Investigating negative reviews and detecting negative influencers in Yelp through a multi-dimensional social network based model

(Article begins on next page)

DOI: https://doi.org/10.1016/j.ijinfomgt.2021.102377

Investigating negative reviews and detecting negative influencers in Yelp through a multi-dimensional social network based model

Enrico Corradini[1], Antonino Nocera[2], Domenico Ursino[1*], and Luca Virgili[1]

[1] DII, Polytechnic University of Marche,
[2] DIII, University of Pavia
[*] Contact Author
e.corradini@pm.univpm.it; antonino.nocera@unipv.it; d.ursino@univpm.it;
l.virgili@pm.univpm.it

Abstract

In this paper, we propose an investigation of negative reviews and define the profile of negative influencers in Yelp. The methodology adopted to achieve this goal consists of two phases. The first one is theoretical and aims at defining a multi-dimensional social network based model of Yelp, three stereotypes of Yelp users, and a network based model to represent negative reviewers and their relationships. The second phase is experimental and consists in the definition of five hypotheses on negative reviews and reviewers in Yelp and their verification through an extensive data analysis campaign. This was performed on Yelp data represented by means of the models introduced during the first phase. Its most important result is the construction of the profile of negative influencers in Yelp. The main novelties of this paper are: *(i)* the definition of the two social network based models of Yelp and its users; *(ii)* the definition of three stereotypes of Yelp users and their characteristics; *(iii)* the construction of the profile of negative influencers in Yelp.

Keywords: Yelp; Multi-dimensional model; Negative influencers; Negative reviews; Social Network Analysis; User stereotypes; Homophily

1 Introduction

Yelp[1] is a business directory service and a crowd-sourced platform designed to help users find businesses like restaurants, hotels, pet stores, spas, and many more. It is one of the most widely used review platforms on the Web. It ranks 9^{th} on the RankRanger list of the top 100 leading websites by traffic[2], with approximately 800 million visits per month. In addition of being a business search and review platform, Yelp is also a social network, because it allows its users to specify their friendships. Finally, it is also a business directory, because it groups businesses into categories and sub-categories.

[1] `https://www.yelp.com`
[2] `https://www.rankranger.com/top-websites`

The success of Yelp has prompted many researchers to investigate this platform (Agarwal et al., 2019; Arthur et al., 2019; Lim and Heide, 2014; Nokhiz and Li, 2017; Gulati and Eirinaki, 2019). Several studies have striven to understand how rates are assigned to businesses (Hu et al., 2014; Lei and Qian, 2015; Sun and Paule, 2017; Singh et al., 2019), and many others have focused on the analysis of the content of text reviews from both a structural and a linguistic viewpoint (Parikh et al., 2014, 2017; Bauman and Tuzhilin, 2014). Some papers have studied Yelp reviews by adopting sentiment analysis-based techniques (Nakayama and Wan, 2019; Angelidis and Lapata, 2018; Guerreiro and Rita, 2019. Forthcoming). Others have focused on identifying strategies for the detection of fake reviews and rates (Luca and Zervas, 2016; Mukherjee et al., 2013; Malbon, 2013; Lee et al., 2018), or have investigated how people search for information (Hicks et al., 2012). Furthermore, some authors have investigated Yelp through the concepts provided by Social Network Analysis, like homophily (McPherson et al., 2001), to study the social influence existing among friends (Qiu et al., 2020). Some researchers have employed these results to outline the decision making of users on purchases (Zhang et al., 2014), while other ones have studied the possible impacts of electronic Word of Mouth (eWOM) in online businesses (Cheung and Lee, 2012; Cheung and Thadani, 2012; Aggarwal et al., 2012). Several studies have explored the causes leading people to publish reviews (Ho et al., 2008), while others have analyzed reviewer strategies to improve their effectiveness (Forman et al., 2008; Shen et al., 2015; Li et al., 2017). Finally, some studies have focused on the analysis of review usefulness (Yin et al., 2014; Schuff and Mudambi, 2012; Kumar and Benbasat, 2006), while others have investigated the differences between positive and negative reviews (Yin et al., 2016; Knoll and Matthes, 2017).

A phenomenon that represents a hot topic for both Yelp and all review platforms is the analysis of negative reviews (Berger et al., 2010). This topic is extremely important not only for the consequences it has in practice, but also from a more theoretical point of view. In fact, it is well known that the Likert scale, which the Yelp reviews and the corresponding scores are based on, is positively biased (Alexandrov, 2010; Peeters and Czapinski, 1990; Bertram, 2007). As a consequence, the presence of negative reviews is a really important problem indicator for a business and, consequently, a valuable piece of information (Kumar and Benbasat, 2006; Li et al., 2017). Indeed, negative reviews can provide much more information, knowledge and improvement possibilities than positive ones (Chang et al., 2019). For this reason, many researchers have already investigated the role of ratings and reviews on businesses, along with their social implications (Ting et al., 2017; Luca, 2016).

Despite the numerous studies on Yelp that have been presented in the past literature, to the best of our knowledge, no paper has proposed a multi-dimensional model capable of best capturing the specificity of Yelp to be at the same time a review platform, a social network and a business directory. Moreover, no paper has proposed a study focused entirely on negative reviews on Yelp that, starting from a representative model of them, could define several stereotypes of users and, hence, build the profile of negative influencers. This paper aims at filling this gap.

Specifically, we first define a multi-dimensional social network based model for Yelp and then use this model to study negative reviews and build a profile of negative influencers in this social medium. We decided to adopt this model because it perfectly fits the specificities of Yelp mentioned above. In fact, our model represents Yelp as a set of 22 communities, one for each macro-category of this social platform (modeling Yelp as a business directory). At the same time, it represents Yelp as a social network, whose nodes indicate users and whose arcs denote the relationships between them. These can be of different types. For example, they can denote friendships between users (modeling Yelp as a social network), or the action of co-reviewing the same business (modeling Yelp as a review platform). Through the concepts and techniques of Social Network Analysis applied to our multi-dimensional model, our approach defines three stereotypes of Yelp users, namely the bridges, the double-life users and the power users. These stereotypes can help the detection of the negative influencers in Yelp and the definition of a profile for them. Both our model and the user stereotypes represent theoretical contributions of our paper. These last are completed by a Negative Reviewer Network, which allows us to investigate the main characteristics of the negative influencers in Yelp.

Among the possible questions that can be answered thanks to our approach, in this paper we focus on the following ones: *(i)* What about the dynamics leading a Yelp user to publish a negative review? *(ii)* How can the interaction of these dynamics increase the "power" of negative reviews and people making them? *(iii)* Who are the negative influencers in Yelp?

The practical implications of negative reviews and influencers we find in this paper have a large variety of applications. First of all, it was proved that negative reviews have a stronger effect on businesses than positive ones (Aggarwal et al., 2012). Furthermore, influencers play a crucial role for the successful placement of products in a social network. So, it is important to know who are the negative influencers that could damage a business, in order to strive to turn them into neutral, or even positive, influencers (Yin et al., 2016; Zhang et al., 2014). Finally, gaining trust through online reviews can help a business gather venture capitals for its growth (Fogel and Zachariah, 2017; Kumar and Benbasat, 2006). As a matter of fact, reviews are consumer opinions, unfiltered by traditional media, more sincere and imperfect (Aggarwal et al., 2012; Cheung and Lee, 2012). For this reason, a proper coverage of positive reviews can attract more financiers (Aggarwal et al., 2012; Cheung and Thadani, 2012; Knoll and Matthes, 2017). On the other hand, negative reviews and influencers can drive potential investors away from investing in a company (Luo, 2009).

The outline of this paper is as follows. In Section 2, we present related literature and highlight the main novelties of our approach with respect to the past ones. In Section 3, we describe the theoretical background and hypotheses development. In Section 4, we present the methodology we adopted during the investigation activity. In Section 5, we illustrate the results obtained. In Section 6, we propose a discussion and a synthesis of them, their implications, and possible future research directions. Finally, in Section 7, we draw our conclusions.

2 Related Literature

Over the years, researchers have focused on Yelp as a reference platform for studying how users interact with each other and build cooperative social groups. Their research efforts have also been supported by the social medium itself, which has made available a complete snapshot of its data to foster comprehensive analyses on it (Cui, 2015). Many authors have used this snapshot to investigate the role of ratings and reviews on businesses and their social implications (Ting et al., 2017; Luca, 2016). Researchers have also analyzed how people search for information on Yelp (Hicks et al., 2012) and what aspects (including uses and rewards) lead them to employ this platform.

Several authors have investigated Yelp using Social Network Analysis (SNA, for short) (Qiu et al., 2018, 2020). For instance, the authors of (Qiu et al., 2020) rely on the concept of homophily (McPherson et al., 2001) to study the social influence possibly existing between users and, in particular, between friends. Starting from the results obtained, they propose the construction of the profile of an influencer in Yelp. The authors of (Qiu et al., 2018) focus on the role of friendship in this social medium. Specifically, they investigate the impact of social relationships from the consumer's side and find that these relationships exert a significant impact in those consumers having at least one common purchase.

As for the analysis of social relationships, several studies have been conducted in both Yelp and other social platforms to understand how users perceive their social contacts and how they influence their acquaintances (Lim and Heide, 2014; Nokhiz and Li, 2017; Gulati and Eirinaki, 2019; Nam et al., 2017; Shen et al., 2015; Kang et al., 2013; Zhou et al., 2019; Zhang et al., 2021). For example, the authors of (Nam et al., 2017) propose an approach to analyze a large set of brand associations obtained from social tags for marketing research. They apply well-known text mining techniques to understand consumers' perceptions of brands starting from social tagging data. The authors of (Cheung and Lee, 2012) analyze a dataset obtained from `OpenRice.com`, a crowd-sourced social medium for restaurant reviews in Hong Kong and Macau. The authors of (Forman et al., 2008) show that online community members rate reviews containing descriptive identity information more positively. Indeed, a disclosure of personal information on an online review system leads to a greater volume of sales. The authors of (Shen et al., 2015) aim at understanding how online reviewers compete to acquire the attention, typically scarce, of users. They propose a theory explaining the strategies adopted by online reviewers in choosing the right product and the right rate when posting reviews. As far as Yelp is concerned, the authors of (Lim and Heide, 2014) investigate the effects of the review rate, the reviewer profile, and the receiver familiarity with the platform, on the credibility of a review on this social medium. Moreover, the authors of (Nokhiz and Li, 2017) find a strong correlation between the moral attitude of a community of users and their tendency to express low rates and negative reviews in case some moral foundation is violated. As for the investigations of social relationships in social media, another interesting topic concerns information diffusion (Aslay et al., 2018; Xuan et al., 2019. Forthcoming; Kim et al., 2018; Bhowmick et al., 2017; Lin and Wang, 2020). In the analysis of this topic, an increasing number of researchers are studying the role not only of classic and direct relationships, such as friendship, but also several other ones, such as co-posting or homophily of interests (i.e., having interest in the same topics) (Saxena et al., 2019. Forthcoming; Bhanodia et al., 2019).

In all previous approaches, the reviews considered are general (i.e., they could be positive or negative). However, to our end, negative reviews and reviewers are worth a special attention. The importance of negative reviews in the analysis of social platforms has been investigated in the recent scientific literature by highlighting their impact in social contexts, along with the mechanisms leading users to make them (Nakayama and Wan, 2019; Fogel and Zachariah, 2017; Setyani et al., 2019; Arthur et al., 2019; Agarwal et al., 2019). In these studies, researchers point out that dealing with negative reviews is a fundamental task in review-based platforms for business operators (Kumar and Benbasat, 2006; Li et al., 2017). In fact, it was empirically shown that answers and justifications to negative rates contribute to the increase of trust between users and businesses (Fogel and Zachariah, 2017), and that users tend to perceive reviews confirming their initial beliefs as more helpful (Yin et al., 2016). Several studies focus on the key factors making a review helpful (Schuff and Mudambi, 2012; Fogel and Zachariah, 2017), while others show that negative reviews are more useful and can influence user opinions more than positive ones (Basuroy et al., 2003; Cao et al., 2011). In this perspective, the authors of (Zhang et al., 2014) propose a model to identify the key elements leading customers to make their decisions; this model was empirically tested with 191 users of an existing online review site. Furthermore, the authors of (Aggarwal et al., 2012) use the VentureExpert database to gain knowledge on a sample of famous businesses. The authors of (Ho et al., 2008) formalize a metric, called disconfirmation, measuring the discrepancy between the expected evaluation of a product and the one assigned by experts or other people. The authors of (Fogel and Zachariah, 2017) study a set of variables to evaluate the users' intention of employing Yelp, as well as their behavior in using a service or purchasing a product after reading Yelp reviews. Finally, the authors of (Arthur et al., 2019) analyze the reviews made by hospital patients in order to identify a common language correlated with negative and positive reviews.

An important aspect to consider when using Social Network Analysis for evaluating reviews and reviewers is the fact that user relationships in a social network are often heterogeneous (Cai et al., 2005). For this reason, many studies have proposed to decompose social media into different networks of relationships. Indeed, multi-relationship networks have been extensively studied in the past (Davis et al., 2011; Yang et al., 2012; Zhang et al., 2013). For example, the authors of (Zhang et al., 2013) combine the analysis of the friendship network and the author-topic one, both constructed starting from the information available in an online social network. Instead, the authors of (Yang et al., 2012) focus on a co-authorship network and consider different types of relationships, i.e., co-authorship, co-participation to the same edition of a conference, and geographic proximity.

In multi-relationship networks, the classical definition of influencer is extended because the role of such users is not bound to communities derived from a single category of relationships. Instead, it also includes the capability of providing information diffusion channels among different networks, one for each type of relationships. To refer to this extended definition of influencer, the term "bridge" is often adopted. In the past literature, several studies have been devoted to investigate the role of bridges in the formation of social communities. For instance, the authors of (Kavanaugh et al., 2005) show that users with a weak connection bridging heterogeneous groups have higher levels of community commitment, civic interest, and collective attention than the other ones. Furthermore, the authors of (Granovetter, 1973) prove that Internet users, who bridge heterogeneous online communities by means of weak ties, have a high social engagement, use the Internet for social purposes, and are prone to become members of new social communities. The interest towards users serving as bridges among communities has increased over the years and, indeed, several studies have been done to analyze the behavior and peculiarities of such users in complex networks (Franks et al., 2008; Shi et al., 2007; Leskovec et al., 2007; Berlingerio et al., 2011, 2013).

Some studies have also analyzed the behavior of users serving as bridges among different social networks (Buccafurri et al., 2013, 2014, 2015). Here, the concept of community is brought to the edge, because it is mapped to a whole social network. Specifically, the authors of (Buccafurri et al., 2013) report a complete identikit of users bridging different social networks. The authors of (Buccafurri et al., 2014) leverage the peculiarities of bridge users to define a new crawling strategy to sample a multi-social network environment. Finally, the authors of (Buccafurri et al., 2015) perform a comparative study of users serving as bridges among two of the most famous social networks, namely Facebook and Twitter.

From the above description, it can be seen that, in the literature, there is an impressive number of papers dealing with issues similar to those analyzed in this paper. However, none of them proposed a multi-dimensional social network based model for Yelp, capable of representing the specificity of this social platform of being simultaneously a review platform, a social network and a business directory. The presence of this model would allow us to answer the following research question: What about the dynamics leading a Yelp user to publish a negative review? Furthermore, no paper proposed a study focused entirely on negative reviews and reviewers in Yelp, which, starting from a social network based model representing them, could define a set of stereotypes of users publishing negative reviews. Having all this available would allow us to answer the following research question: How can the interaction of the dynamics driving negative reviewers increase their "power" and the one of their reviews? Finally, no past paper built a profile of a negative influencer in Yelp. Reaching this result would allow us to answer the following research question: Who are the negative influencers in Yelp? This paper aims at filling this gap and answer the three research questions mentioned above.

