Feature-based opinion mining: A proposal to enhance customer focus based on text feedback

UNIVERSITEIT VAN AMSTERDAM

Business Information Systems
MSc. INFORMATION STUDIES

Author: Antigoni Kourou
Student number: 11118172
Supervisor: Dr. Maarten Marx
Signature:

JUNE 2016

Dedication

Here goes the dedication.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

ABSTRACT

Table of Contents

List of Figures

CHAPTER 1

INTRODUCTION

The exponential growth of Information and Communication Technologies (ICTs) has played a great role in the development of tourism industry. Electronic tourism (e-tourism) is the application of ICTs for tourism purposes, including the digitalization of all its processes and value chains [2]. Increasingly, ICTs provide users with access to many sources of information and has eventually affected the consumer behavior in tourism industry [18]. In order to remain a strong competitor, service providers have to keep their customers happy. Nowadays there is a huge variety of web services which provide a great complexity and diversity of recommended offers for meeting the demands of travelers. However, the personalized consumption patterns and individualistic lifestyles make it incredibly difficult for the service providers to anticipate tourists' behavior [19]. For decision makers to successfully understand the requirements, needs, desires and preferences of the customer, detailed information has to be obtained. Travelers, making use of the ICT tools that facilitate the information retrieval and decision making processes, have now direct access to all types of information provided by tourism agencies, companies, marketers, enterprises or other users. Furthermore, ICTs and Internet have transformed e-tourism markets from customer-centric to customer-driven [3], meaning that users play a major role in creating and sharing traveling information through blogs and review websites. Online feedback mechanisms, also known as reputation systems, *have emerged as a viable mechanism for fostering cooperation among strangers in such settings* [5]. Examples of these systems, for instance TripAdvisor, Booking.com or AirBnb, after each trade encourage both parties to give feedback about their trading partner based on their own experience. Customer feedback is an essential component in every modern business‚Äô tool kit. A big number customer feedback software tools exists for helping service providers to measure and improve customer satisfaction, identify unhappy customers, reduce churn and get valuable insight from customers. However, most of these tools are very expensive, considering that most of service providers can make use of their own feedback mechanisms.

1.1 Problem statement

The most common types of consumer generated feedback are ratings from 1-5 stars and general text comments. An important separation exists between the distinct role of text comments as tacit knowledge and ratings as explicit knowledge and the ways they are analyzed. For online marketplaces to succeed, their feedback technologies must be able to not only collect users feedback, but to properly analyze it and utilize for decision making purposes [20]. However, current online travel systems, aiming to assist the consumer in finding suitable offers, filter the information based on location-price factor and on the overall ratings accumulated from the feedback system, meaning that text reviews are revealed for the public to read but they do not directly affect the overall analysis. Focusing on solely numerical ratings and ignoring the importance of text feedback leads to two major issues for feedback systems.

First, many academic papers on online reputation systems and building trust in the online marketplace report the existence of bias on online reviews [1, 6–9, 22], thus reducing the bias of these systems is an important issue towards a more efficient online feedback system. Utilizing only biased ratings does not necessarily mean that the top result is the most suitable option for a certain user considering the personal requirements. On the other hand, the analysis of ratings does not indicate much information for the service providers, who aim to acquire knowledge on how to improve their services. In order to gain detailed insights from the feedback system, some service providers including Airbnb, ask its users to rate not only the overall quality of the listing, but also six accommodation features. However, since the overall ratings are biased [8], how do we make sure that the ratings for specific features are objective? From this point of view, the bias on quantitative data of the system raises the issue of reliability on the system itself.

Second, the incompleteness of the analysis based solely in quantitative data leads to the need for complementary analysis. Customers' needs are considered multi-dimensional and difficult to measure on discrete scales such as ratings [16], therefore a customer has to extract the needed information from different sources and types of information provided by agencies, companies or other users. According to [20], text comments are particularly interesting for the audience as a new trust-building means in online marketplaces by revealing hidden knowledge, which is often underestimated from their owners and cannot be described by negative/positive ratings. Furthermore, [8] suggests that text opinions influence the decision making process even when the ratings are high. In the Airbnb feedback system a negative rating is followed by a text in 45% of the cases, which implies the great power of text analysis for discovering deeper insights for the listing [8]. Acknowledging the importance of text comments, some feedback systems often offer summaries to all text comments, which mostly consist on a bunch of most used words. However, this bag of words does not necessarily cover the features that a certain user is interested in, neither the features that need to be improved. The users or service providers still have to read all the text comments related to the feature, meaning that it still does not reduce much of the work. A survey by [21] asked the respondents to indicate how many feedback comments they examined before each online transaction. The result showed that 81% reported examining 25 comments (one webpage), 5% viewed 50 comments, 11% more than 50 ones, and only 3% did not examine any text comments. These findings reveal that despite the importance of text feedback to the users, it is difficult for them to access the meaning of numerous text comments [21]. Given this situation, the average human reader will have difficulties on identifying and extracting the relevant information from the opinions in them. Automated analysis systems are thus needed [14].

