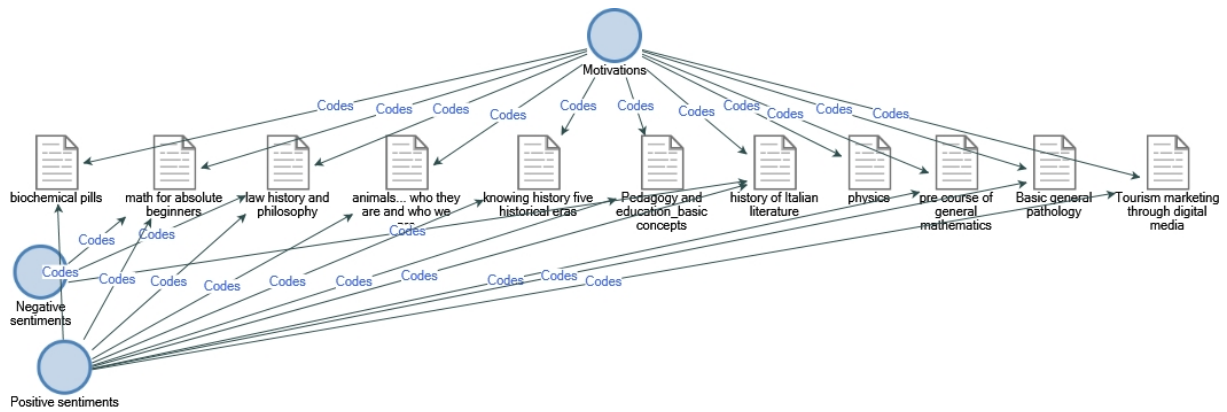


Figure 6. "Emotions to topic - sources" project map

With further analysis, the figures of them are not showed here because of the small space, we clusterized both sources and codes by coding similarity. As results, it emerged that "Motivation" and "Positive emotions" are more similar categories, and that "Physics" and "Basic genetic pathology" are the most distant sources form the others. These further results obtained by the cluster analysis mainly confirm the previous ones.

Conclusions and implications

In this contribution, we made a sentimental analysis in terms of both negative and positive opinions students have about the learning experience they are going to attend or just began on EduOpen. We also realized a more specific emotional analysis about the feelings learners have for the specific topic of the course they choose. We used a grounded theory approach to grasp the set of dimensions, categories and subcategories about emotions arising from the texts through a bottom up research process. According to the main results, the emerging set of categories is a very complex one and is composed by some clusters of similarity coding. By looking at the hierarchical graph about sentimental analysis, we can see that in general positive sentiments characterize the learners' perception about the experience in EduOpen. At the same time, the mesodimension "Motivations" has a prominent space in the hierarchical graph about the emotions connected to the topic of the course. By going in depth in the categories, there emerges that some of them are about intrinsic motivations (e.g. to deepen the student's knowledge) and others are about external ones (e.g. to have a support for the university exams). However, cluster analysis shows that this last category is quite similar to category "Positive feelings" in terms of coding similarity. It seems, therefore, that students attending the courses have different motivations to participate in them, but they also feel positive emotions related to such a participation. Particularly interesting are the courses "Pedagogy and education" and "Physics". The first one, indeed, does not have negative references in the dimension "Sentiment analysis", whereas the second one is coded just by using the category "Motivation". Furthermore, there are three courses having references about negative sentiments related to the topic. We find all these results very much interesting for different reasons. Far from generalize a so specific study, we do claim that the entire set of categories shows how complex is the emotional experience of students. This is not just due to the number of categories shaping the set, but also to the three levels characterizing it, the relationships among them and the contextualized value they have in the different educational experiences. These results can have implications in the arrangement of the activities and in the personalization of the learning process, since an organization taking care of the specific emotions students feel can make the learning aims more effective. At the same time, further more specific analysis can give justice to the complexity of the students' emotional presence. Indeed, next studies will analyse the possible statistical correlation in the relations codes-dimensions and codes-sources, and the direction of such relations. Furthermore, we will analyse the forums of other EduOpen courses in order to broader the study to the macro entire context of EduOpen, and to create methodological tools connecting the usual learning analytics' quantitative perspective and the qualitative dimension shaping the emotional experience of students.

References

1. Cleveland-Innes, M., & Campbell, P. (2012). Emotional Presence, Learning, and the Online Learning Environment. *The international Review of Research in Open and Distance Learning*, 13(4), 269-292.
2. Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45-60). SF: John Wiley & Sons Ltd.
3. Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2, 1–19.
4. Lehman, R. (2006). The role of emotion in creating instructor and learner presence in the distance education experience. *Journal of Cognitive Affective Learning*, 2(2), 12– 26.
5. Lipman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.
6. Mayring, P. (1997). *Qualitative Inhaltsanalyse*. Weinheim: Deutscher Studien Verlag.
7. Plutchik, R. (2013). Emotion: Theory, research, and experience. In R. Plutchik & H. Kellerman (Eds.), *Theories of Emotion* (pp. 10-24). New York: academic Press.
8. Rienties, B., & Rivers, B. A. (2014.). Measuring and understanding learner emotions: evidence and prospects. *Learning Analytics Review*, 1, 2-30.
9. Zembylas, M. (2008). Adult learners' emotions in online learning. *Distance Education*, 29(1), 71–87.