Our paper draws inspiration from the research strands mentioned previously. First of all, our multi-dimensional social network based model of Yelp can be employed to handle different relationships (e.g., friendship, co-review). In particular, it is possible to define an occurrence of the model for each relationship. This way of proceeding falls within the context of multi-relationship networks, but in a new way. In fact, differently from past multi-relationship models, ours does not require the prior and static definition of the relationships to represent, but allows a dynamic choice of them, based on the analysis to be performed. For example, in this paper, we have chosen friendship and co-review between Yelp users. Furthermore, the choice of including in our model the macro-categories in which the businesses are grouped in Yelp represents an additional feature of it. It makes possible a definition of the bridge concept perfectly fitted on Yelp, which, in turn, allows for the definition of three user stereotypes for this social platform. Therefore, the multi-dimensionality of our model enables an analysis of Yelp users and their relationships from multiple orthogonal viewpoints, acting simultaneously and influencing each other.

Our multi-dimensional social network based model makes our definition of bridge possible. Starting from that definition, and operating on the model itself, we define three user stereotypes, namely: *(i)* the k-bridge, i.e., a person who reviewed businesses belonging to k different Yelp macro-categories; *(ii)* the power user, i.e., a person very active in all the macro-categories in which she is interested; *(iii)* the double-life user, i.e., a person exhibiting different behaviors in the different macro-categories in which she operates. Compared to the generic stereotypes presented in the past literature (Buccafurri et al., 2012), those identified in this paper are tailored to Yelp and, therefore, can provide a more specific contribution in the definition of the profile of negative influencers in this social medium.

Having the multi-dimensional model, the three stereotypes and the Negative Reviewer Network at disposal, our approach can investigate negative reviews and reviewers and can build a profile of negative influencers. These tasks are very important because it was shown that the effect of negative reviews and reviewers is much greater than the one of positive reviews and reviewers (Aggarwal et al., 2012). Furthermore, negative reviews and reviewers are not very common because people tend to give high ratings to businesses (Bertram, 2007; Potamias, 2012). But for this very reason, the information they bring is extremely valuable. Indeed, consumers and businesses are prone to rely on negative reviews and reviewers to understand the reasons for possible dissatisfaction caused by a product, a service or a business (Arthur et al., 2019; Agarwal et al., 2019).

Compared to the works on negative reviews and reviewers described above, our approach is more focused on the issue of influence, more specifically on negative influence. In this context, it offers a first important contribution thanks to the definition of the Negative Reviewer Network. This tool allows the exploitation of Social Network Analysis techniques to investigate the influence of a negative reviewer on other users. We point out that the Negative Reviewer Network is general and can be used to investigate the same issue in other review platforms. Starting from it and the multi-dimensional model introduced in this paper, which is instead specific to Yelp, our approach provides a second important contribution, i.e., it constructs the profile of a negative influencer in Yelp. Such a profile is perfectly fitted on this social platform because it takes into account both the partitioning of Yelp into macro-categories and the possibility to specify user friendships, provided by this platform.

3 Theoretical background and hypothesis development

Our multi-dimensional investigation of negative reviews and detection of negative influencers in Yelp is possible thanks to a new multi-dimensional social network based model of Yelp. This model starts from the observation that, in this social medium, businesses are organized according to a taxonomy consisting of four levels. Level 0 includes 22 macro-categories. Each macro-category has one or more child categories; therefore, level 1 includes 1002 categories. A category may have zero, one or more sub-categories; as a consequence, level 2 comprises 532 sub-categories. Finally, level 3, has only 19 sub-sub-categories; indeed, most sub-categories are not further categorized. Our model represents Yelp as a set of 22 communities, one for each macro-category:

$$\mathcal{Y} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{22}\}$$

Given the macro-category $\mathcal{C}_i$, $1 \leq i \leq 22$, a corresponding user network $\mathcal{U}_i = \langle N_i, A_i \rangle$ can be defined. N_i is the set of the nodes of $\mathcal{U}_i$; there is a node n_{i_p} for each user u_{i_p} who reviewed at least one business of $\mathcal{C}_i$. A_i is the set of the arcs of $\mathcal{U}_i$; there is an arc $a_{pq} = (n_{i_p}, n_{i_q}) \in A_i$ if there exists a relationship between the users u_{i_p}, corresponding to n_{i_p}, and u_{i_q}, corresponding to n_{i_q}.

Finally, an overall user network $\mathcal{U} = \langle N, A \rangle$ corresponding to $\mathcal{Y}$ can be defined. There is a node $n_i \in N$ for each Yelp user. There is an arc $a_{pq} = (n_p, n_q) \in A$ if there exists a relationship between the users u_p, corresponding to n_p, and u_q, corresponding to n_q.

In the definition of $\mathcal{U}$ (and, consequently, of $\mathcal{U}_i$), we do not specify the kind of relationship between u_p and u_q. Actually, it is possible to define a specialization of $\mathcal{U}$ for each relationship we want to investigate. In this paper, we are interested in two relationships existing between Yelp users, namely friendship and co-review. As a consequence, we define two specializations of $\mathcal{U}$, namely $\mathcal{U}^f$ and $\mathcal{U}^{cr}$. $\mathcal{U}^f$ is the specialization of $\mathcal{U}$ when we consider friendship as the relationship between users, whereas $\mathcal{U}^{cr}$ denotes the specialization of $\mathcal{U}$ when co-review (i.e., reviewing the same business) is the relationship between users.

Starting from this model, it is possible to define some Yelp stereotypes, namely: *(i) the k-bridge*, i.e., a person operating in k categories of Yelp; *(ii) the power user*, i.e., a person very active in all the categories that she is interested in; *(iii) the double-life user*, i.e., a person showing different behaviors in the different categories she attends. Her different behaviors can regard the activity level (*access-dl-user*) or the severity of her reviews (*score-dl-user*). These stereotypes can lead to the detection of negative influencers in Yelp. We formalize them in Section 4. We have introduced them here in that their concepts are necessary to understand the following of this section.

Starting from this theoretical background, we aim at answering the three questions mentioned in the Introduction. In particular, we use the above model and stereotypes to design and perform a social network analysis-based campaign aiming at evaluating some hypotheses that we synthesize in the following:

- First of all, the review mechanism of Yelp is based on a scale from 1 to 5 stars. This is similar to the review mechanisms encountered in several other social media. In this context, we formulate the following:

  Hypothesis 1 (H1) - The star-based review system of Yelp is positively biased.

  In the scale adopted by Yelp, 1 means "absolutely bad" and 5 means "fantastic". A review with 2 stars is still negative, but 3 stars already denote a positive review. In other words, the review mechanism of Yelp makes it more probable that users release positive reviews. Unless the experience was really bad, the review will almost always be positive. This is confirmed by how Yelp itself labels the stars (1 - "Eek! Methinks not"; 2 - "Meh. I've experienced better"; 3 - "A-OK"; 4 - "Yay! I'm a fan"; 5 - "Woohoo! As good as it gets!").

  On the other hand, if we consider this review mechanism from a more formal and theoretical viewpoint, we can observe that it is based on a Likert scale, which was already shown to be asymmetric and positively biased (Alexandrov, 2010; Peeters and Czapinski, 1990; Bertram, 2007).

- We think that the stereotypes introduced above can help very much in evaluating negative reviews and influencers. As for a specific kind of stereotype, i.e., the double-life users, we formulate the following:

  Hypothesis 2 (H2) - access-dl-users and score-dl-users play a key role in negative reviews.

  To understand the reasoning behind this hypothesis, consider score-dl-users. Clearly, they can be partitioned into two sets. The former is made up of users who mainly write positive reviews and few negative reviews. These are basically positive users who, for some reasons, had a bad experience with some businesses. So, what drove them to write negative reviews, considering that they are keen to write positive ones? A user assigns a 1-star score to a business when her expectations were not satisfied. This was already investigated in literature (see, for instance, (Ho et al., 2008)), where it was proved that a high discrepancy between the others' opinions and the experience of a user is the main driver for her to write a negative review.

  The latter set of access-dl-users is much more peculiar. It comprises those users who generally write negative reviews but, in some cases, release positive ones. These users have probably developed very severe criteria for evaluating businesses, leading them to be satisfied only rarely.

- We have already discussed about the multi-dimensionality of our model. One of its main dimensions is friendship. Actually, it is well known that this relationship plays a key role in social networks (Bhowmick et al., 2017; Saxena et al., 2019. Forthcoming; Bhanodia et al., 2019). Starting from these results, it is reasonable to formulate the following:

  Hypothesis 3 (H3) - A user has a strong influence on her friends when doing negative reviews.

  This could seem obvious. In past literature it has been proved that users are influenced by others when writing reviews. In particular, it has been found that users tend to have a positive opinion of a product/service if it has been positively commented by other users (Cheung and Lee, 2012).

  In addition, people generally trust more those users sharing their personal profile on online review platforms (Forman et al., 2008). It was found that a personal information disclosure is crucial for the spread of positive comments about a product/service, because the possibility of associating information with a particular person gives a boost in the overall perceived confidence. All of this is amplified when users share a common geographical location. This reasoning can also be applied to relationships like friendship, because personal information is certainly disclosed between friends.

  Here, we hypothesize that the influence exerted by friends is valid not only for positive reviews but also for negative ones, possibly leading to a phenomenon of negative influence between friends.

- Another stereotype introduced above that could play an important role as negative influencer is the bridge one. As for it, we formulate the following:

  Hypothesis 4 (H4) - Bridges have a much greater influence power than non-bridges.

  If Yelp can be modeled as a network of different communities, each corresponding to a given business macro-category, it is immediate to think of bridge users as special ones, capable of facilitating information diffusion from a community to another. Bridge users have a position of power in the network, and this power can even be measured (Ke-Jia et al., 2020). If we look at classical centrality measures in social network analysis, it is easy to argue that bridge users have a high betweenness centrality value. On the other hand, if we look at reviews, it is plausible that a bridge could expand the negative conception of a brand from a category to another which both the bridge and the brand belong to.

- The previous reasoning about the correlation between bridges and betweenness centrality paves the way to think that centralities play a key role in the diffusion of negative reviews. In particular, it is reasonable to make the following hypothesis:

  Hypothesis 5 (H5) - There is a correlation between degree and/or eigenvector centrality and the capability of being negative influencer.

Degree centrality tells us which nodes have the highest number of relationships in a network. These are probably power users, if we consider our stereotypes. They certainly are important users, because they are densely connected. On the other hand, eigenvector centrality can help us to identify influential users, who do not like to appear as such (the so called grey eminences or grey cardinals). Those kinds of users are often connected to few nodes, each having a high number of relationships with the other users (Maharani et al., 2014). These two centrality measures can be useful to find negative influencers in Yelp.

4 Methodology

As we have seen above, our methodology starts from the multi-dimensional social network based model introduced in Section 3, formulates some hypotheses and aims at verifying them using an inferential campaign based on social network analysis. This campaign makes use of a number of concepts, stereotypes and definitions that we introduce in this section. Instead, the way they are exploited to prove the hypotheses and, more in general, to extract useful knowledge is described in Section 5.

The first concept we introduce is a stereotype, namely the *k-bridge*. Specifically, a *k-bridge* is a Yelp user who reviewed businesses belonging to exactly k different macro-categories of Yelp. A user who reviewed businesses of only one macro-category is a *non-bridge*. Finally, we use the generic term *bridge* to denote a k-bridge such that $k > 1$. Given a k-bridge u_p of $\mathcal{U}$, where $\mathcal{U}$ is the overall user network corresponding to Yelp and introduced in Section 3, there are k nodes $n_{1_p}, n_{2_p}, \cdots, n_{k_p}$ associated with her, one for each macro-category containing at least one review performed by her.

After having introduced the *k-bridge*, we present some other stereotypes, namely the power user and the double-life user. More specifically, let $\mathcal{C}_i \in \mathcal{Y}$ be one of the macro-categories of Yelp. Let rn_i be the average number of reviews of $\mathcal{C}_i$. Let b_p be a Yelp bridge and let $CSet_p$ be the set of the macro-categories that received reviews from b_p. Then:

- b_p is defined as a *power user* if, for each macro-category $\mathcal{C}_j \in CSet_p$, the number of her reviews is greater than or equal to $2 \cdot rn_j$.
- b_p is defined as a *(x,y) access double-life user* (*access-dl-user*, for short) if both the following conditions hold:
  - for a subset $CSet_{p_x} \subset CSet_p$ of x macro-categories, the number of reviews of each $\mathcal{C}_j \in CSet_{p_x}$ is greater than or equal to $2 \cdot rn_j$;
  - for a subset $CSet_{p_y} \subset CSet_p$ of y macro-categories, such that $CSet_{p_x} \cap CSet_{p_y} = \emptyset$, the number of reviews of each $\mathcal{C}_k \in CSet_{p_y}$ is less than or equal to $\frac{1}{2} \cdot rn_k$.

Double-life users play an extremely interesting role because they are very rare. Therefore, we deepen our investigation on them and introduce a second kind of double-life users. Specifically, let b_p be a Yelp bridge. Then b_p is defined as a (x, y) *score double-life user* (*score-dl-user*, for short) if both the following conditions hold:

- for a subset $CSet_{p_x} \subset CSet_p$ of x macro-categories, the average number of stars that b_p assigned to the corresponding businesses is higher than or equal to 4;

- for a subset $CSet_{p_y} \subset CSet_p$ of y macro-categories, such that $CSet_{p_x} \cap CSet_{p_y} = \emptyset$, the average number of stars that b_p assigned to the corresponding businesses is lower than or equal to 2.

In order to make our inferential campaign on negative reviews and reviewers complete, we need to introduce a further network that we call *Negative Reviewer Network* $\overline{\mathcal{U}} = \langle \overline{N}, \overline{A} \rangle$. $\overline{N}$ is the set of nodes of $\overline{\mathcal{U}}$. There is a node $n_i \in \overline{N}$ for each Yelp user who made at least one negative review. There is an arc $a_{pq} = (n_p, n_q)$ if there exists a friendship relationship between the user u_p, corresponding to n_p, and the user u_q, corresponding to n_q.

In the next section, we show how all the concepts presented here can be exploited to prove the hypotheses formulated in Section 3. This allows us to extract knowledge about negative reviews and negative influencers in Yelp.

5 Results

5.1 General characteristics of Yelp

We collected the data necessary for the activities connected with our inferential campaign from the Yelp website at the address `https://www.yelp.com/dataset`. In order to extract information of interest from available data, we had to carry out a preliminary analysis. A first result concerns the presence of 10,289 businesses whose category did not belong to any of the Yelp macro-categories, and 482 businesses that did not have any category associated with them (recall that in Yelp a business can belong to one or more categories). Since the total number of businesses was 192,609, we decided to discard these two kinds of businesses, because the amount of data removed was insignificant while their presence would have led to procedural problems.

At this point, we analyzed the distribution of the categories among the macro-categories. We report the result obtained in Figure 1. As we can see from this figure, the macro-category "Restaurants" has a much greater number of categories than the other ones.

Figure 2 shows the average number of reviews per user for each macro-category. As we can see, the three macro-categories with the highest average number of reviews are "Restaurants", "Food" and "Nightlife". Furthermore, in Figure 3, we show the same distribution for bridges only. We can see that the three macro-categories with the highest number of reviews are always the same. However, the average number of reviews is generally higher for bridges than for normal users. Therefore, we can conclude that bridges not only tend to review businesses of different macro-categories (and this happens by definition of bridge itself) but also to do more reviews than non-bridges.

In Figure 4, we report the distribution of access-dl-users against k. From the analysis of this figure, we observe that the number of access-dl-users is already very high for $k = 2$ and further increases for $k = 3$; then, it decreases very quickly and becomes almost negligible for $k > 4$.

We start looking at the access-dl-users corresponding to the simplest case of bridges, namely 2-bridges. Table 1 shows the total number of 2-bridges, the number of (1,1) access-dl-users and the number of power users, together with their corresponding percentage of the overall number of 2-bridges. This table shows that (1,1) access-dl-users and power users represent very small fractions of the overall set of 2-bridges.