1.2 Research question

Natural language processing (NLP) enable computers to derive meaning from any human written input, including their opinions. The NLP methods for doing so fall into the category of sentiment mining methods, known also as *opinion mining*. Examples of their application include mainly the movie rating systems (Netflix, IMDb) and the product rating systems (eBay). However, the importance of extracting sentiment of features from comments, besides their overall positive/negative sentiment is often ignored in the literature. This research proposes the implementation of feature-based opinion mining methods from complementing the analysis of customer feedback in e-tourism and accommodation market. The proposed approach uses an ontology based approach combined with sentiment mining techniques for generating opinion scores for each accommodation feature mentioned in the reviews of a feedback system. This solution deals with the two issues mentioned above, bias of ratings and the need for complementary text analysis for feature extraction. From its point of view, both issues can be brought together as one, since text analysis is believed to contribute to a better rating systems by reducing its bias [8].

This paper answers firstly the question of *how can opinion mining methods be aligned with the quantitative data analysis in order to enhance focus on customer feedback and produce detailed analysis results.* By estimating sentiment scores for the text reviews, the pipeline transforms the text data into discrete quantitative form, which can easily be analyzed for different purposes. Some of the implications of the analysis are treated in the next sections. In addition to this analysis, it is important to find out *how good can feature-based opinion mining estimate the quality of accommodation features in e-tourism feedback systems,* which measures the reliability of the analysis. By answering these two questions, the purpose of the paper is to offer to service providers a new reliable approach on enhancing focus on customers, based on users' generated content.

1.3 Research rationale and structure

Customer Focus Theory is one of the essential parts of Information Studies (IS) and it serves as a guidance for business on how to put focus on their customers, as their most valuable asset. From Figure 1.1 can be clearly seen that Customer Requirement, Information and Feedback are key factors on improving the relationships [15]. The added value of this research consist on automation of feedback analysis, which consequently enhances focus on the customer. From the literature point of view, it can be noticed a gap between the alignment of text analysis with quantitative data, where the opinions from text are mostly perceived as positive, negative or neutral.

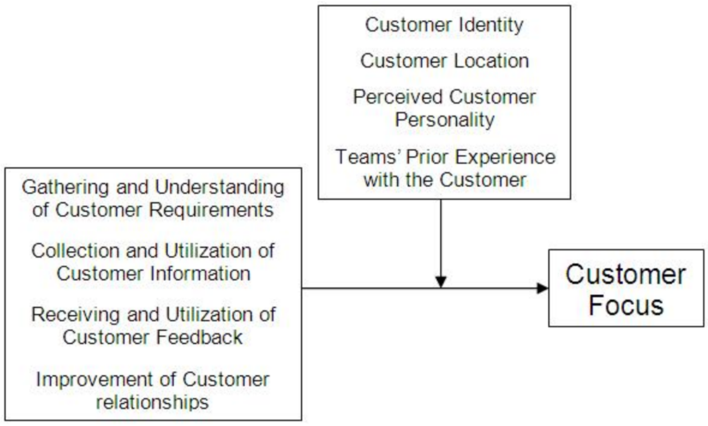

Figure 1.1: Customer Focus Theory

To provide an answer to the research questions, this paper is organized in seven chapters. The first part of the paper introduces the importance of gathering and analyzing customer feedback in e-tourism. The current state-of-the-art of feature-based opinion mining used for analyzing customer feedback is covered in the second part. The literature review includes articles published in a time frame of the last ten years and it leads the reader to the approach proposed for filling the literature gaps. The proposed approach is discussed by explaining each step of the pipeline developed for this research. The fourth chapter covers the methodology used in the research, from collecting the data to the analysis of the output. In chapter five, the whole proposed approach is evaluated based on human logic. Afterwards the results are presented. Finally, the paper discusses the implications of the proposed approach, limitations and further work to be done.

Chapter 2

Current state of knowledge

This chapter is a literature review of the research done until now on the field of sentiment mining for analysing customer feedback.

Chapter 3

The proposed approach

For transforming the text feedback into quantitative data, which will serve as input for analysis purposes, this paper proposes an approach as shown in Figure 3.1. The pipeline consists of four main steps: *pre-processing, feature identification, sentiment detection* and *data analysis*. Initially, the whole input of the pipeline is the huge corpus of text reviews stored in Neo4J database. This big amount of data from social is considered to be very noisy, therefore it will be cleaned up as be described in the following section. Features and sentiment are then identified in text based in sentence level. The output of these three stages consists of the quantitative sentiment scores for sentences of reviews and the features identified in them. This output will then serve as the input data of the final stage of the pipeline: data analysis. The whole pipeline is built using Python programming language and its related packages. All the data analysis and visualization are written in Jupyter notebook. The following sections will explain each step of the proposed approach in further details.