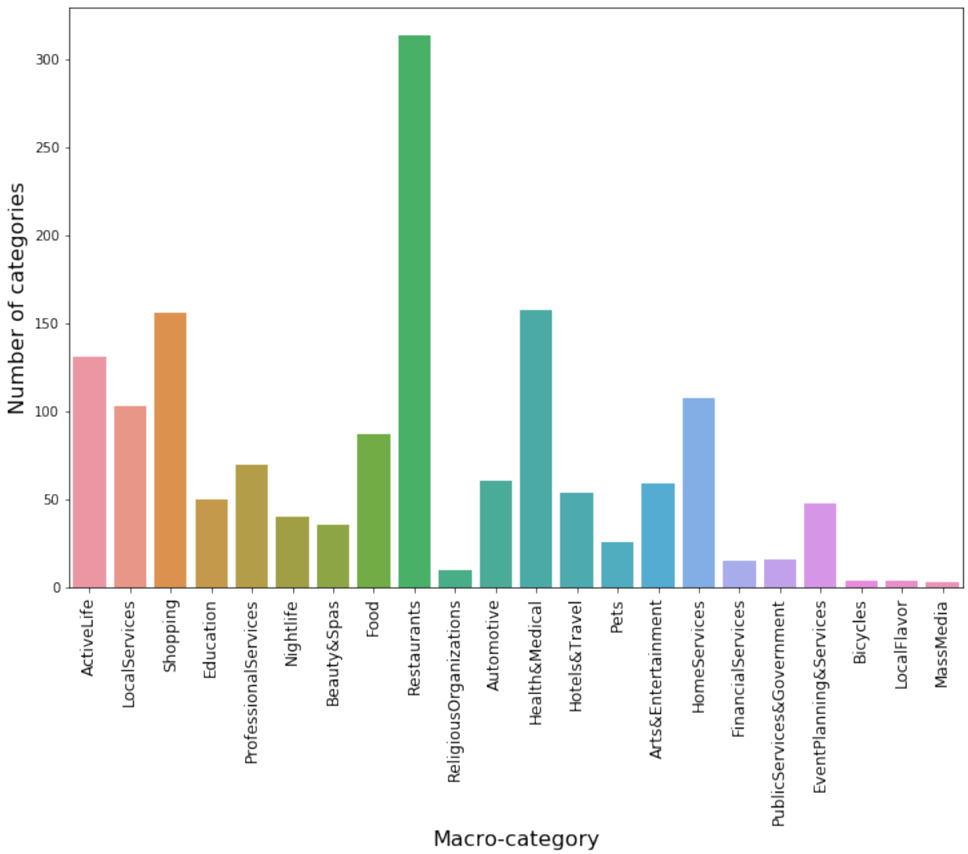

Figure 1: Distribution of the categories inside the Yelp macro-categories

Type of users	*Number and percentage*
2-bridges	427130 (100%)
(1,1) access-dl-users	745 (0.17%)
power users	375 (0.087%)

Table 1: Numbers and percentages of 2-bridges, access-dl-users and power users in Yelp

We continue by examining all the k-bridges as k grows, until at least one of them is an access-dl-user or a power user. We can observe that this condition occurs for $k \leq 6$. The corresponding numbers and percentages are shown in Tables 2 - 5. From the analysis of these tables, we can see how the number of k-bridges decreases as k increases, but the decrease is not fast. On the other hand, the number of access-dl-users decreases very rapidly, about one order of magnitude at each step. The number of power users decreases more slowly.

Type of users	*Number and percentage*
3-bridges	245123 (100%)
(1,2) access-dl-users	450 (0.18%)
(2,1) access-dl-users	374 (0.15%)
power users	200 (0.081%)

Table 2: Numbers and percentages of 3-bridges, access-dl-users and power users in Yelp

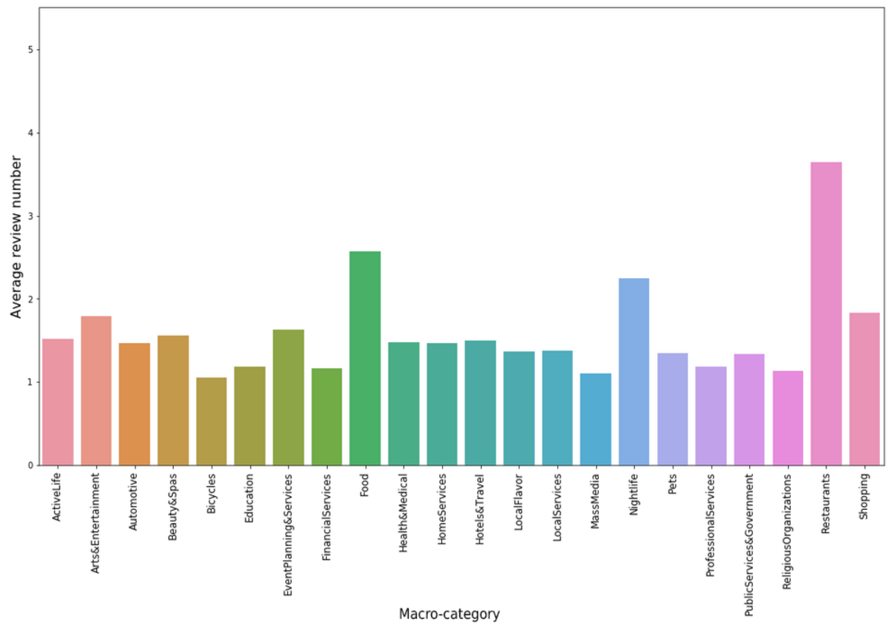

Figure 2: Average number of business reviews made by Yelp *users* for each macro-category

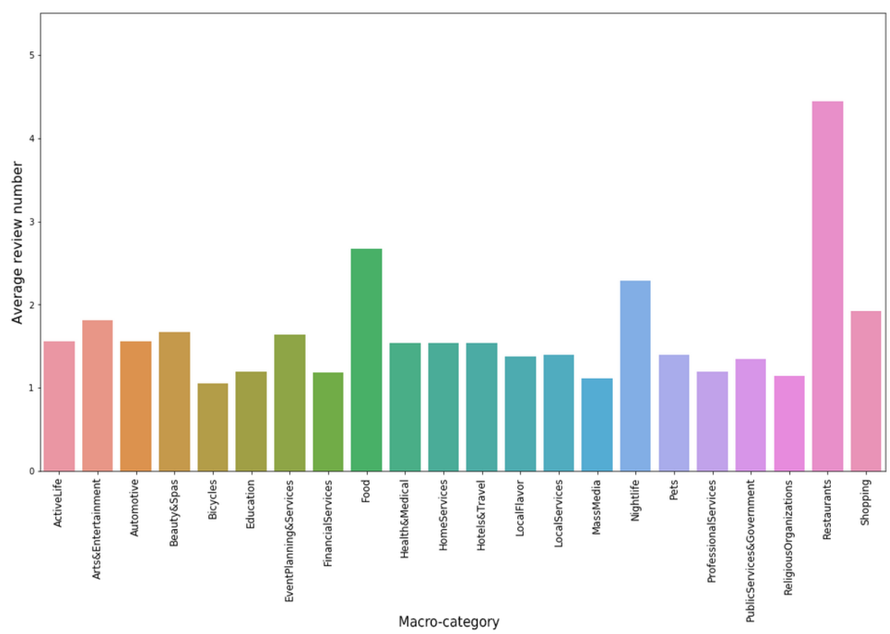

Figure 3: Average number of business reviews made by Yelp *bridges* for each macro-category

5.2 Investigating the correctness of the Hypothesis H1

In Section 3, we have seen that a user can assign a number of stars between 1 and 5 to a business in Yelp. The higher the number of stars, the better her rating is. Therefore, we decided to study the

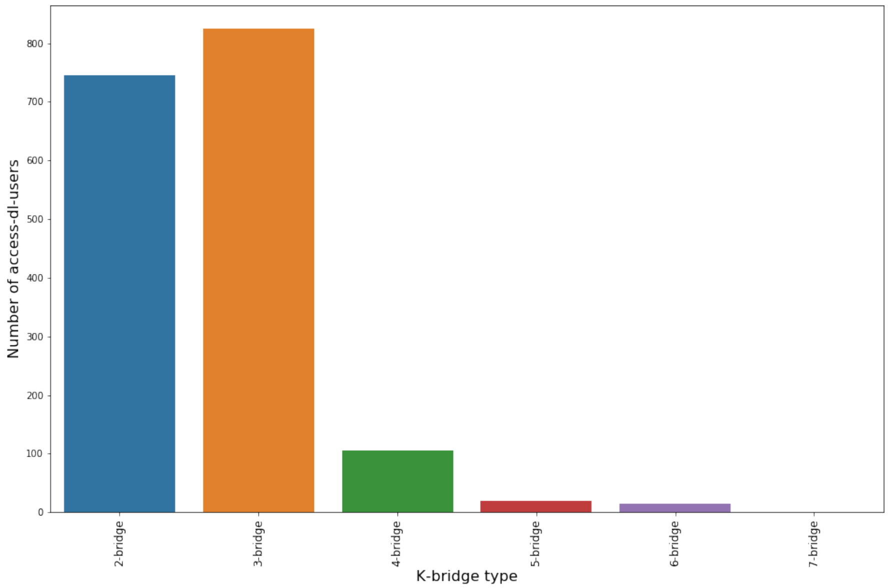

Figure 4: Distribution of access-dl-users against k

Type of users	*Number and percentage*
4-bridges	147101 (100%)
(1,3) access-dl-users	19 (0.013%)
(2,2) access-dl-users	59 (0.040%)
(3,1) access-dl-users	28 (0.019%)
power users	35 (0.023%)

Table 3: Numbers and percentages of 4-bridges, access-dl-users and power users in Yelp

Type of users	*Number and percentage*
5-bridges	91680 (100%)
(1,4) access-dl-users	6 (0.007%)
(2,3) access-dl-users	11 (0.012 %)
(3,2) access-dl-users	3 (0.003%)
(4,1) access-dl-users	0 (0%)
power users	14 (0.015%)

Table 4: Numbers and percentages of 5-bridges, access-dl-users and power users in Yelp

Type of users	*Number and percentage*
6-bridges	63708 (100%)
(1,5) access-dl-users	0 (0%)
(2,4) access-dl-users	0 (0%)
(3,3) access-dl-users	1 (0.002%)
(4,2) access-dl-users	2 (0.003%)
(5,1) access-dl-users	11 (0.017%)
power users	11 (0.017%)

Table 5: Numbers and percentages of 6-bridges, access-dl-users and power users in Yelp

reviews of users focusing on the number of stars that they assigned to businesses.

Figure 5 shows the average number of stars that users assigned to the businesses of each macro-category. As we can see from this figure, this number is very high as it is always greater than 3. As previously pointed out, this is actually not very surprising because the mechanism based on stars follows a Likert scale and, in literature, it is well known that this scale is generally positively biased

(Alexandrov, 2010; Peeters and Czapinski, 1990; Bertram, 2007).

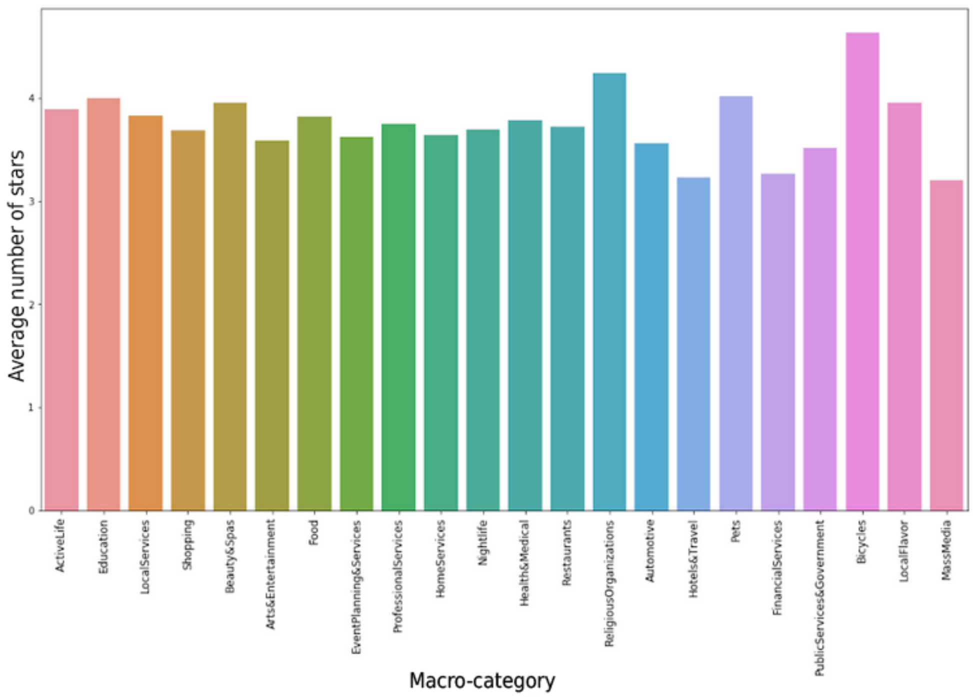

Figure 5: Average number of stars for each macro-category of Yelp

In Table 6, we report the mean, standard deviation and mode of the number of stars assigned by bridges and non-bridges to all businesses. As we can see from this table, there is no substantial difference in this type of behavior between bridges and non-bridges.

Statistical Parameter	*Bridges*	*Non-bridges*
Mean	3.73	3.57
Standard Deviation	1.44	1.72
Mode	5	5

Table 6: Values of mean, standard deviation and mode of the number of stars assigned by bridges and non-bridges to all businesses

From the results of Table 6, it is clear that it makes no sense to talk about power users in the star-based analysis, because almost all users have the same behavior and assign a high number of stars to almost all businesses. All these tests allow us to define the following:

Implication 1: The star-based review system of Yelp is positively biased. Indeed, almost all users assign a high number of stars to almost all businesses.

Implication 1 is clearly a confirmation of the correctness of the Hypothesis H1.

5.3 Investigating the correctness of the Hypothesis H2

In Figure 6, we report the distribution of score-dl-users against k. From the analysis of this figure we note that it follows a power law. If we compare this figure with Figure 4, we observe that for $k = 2$, the number of score-dl-users is much smaller than the one of access-dl-users. However, the decrease of the number of score-dl-users when k increases is much smaller because they are different from 0 until to $k = 14$.

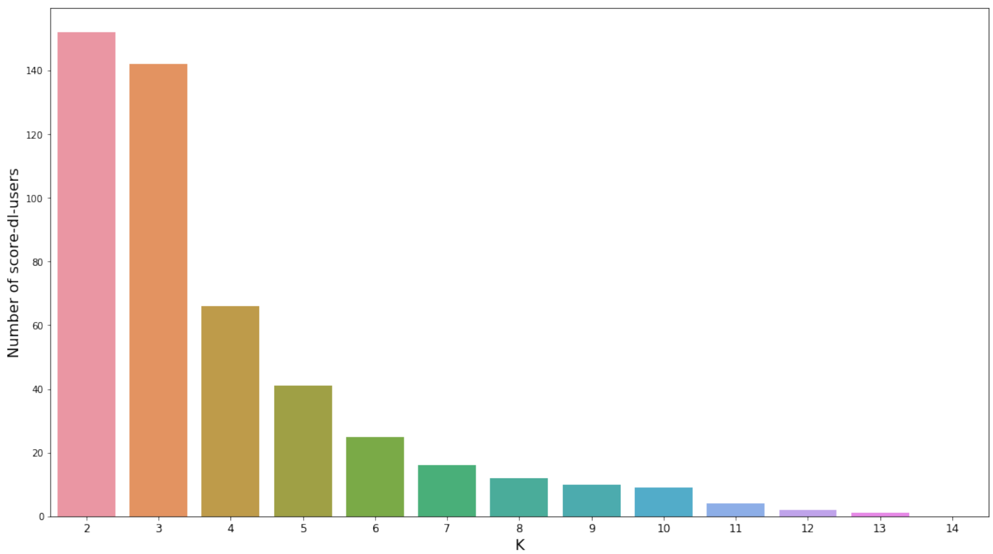

Figure 6: Distribution of score-dl-users against k

We continued our analysis by verifying whether score-dl-users and access-dl-users were the same people or not. We carried out this analysis with $k = 6$, because we had no access-dl-users with higher values of k. In this case, we could see that the intersection of the two sets was empty.

To better understand the main features of score-dl-users we considered those corresponding to 7-bridges. These users were 16 (see Figure 6), a number that allowed us to examine in detail each review carried out by them. During this analysis we found several interesting knowledge patterns. More specifically, we observed that (1,6) and (6,1) score-dl-users show a completely different behavior from the other 7-bridges. In fact, in this case, each (1,6) score-dl-user assigned positive scores to all the business of the only macro-category that she positively reviewed. Similarly, each (6,1) score-dl-user assigned negative values to all the businesses of the only macro-category that she negatively reviewed. This can be justified thinking that users have a strong interest in that macro-category and so they developed more accurate and stable evaluation criteria for the businesses belonging to it.

As for the other 7-bridges, we found that (2,5), (3,4), (4,3) and (5,2) score-dl-users show a less extreme behavior, in the sense that they do not tend to give always positive or always negative ratings to all the businesses of a given macro-category.

We then repeated the previous analyses for the last category of access-dl-users that we had available, namely the 6-bridges, to verify if the particular behavior of score-dl-users was typical of this kind of double-life user or if it was something common. Actually, 6-bridge access-dl-users were 13;

therefore, we were able to make a detailed analysis of each review performed by each user also in this case. We examined (1,5), (2,4), (3,3), (4,2) and (5,1) access-dl-users and we did not find substantial differences in the behavior of these five categories of users. This appeared as a confirmation of the singularity of the behavior observed for (1,6) and (6,1) score-dl-users. The previous analyses suggest the following:

> Implication 2: *(a)* Score-dl-users play a key role in negative reviews. *(b)* They are very keen on negatively judging the macro-category they mostly attend.

Implication *2(a)* confirms the correctness of our Hypothesis H2. But there is much more. In fact, Implication *2(b)* was an unexpected result that prompted us to carry out a further experiment to have a confirmation. In it, we considered k-bridges, with $3 \leq k \leq 8$, and computed the percentage of them who negatively reviewed the macro-category of businesses they attended the most. Afterwards, we computed the same percentage taking into account only k-bridges that were score-dl-users. The results obtained are shown in Table 7. They represent an extremely strong confirmation of the previous qualitative analysis.

k	*Percentage of k-bridges*	*Percentage of score-dl-users k-bridges*
3	4.35%	91.5%
4	4.03%	79%
5	3.65%	61%
6	2.40%	63%
7	2.11%	56%
8	1.55%	33%

Table 7: Percentages of k-bridges and score-dl-users k-bridges who negatively reviewed the macro-category they mostly attended

As we have seen, the definition and behavior of score-dl-users are based on the number of stars assigned by a user to a business during a review. We have already said that this type of score is based on a Likert scale and, therefore, it is positively biased (Alexandrov, 2010; Peeters and Czapinski, 1990; Bertram, 2007). In order to overcome this problem, in the literature authors suggest to evaluate the text of the reviews and to make a sentiment analysis on it (Kaviya et al., 2017; Kasper and Vela, 2011). We carried out this activity using two well-known sentiment analysis tools. The first is TextBlob[3], which, given a text, specifies if the corresponding polarity is positive, negative or neutral. We applied TextBlob to users' review texts. The results obtained are reported in Table 8. From the analysis of this table we can see that the difference between the score based on stars and the polarity based on sentiment analysis is equal to 15%.

The second sentiment analysis tool we considered is Vader (Hutto and Gilbert, 2014). Also in this case, we applied it to the users' review texts. The results obtained are shown in Table 9. The analysis of this table confirms the very low difference between the score of the star-based reviews and the polarity of the review texts (in fact, in this case, this difference is equal to 14%).

This allows us to conclude that score-based evaluations are generally confirmed by the sentiment analysis performed on the corresponding reviews.

[3] `https://textblob.readthedocs.io`

Parameters	*Value obtained by applying TextBlob*
Reviews	6,685,902
Reviews with a number of stars less than or equal to 2 (negative reviews)	1,544,553
Reviews classified as negative by TextBlob	847,359
Reviews with a number of stars greater than or equal to 3 (positive reviews)	5,141,347
Reviews classified as positive by TextBlob	5,781,007
Reviews classified as neutral by TextBlob	57,536
Negative reviews classified as positive	823,414
Positive reviews classified as negative	154,176
Positive reviews classified as neutral	30,914
Negative reviews classified as neutral	26,620

Table 8: Comparison between the review score based on stars and the review polarity obtained by applying TextBlob

Parameter	*Value obtained by applying Vader*
Reviews	6,685,902
Reviews with a number of stars less than or equal to 2 (negative reviews)	1,544,553
Reviews classified as negative by Vader	982,102
Reviews with a number of stars greater than or equal to 3 (positive reviews)	5,141,347
Reviews classified as positive by Vader	5,649,489
Reviews classified as neutral by Vader	54,311
Negative reviews classified as positive	724,241
Positive reviews classified as negative	184,557
Positive reviews classified as neutral	31,542
Negative reviews classified as neutral	22,767

Table 9: Comparison between the review score based on stars and the review polarity obtained by applying Vader

5.4 Investigating the correctness of the Hypothesis H3

At this point, we analyzed how users influence each other with regard to negative reviews. We took into consideration the network of friendships $\mathcal{Y}^f$ since it is easier for a user to have characteristics more similar to her friends than to people she does not know, due to the principle of homophily (McPherson et al., 2001). Therefore, the ability to influence someone and/or to be influenced by her is presumably greater with friends than with others.

As a first analysis, for each macro-category, we considered the percentage of users such that they, and at least one of their friends, reviewed the same business negatively. The results obtained are shown in Figure 7. From the analysis of this figure we can see how the percentages are extremely low. The macro-category with the highest percentage is “Restaurant”, followed by “Nightlife” and “Food”. This result can be explained taking into account that a person often attends restaurants or nightclubs with her friends. Therefore, it is not unlikely that her negative judgement of a business may lead to (or, on the contrary, may be caused by) a negative judgement of one or more of her friends.

We repeated the analysis by distinguishing bridges from non-bridges. The corresponding results are shown in Figures 8 and 9. From the analysis of these figures we observe higher values for bridges than for non-bridges. For example, the value of “Nightlife” for bridges is more than 4 times the value for non-bridges. Similarly, “Food”, in case of bridges, has a percentage more than 7 times higher than for non-bridges.

To prove the statistical significance of our results we adopted a null model to compare our findings with those obtained in an unbiasedly random scenario. Specifically, we built our null model by shuffling the negative reviews among users in our dataset. In this way, we left unaltered all the original features with the exception of the distribution of negative reviews, which became unbiasedly random in the

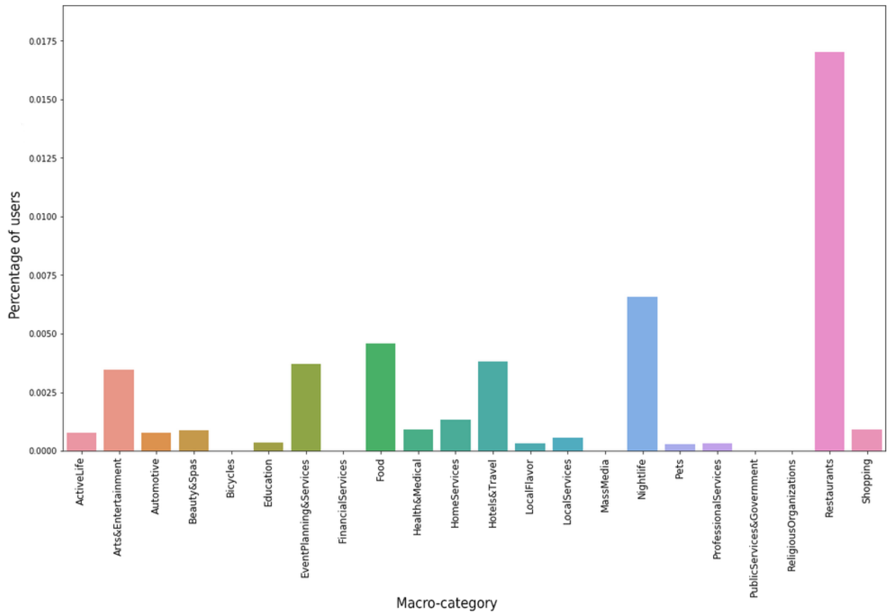

Figure 7: Percentages of *users* such that they, and at least one of their friends, reviewed the same business negatively

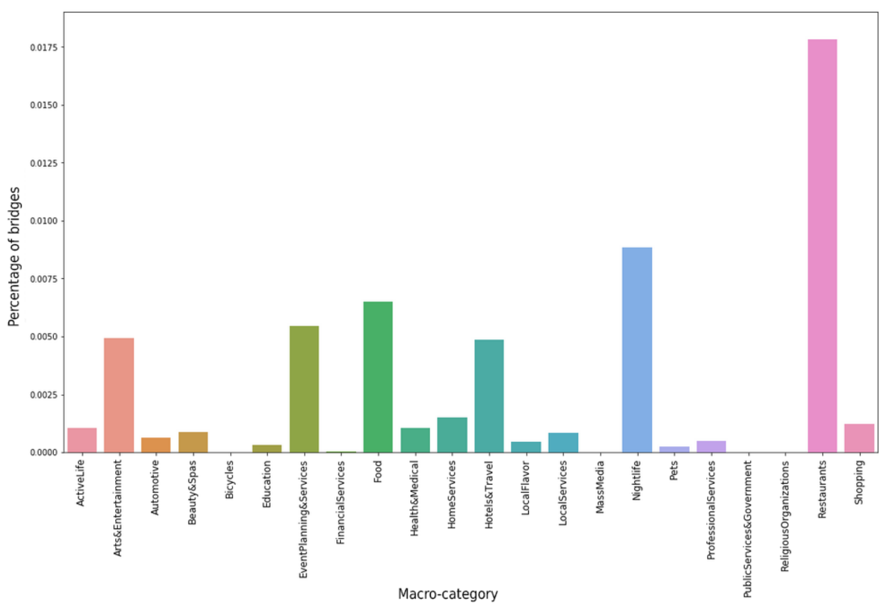

Figure 8: Percentages of *bridges* such that they, and at least one of their friends, reviewed the same business negatively

null model. After that, we repeated our analysis on the null model. The results obtained are reported in Figure 10. Comparing this figure with Figure 7, we can see that there is a certain similarity in the distributions; indeed, many of the macro-categories that had the highest values in Figure 7 continue

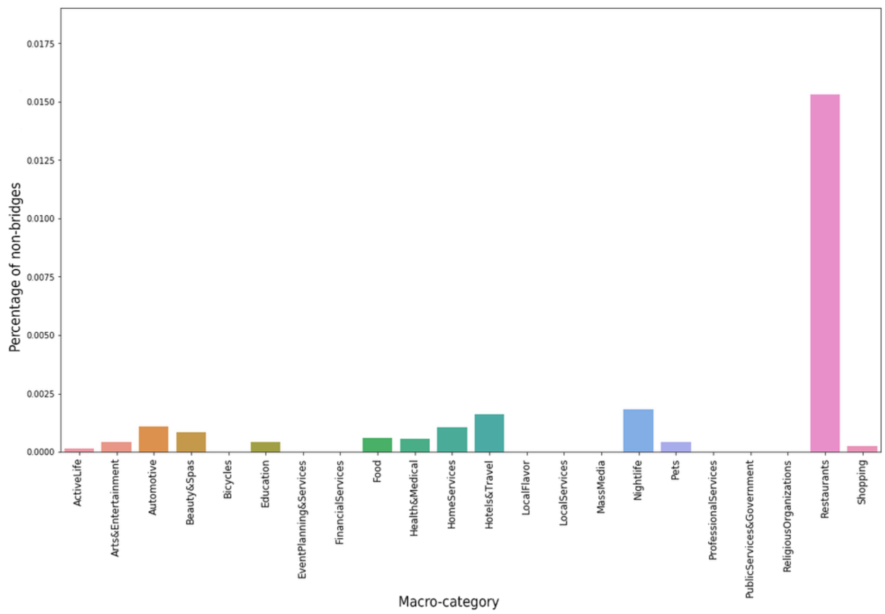

Figure 9: Percentages of *non-bridges* such that they, and at least one of their friends, reviewed the same business negatively

to have the highest values in Figure 10. However, in this last case, the values of the percentages are several orders of magnitude smaller. Therefore, we can conclude that the behavior observed in Figure 7 is not random but it is the result of the reference context.

At this point, for each macro-category, for each user who reviewed a given business negatively, we computed the percentage of her friends who, having reviewed the same business, made a negative review. The results obtained are shown in Figure 11. As we can see from this figure, the percentage values are very high for almost all macro-categories.

Figures 12 and 13 show the same distributions, but for bridges and non-bridges. From the analysis of these figures, it can be observed that the phenomenon is always strong, regardless of whether or not a user is a bridge. An interesting knowledge pattern to observe is that there is a strong polarization on the macro-categories especially in the case of non-bridges. In fact, the percentages of friends influenced by them are either above 90% or null.

All the results shown above allow us to deduce the following:

Implication 3: A user has a very high influence on her/his friends when doing negative reviews.

This implication represents a confirmation of the correctness of our Hypothesis H3.

5.5 Investigating the correctness of the Hypothesis H4

In order to evaluate the Hypothesis H4, we started with the computation of the average percentage of users who, having made a negative review in a category, have at least X‰ of their friends who negatively reviewed a business in the same category. The values of X that we considered are 1, 2, 3,

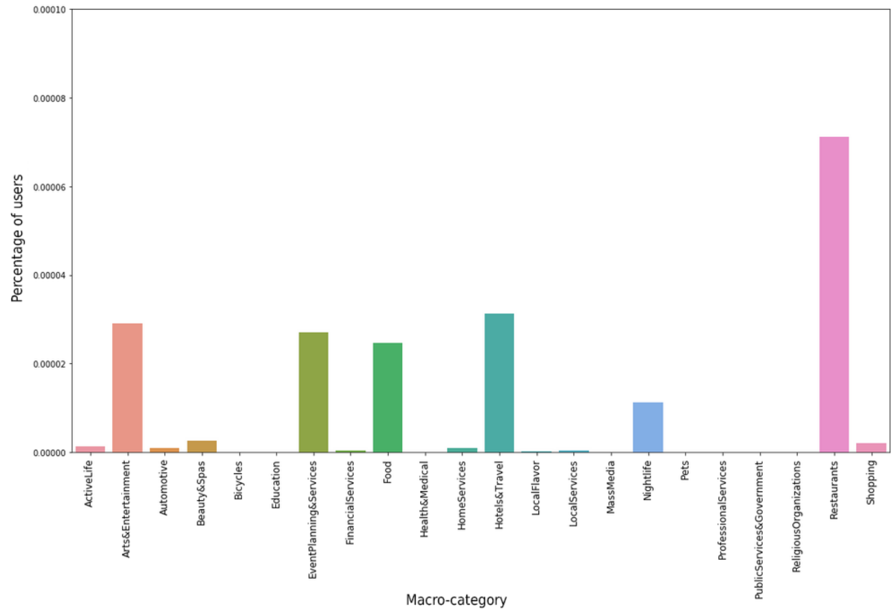

Figure 10: Percentages of *users* in the null model such that they, and at least one of their friends, reviewed the same business negatively

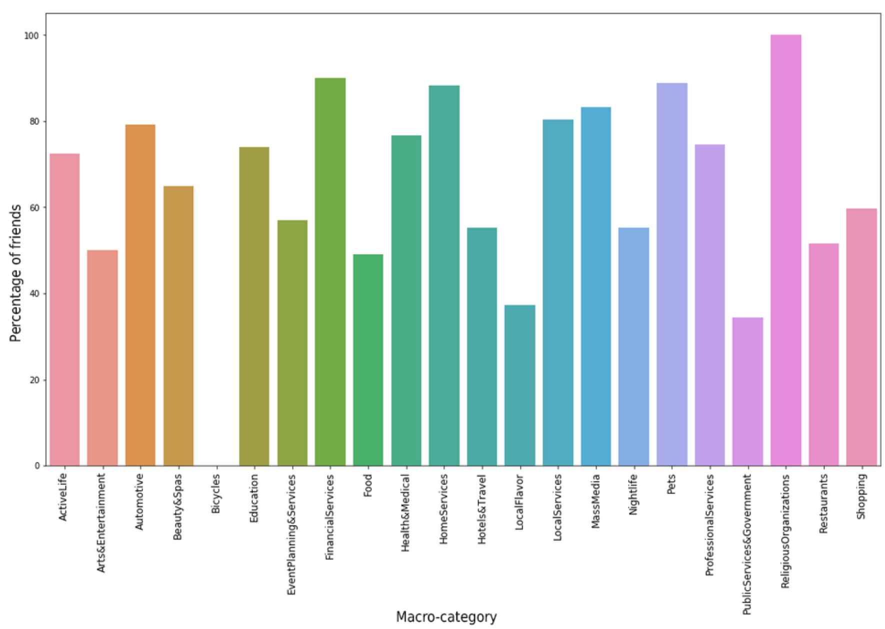

Figure 11: Percentages of friends who, having reviewed the same business as a *user* who reviewed a business negatively, also provided a negative review

5, 10 and 100. As an example, in Figure 14, we report the results obtained in the case of $X = 5$. As we can see from this figure, the percentages are some orders of magnitude greater than the ones of

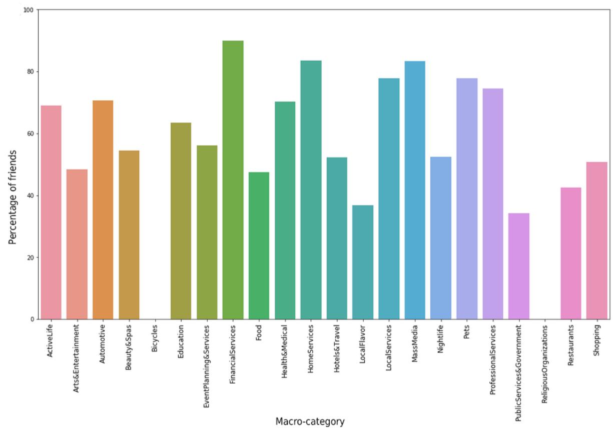

Figure 12: Percentages of friends who, having reviewed the same business as a *bridge* who reviewed a business negatively, also provide a negative review

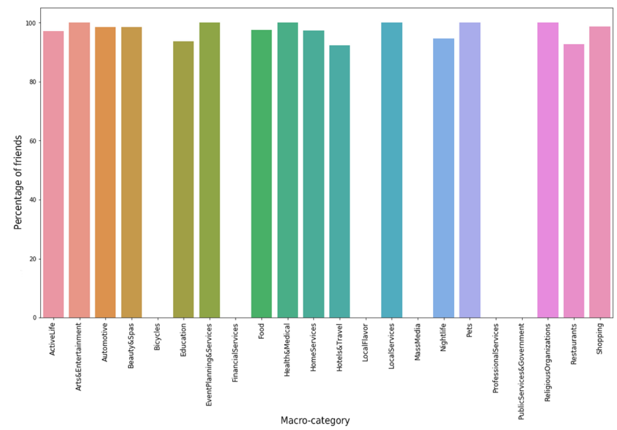

Figure 13: Percentages of friends who, having reviewed the same business as a *non-bridge* who reviewed a business negatively, also provide a negative review

Figure 10. The macro-categories with the highest values are the same as before, i.e., "Restaurants", "Food" and "Nightlife".

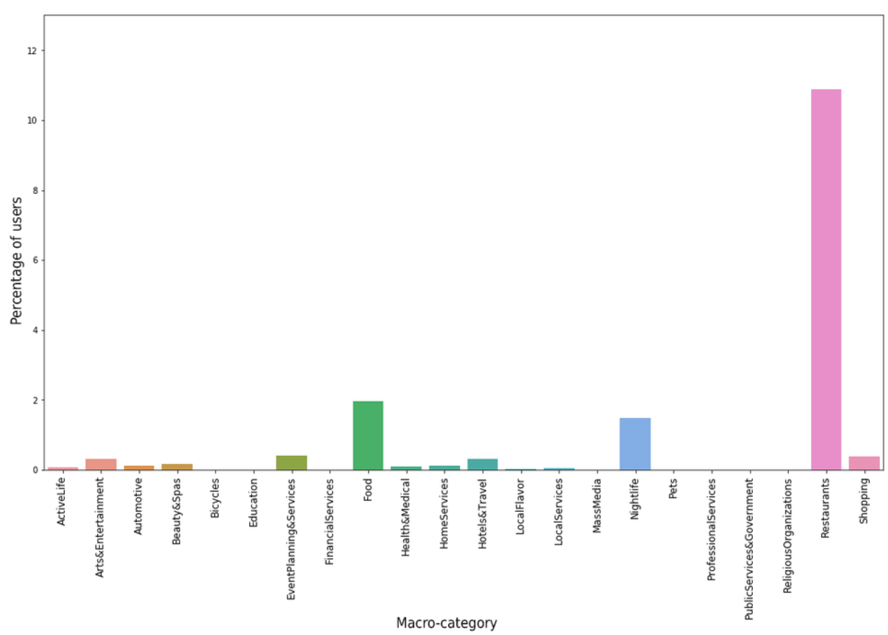

Figure 14: Average percentages of *users* who, having made a negative review in a macro-category, have at least $X‰$ of their friends who reviewed a business in the same macro-category negatively

As in the previous case, we distinguished bridges from non-bridges. The results of the corresponding analysis are shown in Figures 15 and 16. These figures, along with the previous ones involving bridges and non bridges, allow us to define the following:

Implication 4: Bridges have a much greater power of influence than non-bridges.

Again, we made the comparison with the null model. The results obtained for $X = 5$ are reported in Figures 17, 18 and 19. From the examination of these figures, we can see how results obtained are not random but they are intrinsic to Yelp. Note that the non-randomness can be observed for *bridges* but generally not for *non-bridges*; this is important because it allows us to conclude that this property characterizes bridges against non-bridges.

Implication 4 represents a confirmation that our Hypothesis H4 was correct.

5.6 Investigating the correctness of the Hypothesis H5 and defining a profile of negative influencers in Yelp

To investigate the correctness of the Hypothesis H5 we considered the *Negative Reviewer Network* $\overline{\mathcal{U}} = \langle \overline{N}, \overline{A} \rangle$ introduced in Section 4.

The analysis of this network allowed us to focus on users who reviewed some businesses negatively, because, as we saw in the previous analysis, they are uncommon. Firstly, we computed the number of nodes, the number of edges, the clustering coefficient and the density of $\overline{\mathcal{U}}$ and we compared them with the same parameters as $\mathcal{U}$. Results are shown in Table 10.

	$\mathcal{U}$	$\overline{\mathcal{U}}$
Number of nodes	1637138	743178
Number of edges	7392305	2199987
Average clustering coefficient	0.043	0.039
Density	0.00000551619	0.00000796645

Table 10: Characteristics of $\mathcal{U}$ and $\overline{\mathcal{U}}$

From the analysis of this table we can observe that the number of users who made at least one negative review is 45.39% of total users. As for the average clustering coefficient and the density, we found that their values do not present significant differences between $\mathcal{U}$ and $\overline{\mathcal{U}}$.

At this point, we computed the distribution of users for $\overline{\mathcal{U}}$; it is shown in Figure 20. As we can see from this figure, it follows a power law.

After studying the basic parameters of $\overline{\mathcal{U}}$, we computed the degree centrality of the nodes of this network. In particular, we focused on the users with the highest values of degree centrality. More specifically, we considered the top $X\%$ users, $X \in \{1, 5, 10, 20\}$. Observe that as X decreases, the corresponding top users are increasingly central, i.e., increasingly strong. In Figure 21, we show the distributions against k for the top $X\%$ of users with the highest degree centrality. Note that for $X = 20$, the distribution follows a power law, even if it is flatter than the one of Figure 20, which referred to all users. As X decreases, we can see how the distribution becomes flatter and flatter, moving to the right and tending to a Gaussian shape. This allows us to conclude that more central users (i.e., those with the highest degree centrality) tend to be stronger also as k-bridges (i.e., characterized by an increasingly higher value of k).

Instead, in Figure 22, we show the user distributions against k for the top $X\%$ of users with the highest eigenvector centrality. The trend of these distributions as X decreases is very similar to (although slightly less marked than) the one of the degree centrality.

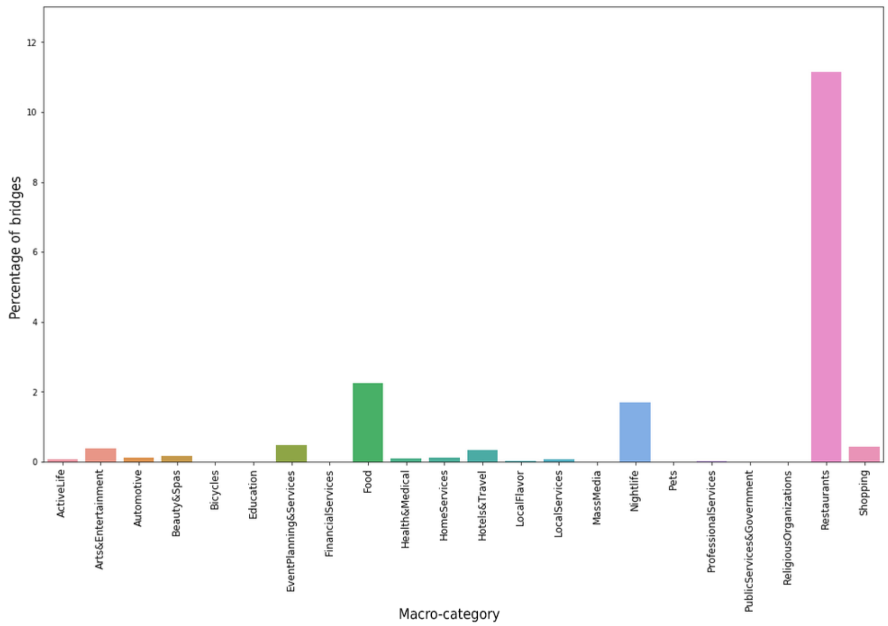

Figure 15: Average percentages of *bridges* who, having made a negative review in a macro-category, have at least $X‰$ of their friends who reviewed a business in the same macro-category negatively

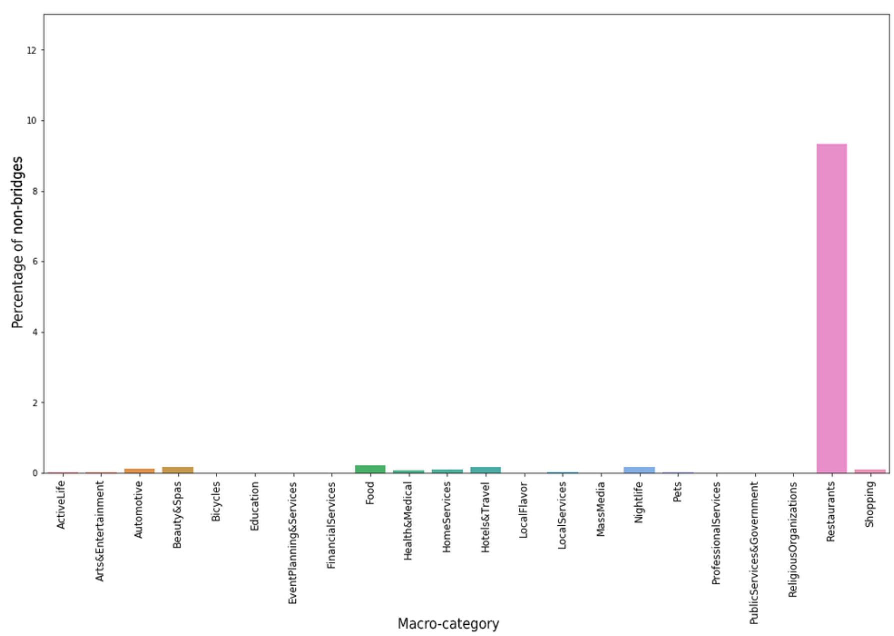

Figure 16: Average percentages of *non-bridges* who, having made a negative review in a macro-category, have at least $X‰$ of their friends who reviewed a business in the same macro-category negatively

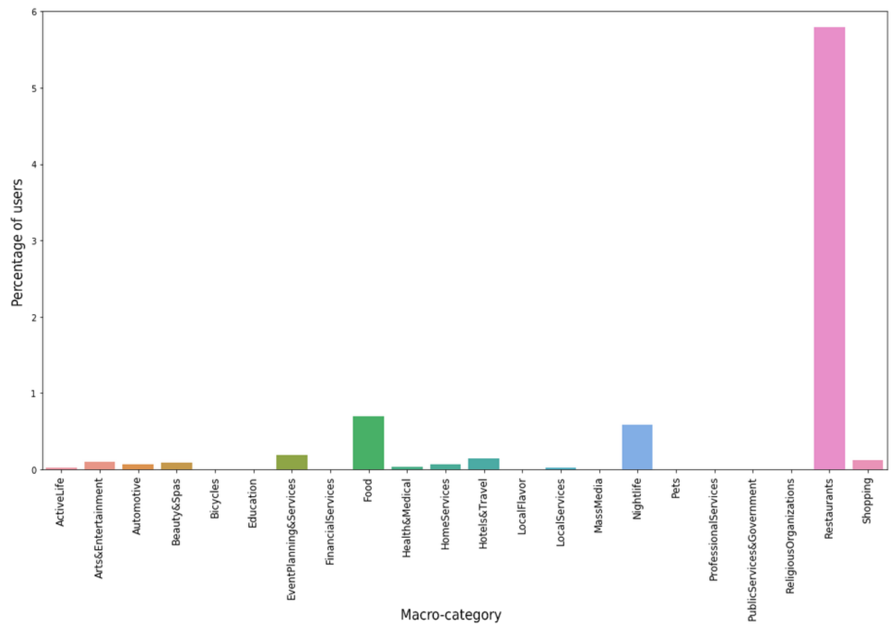

Figure 17: Average percentages of *users* in the null model who, having made a negative review in a macro-category, have at least X‰ of their friends who reviewed a business in the same macro-category negatively

Figure 23 shows the user distributions against k for the top $X\%$ of users with the highest PageRank. Also in this case, we have a similar trend, although the variations of the distributions as X decreases are much more attenuated, compared to the two previous cases. The last three figures allow us to define the following:

Implication 5: There is a correlation between k-bridges and top central users.

Implication 5 is valid especially for the top central users based on degree centrality. This result, along with the previous ones, is extremely important because it allows us to determine which are the main negative influencers in Yelp. In fact, we can define the following:

Implication 6: The main negative influencers in Yelp are score-dl-users who simultaneously are top central users (according to degree and/or eigenvector and/or PageRank centrality measures).

Implication 6 not only confirms the correctness of the Hypothesis H5, but goes much further. In fact, it defines a profile of the negative influencers in Yelp and, consequently, provides a way to detect them.

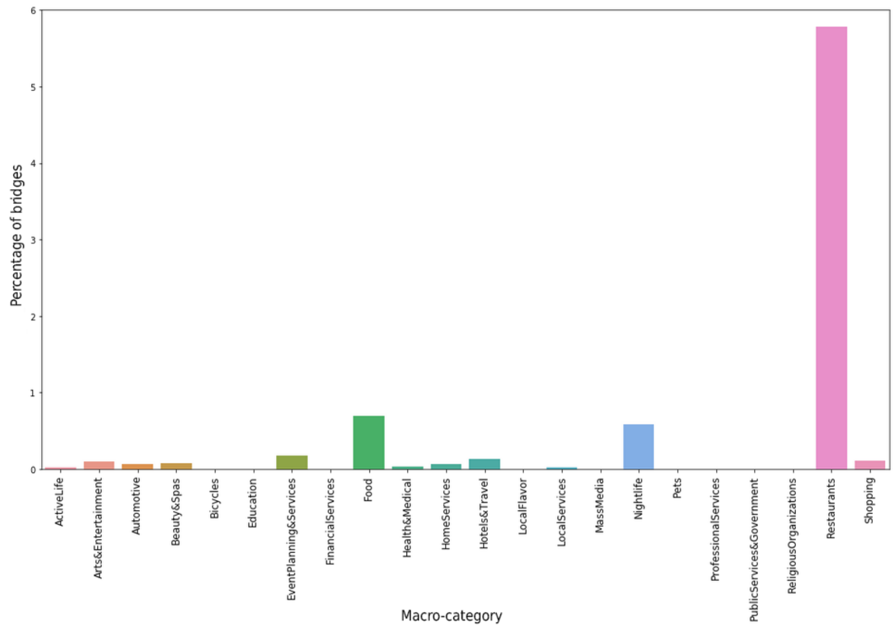

Figure 18: Average percentages of *bridges* in the null model who, having made a negative review in a macro-category, have at least X‰ of their friends who reviewed a business in the same macro-category negatively

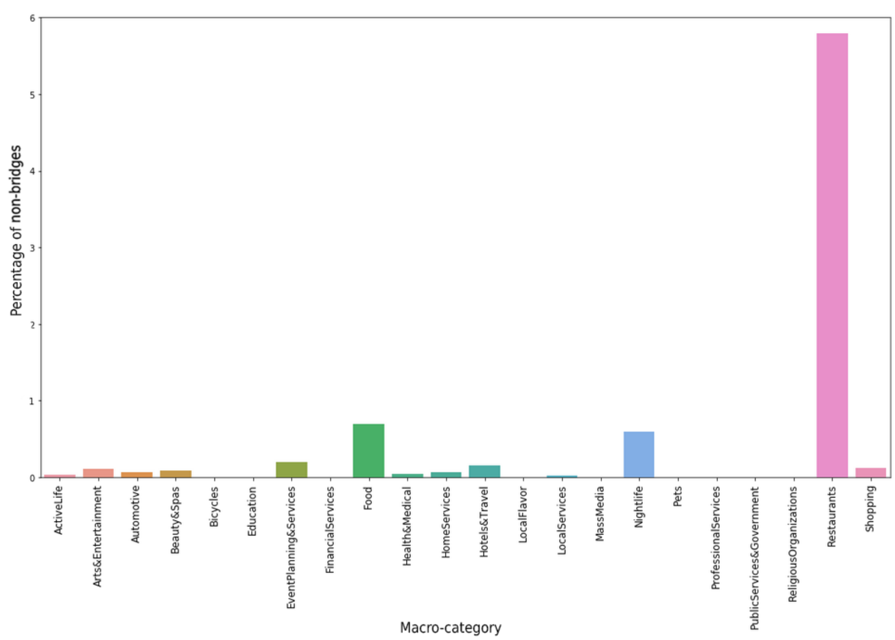

Figure 19: Average percentages of *non-bridges* in the null model who, having made a negative review in a macro-category, have at least X‰ of their friends who reviewed a business in the same macro-category negatively

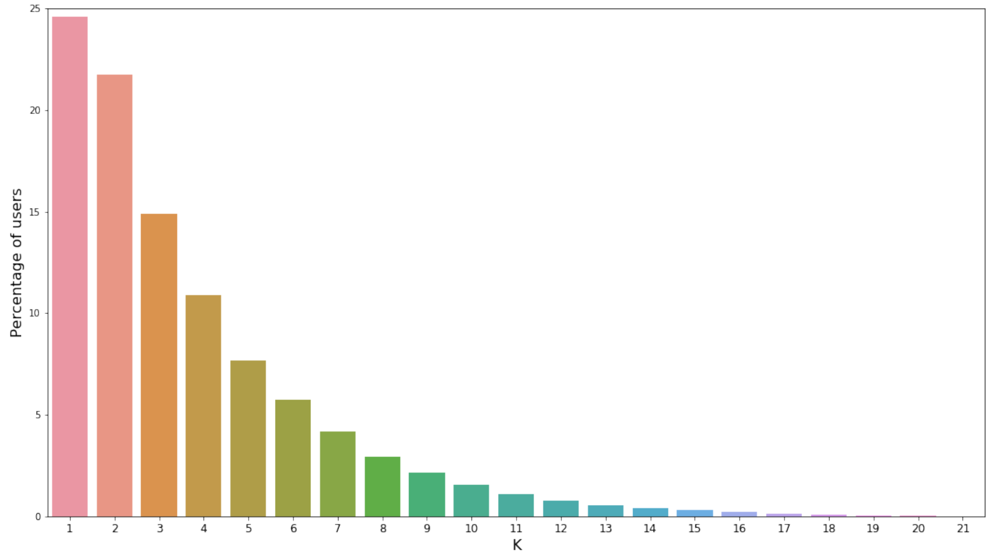

Figure 20: Distribution of users of $\overline{\mathcal{U}}$ against k

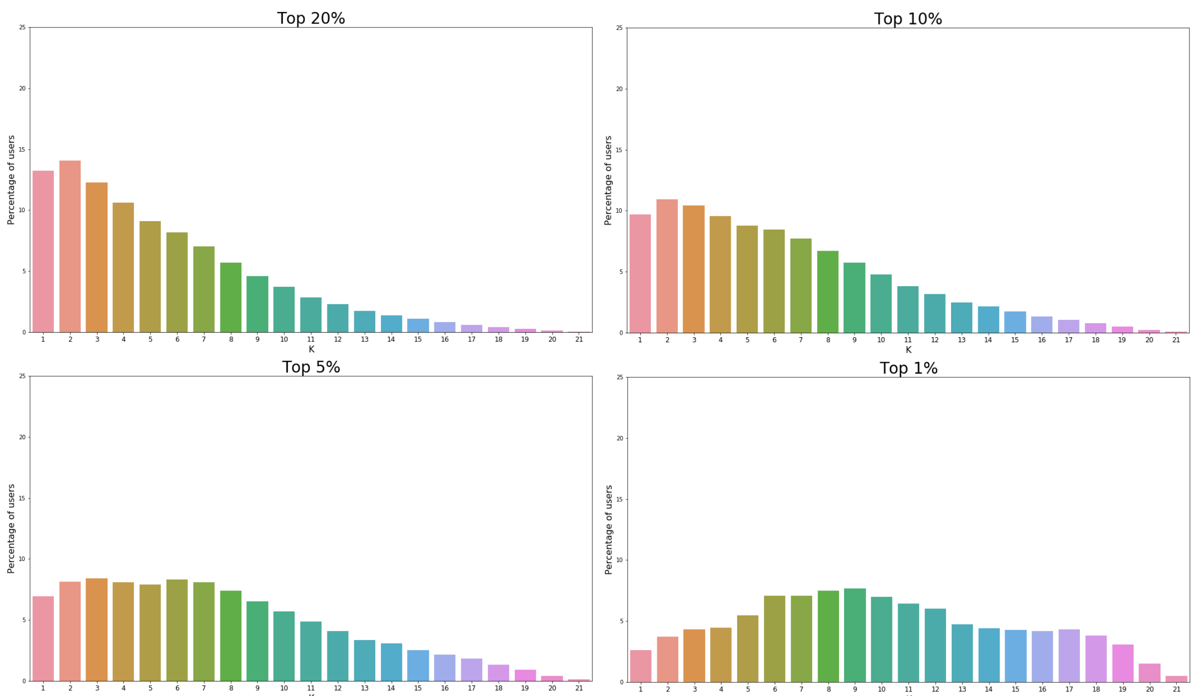

Figure 21: Distributions of the top $X\%$ of users with the highest degree centrality against k

6 Discussion

6.1 Reference context of our paper

In the previous sections, we have investigated the phenomenon of negative reviews in Yelp and, then, we have characterized negative influencers in this social medium. In the past, different research

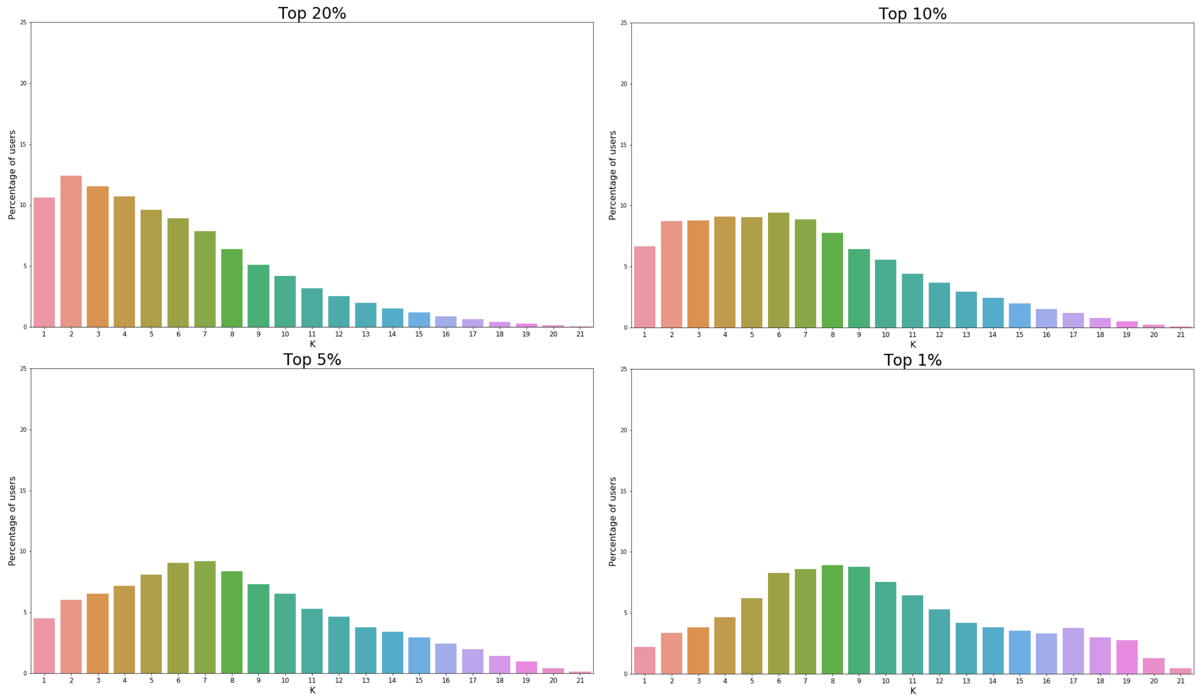

Figure 22: Distributions of the top $X\%$ of users with the highest eigenvector centrality against k

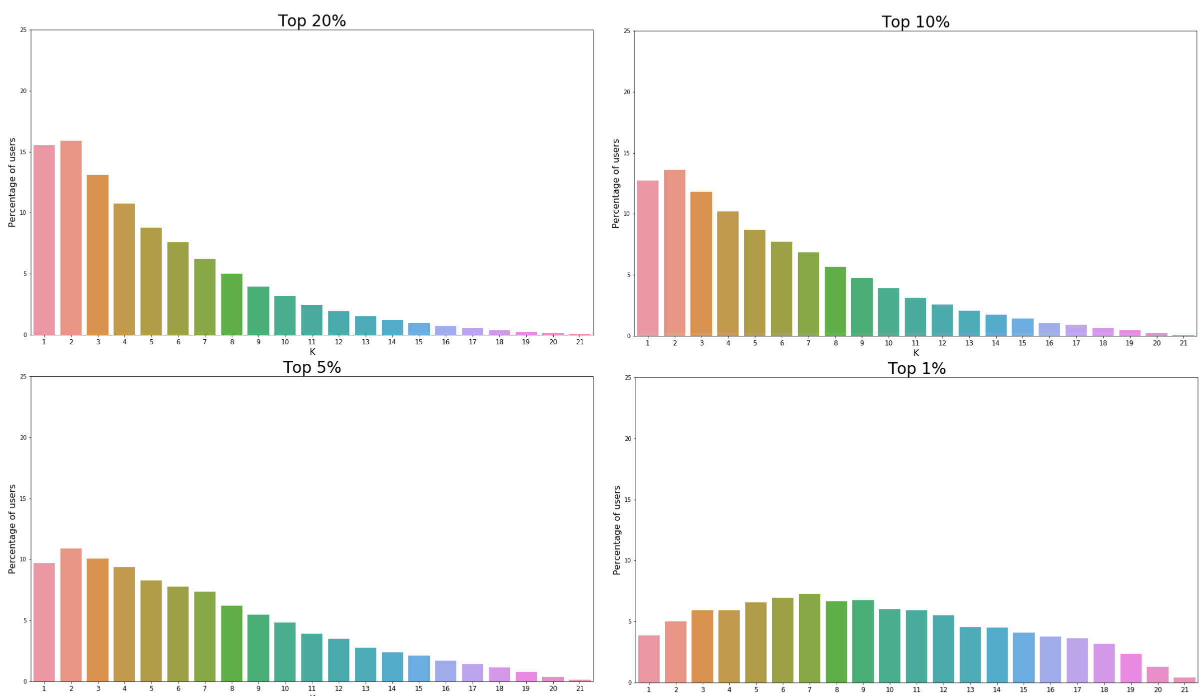

Figure 23: Distributions of the top $X\%$ of users with the highest PageRank against k

papers have focused on the consequences that user-written reviews have on businesses and, generally, on the market. As a first step in this scenario, it is interesting to understand what makes customer reviews helpful to a consumer in her process of making a purchase decision. With regard to this, in (Schuff and Mudambi, 2012), the authors first collect reviews made on `Amazon.com`. Then, they distinguish between two different product types, namely: *(i)* search goods, for which a consumer can obtain information on their quality before purchasing them; *(ii)* experience goods, which are products requiring a purchase before evaluating their quality. This product categorization plays a key role in understanding what a consumer perceives more from a review. Indeed, moderate reviews are more helpful than extreme (i.e., strongly positive or negative) ones for experience goods, but not for search goods. Furthermore, longer reviews are generally perceived as more helpful than shorter ones, but this effect is greater for search goods than for experience goods.

Another interesting contribution in this scenario is reported in (Zhang et al., 2014), in which the authors introduce several factors that can influence the decision making process of consumers about their purchases. Indeed, the authors of (Zhang et al., 2014) strive to understand what are the key elements guiding a user in the purchase of a certain product. They propose a model taking systematic factors (e.g., the quality of online reviews) and heuristic ones (e.g., the quantity of online reviews) into account. They test this model on 191 users and obtain interesting results. In fact, they identify important factors to care about; these are argument quality, source credibility, and perceived quantity of reviews. They empirically prove that consumers receiving reviews from credible sources and perceiving the quantity of reviews as large tend to perceive the topics in online reviews as more informative and persuasive. This means that if consumers find review sources to be credible, their purchase intention is usually higher. Finally, they also show that consumers are more likely to purchase products with many online reviews rather than with few ones.

Several authors have investigated the impact of positive and negative reviews. For instance, the authors of (Cheung and Lee, 2012) examine how a positive Electronic Word of Mouth (hereafter, eWOM) can affect other users' purchasing decisions. Indeed, eWOM is strictly related to the online reviews phenomenon, which can be regarded as a special case of it. Generally, eWOM is based on an analysis of costs and benefits. The authors investigate the psychological motivations beneath the spread of positive reviews. They take a sample dataset from the `OpenRice.com` platform, one of the most successful review platforms in Hong Kong and Macau. Through a questionnaire, they asked people who wrote reviews on this website their motivations. Starting from the received answers, they build a model based on different features, namely the eWOM intention of consumers, the reputation, the reciprocity, the sense of belonging, the pleasure to help, the moral obligation and the self-efficacy of knowledge. They show that their model is capable of representing the behavior of users when they share (positive) personal experiences on such online platforms.

The influence of positive reviews of businesses has been studied from many other points of view. For example, in (Knoll and Matthes, 2017), the authors analyze celebrity sponsorships in the context of for-profit and non-profit marketing. They actually find that famous people can influence the appreciation one has for a product or service, in a positive or negative direction. This suggests that it makes sense studying who negative influencers are, how they behave and how they can be detected in an online platform. Not limited to celebrities, people are more incline to follow users disclosing their personal information (Forman et al., 2008). The members of an online community rate reviews containing descriptive identity information more positively, and the prevalence of identity information disclosure by reviewers is associated with increased subsequent sales of online products. In addition, the shared geographical location increases the relationship between disclosure and product sales.

Wrapping up these important results, we can say that buyers are influenced by positive eWOM, especially if it is performed by nearby identifiable users; even more, celebrities can change the appreciation that people have for a product or a service. But the consequences are not just limited to customers. Even internal decision-making processes of businesses can be influenced by online review systems (Aggarwal et al., 2012). The diffusion of personal opinions through the Internet has radically changed the concept of reviewing a product or a service that one has in traditional media. In fact, online review platforms offer to users a space where they can express their *unfiltered* thoughts on products or services. In particular, eWOM encourages a two-way communication between a source and a reader, thus being more engaging. A very important result of (Aggarwal et al., 2012) is that eWOM helps companies to obtain higher product and service evaluations and, if necessary, higher amounts of funding; furthermore, it influences the decision-making processes of companies, showing that its power is not limited only to buyers. The other important result of (Aggarwal et al., 2012) is that the effect of negative eWOM is much greater than the one of positive eWOM.

Negative reviews open up many research issues. One of them is finding out what drives users to write negative reviews. Discontent, or "disconfirmation", with a product or service has been studied as a cause of this phenomenon. The authors of (Ho et al., 2008) define disconfirmation as the discrepancy between the expected evaluation of a product and the evaluation of the same product performed by experts. In particular, they find that a person is more likely to leave a review when the disconfirmation she encounters is great. They also find that the evaluation published by a person may not reflect her post-purchase evaluation in a neutral manner; indeed, the direction of such polarization is in agreement with disconfirmation.

The authors of (Yin et al., 2016) introduce a theory about the initial beliefs of a consumer when she is looking for a product. According to this theory, a consumer forms an initial judgement about a product based on its summary rating statistics. This initial belief plays a key role in her next evaluation of the review. To prove their conjecture, the authors of (Yin et al., 2016) collected the application reviews from Apple Store from July 1^{st} to August 31^{st}, 2013. By analyzing these reviews they show the existence of a confirmation bias, which outlines the tendency of consumers to perceive reviews confirming (resp., disconfirming) their initial beliefs as more (resp., less) helpful. This tendency is moderated by the consumer confidence in their initial beliefs. This bias also leads to a greater perceived helpfulness of positive reviews when the average product rating is high, and of negative reviews when the average product rating is low.

6.2 Main findings of the knowledge extraction process

In the Introduction, we specified that the main novelties of this paper concern: *(i)* the definition of the two social network based models of Yelp; *(ii)* the definition of three Yelp user stereotypes and their characteristics; *(iii)* the construction of the profile of negative influencers in Yelp. We also pointed out that this paper aims at answering three research questions, namely: *(i)* What about the dynamics leading a Yelp user to publish a negative review? *(ii)* How can the interaction of these dynamics increase the "power" of negative reviews and people making them? *(iii)* Who are the negative influencers in Yelp? In order to obtain these results and answer these questions, we conducted a data analytics campaign that allowed us to formulate six implications.

The first tells that "The star-based review system of Yelp is positively biased. Indeed, almost all users assign a high number of stars to almost all businesses.". It can be explained by taking into account that Yelp's review system is based on a Likert scale, and it is well known that this scale is positively biased (Alexandrov, 2010; Peeters and Czapinski, 1990; Bertram, 2007). This implication does not provide unexpected information, but still represents an important confirmation about the correctness of our knowledge extraction process.

The second implication tells that "Score-dl-users play a key role in negative reviews. They are very keen on negatively judging the macro-category they mostly attend.". Unlike the first one, it was not expected. Its explanation partially comes from the first implication. Indeed, if it is true that the Likert scale is positively biased, then a user must be particularly motivated to give a negative rating. Moreover, if such an evaluation is given by a double life user, then it means that it is provided by a person potentially balanced in her evaluations (indeed, she gave both positive and negative evaluations in the past). If a person with these characteristics gives a negative review, it is reasonable to assume that she did so because she had "something important to say". In that case, she probably provides some well founded justifications for her dissatisfaction. In order to do this, she must be competent in that macro-category, which explains the last part of the implication.

The third implication tells that "A user has a very high influence on her/his friends when doing negative reviews.". The first part of it represents an expected result, and is easily explained by the homophily principle (McPherson et al., 2001). The second part was unexpected and can be explained by considering that several studies in related literature show that negative reviews and reviewers are stronger than positive ones.

The fourth implication tells that "Bridges have a much greater power of influence than non-bridges.". It represents a partially expected result if we consider that bridges generally have a high betweenness centrality and, thus, have the ability to convey an idea, sentiment or opinion from one macro-category to another.

The fifth implication tells that "There is a correlation between k-bridges and top central users.". At first glance, it may appear an expected result, but actually this is not the case. In fact, in some contexts, for example in a Social Internetworking System, bridges connecting different social networks are not necessarily power users (Buccafurri et al., 2013). Actually, the more the communities involved in a (multi-) network scenario are integrated, the more likely a bridge is also a power user. Based on this reasoning, and considering that Yelp's macro-categories are closely related to each other, because both a user and a business can belong to more macro-categories simultaneously, the result obtained is reasonable and motivated.

Finally, the sixth implication tells that "The main negative influencers in Yelp are score-dl-users who simultaneously are top central users (according to degree and/or eigenvector and/or PageRank centrality measures).". It is certainly unexpected and is one of the main findings of this paper. It was obtained by appropriately integrating the previous five implications. For this reason, the justifications underlying it are those that allowed us to explain the implications from which it derives.

6.3 Theoretical contributions and implications

This paper provides several theoretical contributions to the literature on online review systems and eWOM. First of all, it introduces a new multi-dimensional social network based model of Yelp. This model perfectly fits the category-based structure of this social medium. It represents Yelp as a set of 22 communities, one for each macro-category. At the same time, it models this social medium as a user network $\mathcal{U}$ where each node denotes a user and an arc between two nodes represents a generic relationship between the corresponding users. Our model can be used in several different scenarios, depending on the type of relationship one wants to represent. In our study, we have specialized it to two different types of relationships, namely the friendship between users (i.e., $\mathcal{U}^f$) and the co-review of the same business carried out by different users (i.e., $\mathcal{U}^{cr}$).

The usage of our model, together with a set of experiments performed on a Yelp dataset, allowed us to show that the star-based review mechanism of Yelp is positively biased. This fact implies that a user must have a strong motivation to write a negative review. In turn, this implies that all information about negative reviews and negative influencers in Yelp is extremely valuable.

After that, thanks to our multi-dimensional model, we were able to define different stereotypes of users in Yelp. In particular, we considered three different stereotypes, namely the bridges, the power users and the double-life users. Bridges are users connecting different communities in Yelp. They are crucial for the dissemination of information in this social platform. In fact, we have seen that the influence exerted by bridges is greater than the one exerted by non-bridges. Power users are very active in performing reviews in the categories of their interest. The amount of reviews they carry out makes them extremely important in the identification of potential influencers. Double-life users show different behaviors in the different categories in which they operate. They generally show a particular attention and severity in a category in which they are extremely experienced. This means that they can play a valuable role as influencers in this category.

We have defined our multi-dimensional model and these stereotypes with respect to Yelp. However, our model can be easily generalized to other online review platforms, such as TripAdvisor, as well as to other types of social platforms. In case of online review platforms, the extension of our model is immediate. In fact, it is sufficient to know and report in our model the hierarchy of categories underlying the online review platform. In case of other types of social media, the extension is possible and quite simple. In fact, it is sufficient to specify a (possibly hierarchical) mechanism for dividing users into groups, as well as to identify the types of user relationships of interest. It seems quite obvious that friendship is a relationship of interest for any social platform. On the contrary, co-review does not always make sense and could be replaced by other types of relationships.

As for stereotypes, we observe that those considered in this paper are not the only ones possible for an online review platform. In the future, we plan to identify other stereotypes and study their contribution to the extraction of useful knowledge from Yelp. At the same time, the three stereotypes identified in this paper can be directly extended to any other online review platform. The concept of power user can be easily extended to any social platform and any online social network too. The concept of bridge and double-life user can be extended only to those cases where users of a social platform can be organized into communities based on some parameters. In this case, a bridge is a user acting as a link between two communities, while a double-life user is a user having different behaviors in different communities.

The last theoretical contribution of this paper concerns the definition of the Negative Reviewer Network. This model plays an extremely important role in the study of negative reviews and, above all, in the identification of negative influencers, who correspond to nodes with high degree centrality and/or high eigenvector centrality, as we have seen in Section 5.6. Analogously to what happens for the other theoretical tools introduced in this paper, the extension of this model to other online review platforms is immediate. Instead, its extension to other types of social platforms is much less simple than the other models and concepts seen above. In fact, by its nature, the Negative Reviewer Network is specifically designed to model negative reviews and reviewers. Therefore, its extension is only possible by identifying other negative behaviors that one wants to study and by defining a form of co-participation of multiple users to these behaviors.

6.4 Implications for practice

Starting from the theoretical background, the hypotheses made and the implications confirming them, we can outline different applications of the knowledge extracted in this paper to real life scenarios. In particular, we can identify two different perspectives, i.e., the business and the user ones.

The business perspective concerns all the possible actions that a company can take to expand its customer base, to improve its brand image or to extend the products/services it offers. In this context, the user stereotypes identified in this paper and the implications associated with them can be extremely useful. Let us consider, for example, k-bridges. We have seen the extremely important role that they play in disseminating information between different communities. In Section 6, we have also seen that past literature highlights the strong impact that negative reviews can have. In this context, a k-bridge making a negative review could have a disruptive effect on a business image.

Therefore, the possibility of detecting k-bridges provided by our approach can become a valuable tool for a business, which can adopt a variety of policies aiming at improving their evaluation of its products/services from negative to neutral or, even, positive. Another extremely important policy in this sense could regard the promotion of a business to k-bridges who do not know it. This could favor the knowledge of this business in all the communities which the k-bridges belong to. In fact, a k-bridge belonging to a community where a business is well known and another community where this latter is unknown could become a promoter of the business from the former community to the latter one.

Another important application that could leverage k-bridges is the expansion of products/services offered by a business towards new categories, or even new macro-categories, of Yelp. One way to increase the chance of designing new products/services being of interest to users could be as follows. A business could identify all the k-bridges belonging to the categories in which it is already known and its products/services are highly appreciated. Then, it could determine the other categories of products/services where the identified k-bridges have performed revisions; in fact, the products/services of these last categories could be of interest for the potential customers of this business. The greater the number of k-bridges that have shown interest in these categories, the more likely customers belonging to them will be attracted by the business if it expands its offers towards these markets.

A further application of k-bridges, collateral to the one seen above, concerns advertising campaigns.

In fact, knowing which are the most promising communities where proposing new products/services also implies being able to carry out advertising campaigns focusing on them. In this way, the effectiveness and efficiency of the advertisement activity in terms of time and costs are increased.

However, k-bridges are not the only stereotype identified in this paper having important practical applications. In fact, both power users and double-life users are equally important. Since the latter two stereotypes appear within the definition of negative influencers, we now see some possible applications of this last concept that subsumes the other two ones. Negative influencers have two important characteristics. The first concerns the high value of network centrality measures (degree centrality and/or eigenvector centrality and/or PageRank), which makes them very influential in the communities where they operate. The second concerns their behavior in carrying out reviews. In fact, we have seen that a negative influencer, being a score-dl-user, tends to give positive reviews in the categories of lesser interest, while she is very demanding and severe in the categories in which she is more experienced and that interest her the most. This also assumes that such a user generally has a recognized leadership exactly in the category in which she is most severe. Therefore, it becomes crucial for a business in that category taking all possible actions to ensure that she takes a neutral, or hopefully a positive, attitude towards the products/services it offers. On the other hand, as we have seen for k-bridges, it is possible to think of targeted advertising and marketing actions on these users that, if successful, are characterized by a high level of efficiency and effectiveness.

So far we have seen the possible exploitations of our knowledge patterns from the business viewpoint. Now, we want to see how the same patterns can have practical implications for the user as well. In particular, we want to consider what benefits a user can get by looking at other relevant users (such as k-bridges, power users, influencers) in Yelp.

A first benefit can be obtained from the examination of the reviews of negative influencers in Yelp. Based on the knowledge we have extracted, we can assume that these users are very experienced in a certain category and very severe in exactly that category. Therefore, if these users have issued positive reviews on the products/services of a business in that category, it is very likely that they are of high quality.

A second benefit for a user concerns the knowledge of the features characterizing the profile of an influencer in Yelp. This knowledge becomes extremely useful if she wants to become an influencer in that social medium. In fact, based on the implications derived in our paper, the user knows that she has a better chance to become an influencer if she becomes a k-bridge. As a consequence, she will have to be active in making revisions in multiple categories. In addition, she should be a power user; therefore, she must have many friendship and co-review relationships (which implies she has a high degree centrality). Alternatively, she can have a limited number of friendship and co-review relationships as long as the users connected to her are, in turn, power users (which implies she has a high eigenvector centrality). Finally, she must identify one or more categories in which she wants to be an influencer and develop a high experience in them in order to give severe, but correct, reviews.

The knowledge extracted in our paper can also be useful to define recommender systems for users who want to discover new products/services. This can be done, for example, by leveraging k-bridges. In fact, assume that a user follows some categories. It is possible to identify all the k-bridges of these categories and, for these k-bridges, to consider the categories followed by them. In this way, it is possible to identify which categories are the most followed by these k-bridges. If one of these categories is not already followed by the user, it is possible to recommend it to her. This very general approach could be further refined by examining the proximity, in the Yelp hierarchy, of candidate categories to those already followed by the user. A further refinement could assign different weights to the different k-bridges, based on the similarity of their past evaluation to those of the user of interest on the same products/services, or based on the number of categories already followed by both them and the user of interest.

6.5 Limitations and future research directions

The theoretical tools introduced in this paper (i.e., the multi-dimensional social network based model of Yelp, the stereotypes and the Negative Review Network), together with the hypotheses formulated and the implications confirming them, have allowed us to shed light on the phenomenon of negative reviews and negative influencers in Yelp. The tools proposed and the approach followed are sufficiently general to be extended directly to other online review platforms and, after some generalizations, to any social platform. However, they are to be considered simply as a first step in this direction, because they are not free from limitations, whose knowledge paves the way to new future research investigations.

The first limitation of our approach is that it is exclusively structural and does not take semantics into account. Actually, a more focused study on the contents of negative reviews would be necessary to understand the reasons that led users to formulate them. This would increase the effectiveness and efficiency of the applications of our approach discussed in Section 6.4. In fact, given a service/product receiving many negative reviews, we could strive to understand the main reasons for this fact and, therefore, make the appropriate improvements aimed at satisfying as many users as possible in the shortest time.

An in-depth semantic analysis of reviews would also be extremely useful to define one or more taxonomies of negative influencers. This would allow us to classify them based not only on the products/services they criticize, as in the present approach, but also on the main reasons for negativity (which would give us several indications on where intervening first or mainly). Semantic knowledge would also allow us to better evaluate negative influencers in order to understand who give plausible reasons and who, instead, are prevented, regardless it happens. As a matter of fact, a business could make an effective and efficient recovery work on the former category of influencers, while it could decide not to intervene on the latter one, because the possibility of making them neutral or positive is low.

Another limitation of our approach, which is, at the same time, a potential future development of our research concerns stereotypes. In this paper, we have presented three of them, namely the k-bridges, the power users and the double-life users. Their identification was driven by our research needs. However, we believe that several other stereotypes could be defined and that it could be even possible to go so far as to define a real taxonomy of stereotypes for both Yelp and other online (review) platforms. These would become a real toolbox available to decision makers when they need to make decisions regarding the products/services provided by their business (for instance, to determine those ones to be removed from catalogues, new ones to be proposed, existing ones to be modified for making them more in line with user needs and desires, etc.).

A third limitation of our approach, which is also linked to current technological limitations expected to become less impacting in the future, concerns the possibility of studying all these phenomena over time. In fact, our current approach is based on a temporal (albeit wide) photograph of the negative reviews of Yelp. It is not incremental and, if we want to study the evolution of a phenomenon over time, we should take more datasets referring to different times and study them separately. However, this does not allow us to have a continuous monitoring of the phenomenon, in order to capture any changes regarding it (for instance, any change of how some products/services are perceived by users) as soon as possible. The weight of this limitation (and, consequently, the relevance of overcoming it) is smaller in substantially stable socio-economic conditions, because user perceptions of products/services change very slowly over time in this scenario. Instead, it becomes crucial in historical periods characterized by sudden and disruptive phenomena (think, for instance, of the current COVID-19 pandemic), capable of upsetting all previous mental patterns of people's judgement. In this case, having the possibility of immediately understanding the changed perceptions of users about products/services and/or the appearance of new needs, with the consequent demand for new products/services, can allow a business to gain a huge advantage over its competitors. More importantly, this feature would allow the whole ecosystem of public and private product/service providers to be efficient and effective in responding to people demands.

7 Conclusion

In this paper, we dealt with the phenomenon of negative reviews in Yelp and outlined the profile of negative influencers. To this end, we used a new multi-dimensional social network based model of Yelp, several stereotypes of Yelp users derived from it, and a Negative Reviewer Network. Then, we formulated several hypotheses and we evaluated their correctness through an experimental campaign. In particular, at the end of our activities, we obtained the following knowledge patterns: *(i)* the star-based review system of Yelp is positively biased; *(ii)* bridges and double-life users play a key role in negative reviews; *(iii)* a user has a high influence on her friends when doing negative reviews; *(iv)* the main negative influencers in Yelp are score-dl-users who simultaneously are top central users (according to degree and/or eigenvector and/or PageRank centrality measures).

References

A. Agarwal, A. Pelullo, and R. Merchant. "ld": the Word Most Correlated to Negative Online Hospital Reviews. *Journal of General Internal Medicine*, pages 1–2, 2019.

R. Aggarwal, R. Gopal, A. Gupta, and H. Singh. Putting Money Where the Mouths Are: The Relation Between Venture Financing and Electronic Word-of-Mouth. *Information Systems Research*, 23(3):976–992, 2012. INFORMS.

A. Alexandrov. Characteristics of single-item measures in Likert scale format. *The Electronic Journal of Business Research Methods*, 8(1):1–12, 2010.

S. Angelidis and M. Lapata. Multiple instance learning networks for fine-grained sentiment analysis. *Transactions of the Association for Computational Linguistics*, 6:17–31, 2018.

J. Arthur, D. Etzioni, and A. Schwartz. Characterizing extremely negative reviews of total joint arthroplasty practices and surgeons on yelp.com. *Arthroplasty Today*, 2019.

C. Aslay, L. Lakshmanan, W. Lu, and X. Xiao. Influence maximization in online social networks. In *Proc. of the ACM International Conference on Web Search and Data Mining (WSDM'18)*, pages 775–776, Marina del Rey, CA, USA, 2018. ACM.

S. Basuroy, S. Chatterjee, and S. Ravid. How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, 67(4):103–117, 2003. SAGE Publications.

K. Bauman and A. Tuzhilin. Discovering contextual information from user reviews for recommendation purposes. In *Proc. of the International Workshop on New Trends in Content-Based Recommender Systems (CBRecSys @ RecSys 2014)*, pages 2–9, Foster City, CA, USA, 2014.

J. Berger, A. Sorensen, and S. Rasmussen. Positive effects of negative publicity: When negative reviews increase sales. *Marketing science*, 29(5):815–827, 2010. INFORMS.

M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, and D. Pedreschi. Foundations of Multidimensional Network Analysis. In *Proc. of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2011)*, pages 485–489, Kaohsiung, Taiwan, 2011. IEEE.

M. Berlingerio, F. Pinelli, and F. Calabrese. Abacus: frequent pattern mining-based community discovery in multidimensional networks. *Data Mining and Knowledge Discovery*, 27(3):294–320, 2013.

D. Bertram. Likert scales. *Retrieved November*, 2:2013, 2007.

P. Bhanodia, A. Khamparia, B. Pandey, and S. Prajapat. Online social network analysis. In *Hidden Link Prediction in Stochastic Social Networks*, pages 50–63. IGI Global, 2019.

A. Bhowmick, S. Suman, and B. Mitra. Effect of information propagation on business popularity: A case study on yelp. In *Proc. of the International Conference on Mobile Data Management (MDM'17)*, pages 11–20, Daejeon, South Korea, 2017. IEEE.

F. Buccafurri, G. Lax, A. Nocera, and D. Ursino. SISO: a conceptual framework for the construction of "stereotypical maps" in a Social Internetworking Scenario. In *Proc. of the International Workshop on New Frontiers in Mining Complex Knowledge Patterns at ECML/PKDD 2012 (NFMCP 2012)*, Bristol, UK, 2012.

F. Buccafurri, V. Foti, G. Lax, A. Nocera, and D. Ursino. Bridge Analysis in a Social Internetworking Scenario. *Information Sciences*, 224:1–18, 2013. Elsevier.

F. Buccafurri, G. Lax, A. Nocera, and D. Ursino. Moving from social networks to social internetworking scenarios: The crawling perspective. *Information Sciences*, 256:126–137, 2014. Elsevier.

F. Buccafurri, G. Lax, S. Nicolazzo, and A. Nocera. Comparing Twitter and Facebook user behavior: Privacy and other aspects. *Computers in Human Behavior*, 52:87–95, 2015. Elsevier.

D. Cai, Z. Shao, X. He, X. Yan, and J. Han. Community mining from multi-relational networks. In *Proc. of the European Conference on Principles of Data Mining and Knowledge Discovery (PKDD'05)*, pages 445–452, Porto, Portugal, 2005. Springer.

Q. Cao, W. Duan, and Q. Gan. Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, 50(2):511–521, 2011. Elsevier.

Y. Chang, C. Ku, and C. Chen. Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*, 48:263–279, 2019. Elsevier.

C. Cheung and M. Lee. What Drives Consumers to Spread Electronic Word of Mouth in Online Consumer-Opinion Platforms. *Decision Support Systems*, 53(1):218–225, 2012. Elsevier.

C. Cheung and D. Thadani. The impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model. *Decision Support Systems*, 54(1):461—470, 2012. Elsevier.

Y. Cui. An Evaluation of Yelp Dataset. *arXiv preprint arXiv:1512.06915*, 2015.

D. Davis, R. Lichtenwalter, and N. Chawla. Multi-relational link prediction in heterogeneous information networks. In *Proc. of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2011)*, pages 281–288, Kaohsiung, Taiwan, 2011. IEEE.

J. Fogel and S. Zachariah. Intentions to use the yelp review website and purchase behavior after reading reviews. *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1):53–67, 2017.

C. Forman, A. Ghose, and B. Wiesenfeld. Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets. *Information Systems Research*, 19(3):291—-313, 2008. INFORMS.

D. Franks, J. Noble, P. Kaufmann, and S. Stagl. Extremism propagation in social networks with hubs. *Adaptive Behavior*, 16(4):264–274, 2008.

M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973. JSTOR.

J. Guerreiro and P. Rita. How to predict explicit recommendations in online reviews using text mining and sentiment analysis. *Journal of Hospitality and Tourism Management*, 2019. Forthcoming. Elsevier.

A. Gulati and M. Eirinaki. With a Little Help from My Friends (and Their Friends): Influence Neighborhoods for Social Recommendations. In *Proc. of the World Wide Web Conference (WWW'19)*, pages 2778–2784, San Francisco, CA, USA, 2019. ACM.

A. Hicks, S. Comp, J. Horovitz, M. Hovarter, M. Miki, and J. Bevan. Why people use Yelp. com: An exploration of uses and gratifications. *Computers in Human Behavior*, 28(6):2274–2279, 2012. Elsevier.

Y. Ho, J. Wu, and Y. Tan. Disconfirmation Effect on Online Rating Behavior: A Structural Model. *Information Systems Research*, 28(3):626—642, 2008. INFORMS.

L. Hu, A. Sun, and Y. Liu. Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. In *Proc. of the International ACM SIGIR Conference on Research & development in information retrieval (SIGIR'14)*, pages 345–354, Gold Coast, Queensland, Australia, 2014. ACM.

C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proc. of the International AAAI Conference on Weblogs and Social Media (ICWSM'14)*, pages 216–225, Ann Arbor, MI, USA, 2014.

Y. Kang, J. Min, J. Kim, and H. Lee. Roles of alternative and self-oriented perspectives in the context of the continued use of social network sites. *International Journal of Information Management*, 33(3):496–511, 2013. Elsevier.

W. Kasper and M. Vela. Sentiment analysis for hotel reviews. In *Proc. of the International Computational Linguistics-Applications Conference*, volume 231527, pages 45–52, Jachranka, Poland, 2011.

A. Kavanaugh, D. Reese, J. Carroll, and M. Rosson. Weak ties in networked communities. *The Information Society*, 21 (2):119–131, 2005.

K. Kaviya, C. Roshini, V. Vaidhehi, and J. Sweetlin. Sentiment analysis for restaurant rating. In *Proc. of the International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM'17)*, pages 140–145, Chennai, India, 2017. IEEE.

C. Ke-Jia, Z. Pei, Y. Zinong, and L. Yun. iBridge: Inferring bridge links that diffuse information across communities. *Knowledge-Based Systems*, 192, 2020. Elsevier.

J. Kim, J. Bae, and M. Hastak. Emergency information diffusion on online social media during storm Cindy in US. *International Journal of Information Management*, 40:153–165, 2018. Elsevier.

J. Knoll and J. Matthes. The Effectiveness of Celebrity Endorsements: A Meta-Analysis. *Journal of the Academy of Marketing Science*, 45(1):55–75, 2017. Springer.

N. Kumar and I. Benbasat. Research note: the influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research*, 17(4):425–439, 2006. INFORMS.

K. Lee, J. Ham, S. Yang, and C. Koo. Can You Identify Fake or Authentic Reviews? An fsQCA Approach. In *Information and Communication Technologies in Tourism 2018*, pages 214–227, Jonkoping, Sweden, 2018. Springer.

X. Lei and X. Qian. Rating prediction via exploring service reputation. In *Proc. of the International Workshop on Multimedia Signal Processing (MMSP'15)*, pages 1–6, Xiamen, China, 2015. IEEE.

J. Leskovec, L. Adamic, and B. Huberman. The dynamics of viral marketing. *ACM Transactions on the Web*, 1(1):5, 2007. ACM.

M. Li, C. Tan, K. Wei, and K. Wang. Sequentiality of Product Review Information Provision: An Information Foraging Perspective. *MIS Q.*, 41(3):867–892, 2017. Management Information Systems Research Center.

Y. Lim and B. V. D. Heide. Evaluating the wisdom of strangers: The perceived credibility of online consumer reviews on Yelp. *Journal of Computer-Mediated Communication*, 20(1):67–82, 2014. Oxford University Press.

X. Lin and X. Wang. Examining gender differences in people's information-sharing decisions on social networking sites. *International Journal of Information Management*, 50:45–56, 2020. Elsevier.

M. Luca. Reviews, reputation, and revenue: The case of Yelp.com. *Harvard Business School Working Paper*, (12-016), 2016.

M. Luca and G. Zervas. Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427, 2016.

X. Luo. Quantifying the long-term impact of negative word of mouth on cash flows and stock prices. *Marketing Science*, 28(1):148–165, 2009. INFORMS.

W. Maharani, Adiwijaya, and A. Gozali. Degree centrality and eigenvector centrality in twitter. In *Proc. of the International Conference on Telecommunication Systems Services and Applications (TSSA'14)*, pages 1–5, Bali, Indonesia, 2014. IEEE.

J. Malbon. Taking fake online consumer reviews seriously. *Journal of Consumer Policy*, 36(2):139–157, 2013. Springer.

M. McPherson, L. Smith-Lovin, and J. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444, 2001. JSTOR.

A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance. What yelp fake review filter might be doing? In *Proc. of the International AAAI Conference on Weblogs and Social Media (ICDSM'13)*, Boston, MA, USA, 2013.

M. Nakayama and Y. Wan. The cultural impact on social commerce: A sentiment analysis on yelp ethnic restaurant reviews. *Information & Management*, 56(2):271–279, 2019. Elsevier.

H. Nam, Y. Joshi, and P. Kannan. Harvesting brand information from social tags. *Journal of Marketing*, 81(4):88–108, 2017.

P. Nokhiz and F. Li. Understanding rating behavior based on moral foundations: The case of Yelp reviews. In *Proc. of the International Conference on Big Data (Big Data 2017)*, pages 3938–3945, Boston, MA, USA, 2017. IEEE.

A. Parikh, C. Behnke, M. Vorvoreanu, B. Almanza, and D. Nelson. Motives for reading and articulating user-generated restaurant reviews on yelp. com. *Journal of Hospitality and Tourism Technology*, 5(2):160–176, 2014.

A. Parikh, C. Behnke, B. Almanza, D. Nelson, and M. Vorvoreanu. Comparative content analysis of professional, semi-professional, and user-generated restaurant reviews. *Journal of Foodservice Business Research*, 20(5):497–511, 2017.

G. Peeters and J. Czapinski. Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European review of social psychology*, 1(1):33–60, 1990. Taylor & Francis.

M. Potamias. The warm-start bias of Yelp ratings. *arXiv preprint arXiv:1202.5713*, 2012.

J. Qiu, Y. Li, and Z. Lin. Does Social Commerce Work in Yelp? An Empirical Analysis of Impacts of Social Relationship on the Purchase Decision-making. In *Proc. of the Pacific Asia Conference on Information Systems (PACIS'18)*, page 16, Yokohama, Japan, 2018.

J. Qiu, Y. Li, and Z. Lin. Detecting Social Commerce: An Empirical Analysis on Yelp. *Journal of Electronic Commerce Research*, 21(3):168–179, 2020. Journal of Electronic Commerce Research.

A. Saxena, R. Gera, I. Bermudez, D. Cleven, E. Kiser, and T. Newlin. Twitter Response to Munich July 2016 Attack: Network Analysis of Influence. *Frontiers in Big Data*, 2:17, 2019. Forthcoming. Frontiers.

D. Schuff and S. Mudambi. What makes a helpful online review? A study of customer reviews on Amazon.com. *Social Science Electronic Publishing*, 34(1):185–200, 2012. Elsevier.

V. Setyani, Y. Zhu, A. Hidayanto, P. Sandhyaduhita, and B. Hsiao. Exploring the psychological mechanisms from personalized advertisements to urge to buy impulsively on social media. *International Journal of Information Management*, 48:96–107, 2019. Elsevier.

W. Shen, Y. Hu, and J. Ulmer. Competing for Attention: An Empirical Study of Online Reviewers' Strategic Behavior. *MIS Q.*, 39(3):683–696, 2015. Management Information Systems Research Center.

X. Shi, B. Tseng, and L. Adamic. Looking at the blogosphere topology through different lenses. In *Proc. of the International Conference on Weblogs and Social Media (ICWSM'07)*, Boulder, CO, USA, 2007.

R. Singh, J. Woo, N. Khan, J. Kim, H. Lee, H. Rahman, J. Park, J. Suh, M. Eom, and N. Gudigantala. Applications of machine learning models on yelp data. *Asia Pacific Journal of Information Systems*, 29(1):117–143, 2019.

Y. Sun and J. Paule. Spatial analysis of users-generated ratings of yelp venues. *Open Geospatial Data, Software and Standards*, 2(1):5, 2017.

P. Ting, S. Chen, H. Chen, and W. Fang. Using big data and text analytics to understand how customer experiences posted on yelp. com impact the hospitality industry. *Contemporary Management Research*, 13(2), 2017. Academy of Taiwan Information Systems Research.

Q. Xuan, X. Shu, Z. Ruan, J. Wang, C. Fu, and G. Chen. A self-learning information diffusion model for smart social networks. *IEEE Transactions on Network Science and Engineering*, 2019. Forthcoming.

Y. Yang, N. Chawla, Y. Sun, and J. Hani. Predicting links in multi-relational and heterogeneous networks. In *Proc. of the International Conference on Data Mining (ICDM'12)*, pages 755–764, Bruxelles, Belgium, 2012. IEEE.

D. Yin, S. Bond, and H. Zhang. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS quarterly*, 38(2):539–560, 2014. JSTOR.

D. Yin, S. Mitra, and H. Zhang. When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Information Systems Research*, 27(1):131–144, 2016. INFORMS.

K. Zhang, S. Zhao, C. Cheung, and M. Lee. Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, 67:78–89, 2014. Elsevier.

Y. Zhang, S. Shi, S. Guo, X. Chen, and Z. Piao. Audience management, online turbulence and lurking in social networking services: A transactional process of stress perspective. *International Journal of Information Management*, 56:102233, 2021. Elsevier.

Z. Zhang, Q. Li, D. Zeng, and H. Gao. User community discovery from multi-relational networks. *Decision Support Systems*, 54(2):870–879, 2013. Elsevier.

M. Zhou, X. Cai, Q. Liu, and W. Fan. Examining continuance use on social network and micro-blogging sites: Different roles of self-image and peer influence. *International Journal of Information Management*, 47:215–232, 2019. Elsevier.