3.1 Pre-processing

The pipeline reads the text data from cloud with the help of *py2neo*, a toolkit for working with Neo4J from within Python applications, and it formulates the queries in Cypher, the querying language for Neo4J graph database. The reviews in the corpus are read one by one and each of them is checked if it fulfills the language requirements. In my work, the whole concept of the pipeline is developed in English, as the most used language in e-tourism websites and as the easiest language for text mining. Thus, every time that the algorithm runs into a review not in English, it will ignore the review and continue with the next one. For each English review detected, the algorithm will split the text into sentences, with the intention of detecting the

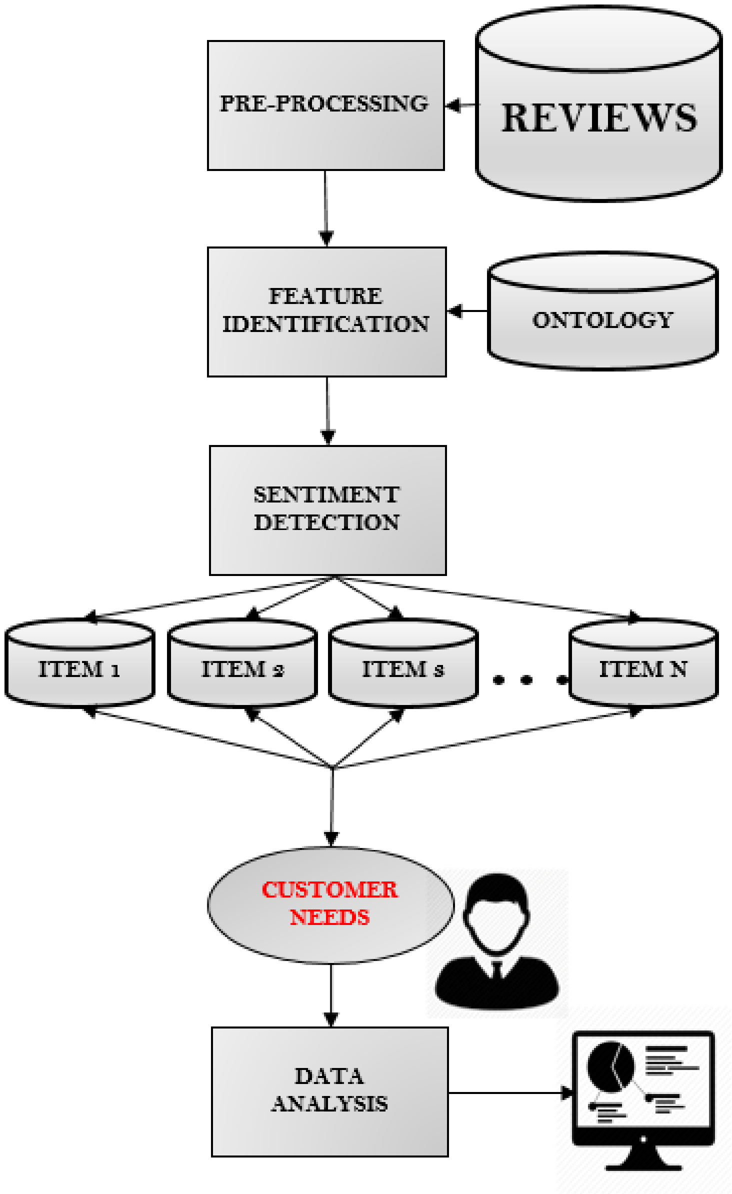

Figure 3.1: The proposed approach

sentiment within each sentence of the review and extracting features within these sentences. Afterwards tokenization and lemmatizing are performed to each sentence of the review, in order to cut them into words, which are candidates for representing an accommodation feature. For tokenization the pipeline uses the *TweetTokenizer*[1] package, part of NLTK and Twitter-aware designed, in order to be able to identify emoticons and adapt to new domains. On the other hand lemmatization is performed using *WordNet*[2], a large lexical database for English terms, in which

[1]http://www.nltk.org/api/nltk.tokenize.html
[2]https://wordnet.princeton.edu/

nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept [17]. The output of these two steps is a list of lemmas for each sentence, which will be used by the next step in the pipeline as explained in the following section.

3.2 Feature identification

As soon as each sentence in the corpus of reviews is represented as a list of lemmas, an ontology based approach is used to identify the accommodation features of each sentence. The ontology is chosen for feature identification as a way for detecting all the terms, concepts and relations linked to accommodation domain. Prior to identifying the features, the pipeline builds a list of synonyms, hyponyms and hypernyms for every term of the ontology. The list is based on the Synsets and relations between concepts as introduced by WordNet library. The lists of related terms is lemmatized, as explained above, and the duplicates are removed. The lemmatization process aims to create a list of related lemmas for each feature in the ontology, which would then be used for feature identification in the sentences. Thus, the feature identification steps is a string match of the list of lemmas for every single sentence, with the list of related lemmas of every feature in the ontology. Considering that the ontology consists of many terms, this sample pipeline is trained to identify only six features that Airbnb asks the customers to manually rate as part of their feedback, namely *accuracy, cleanliness, check-in, communication, location* and *value*. When a word in the sentence is identified to be a feature, the algorithm jumps to the next word of the sentence. Within one sentence more than one feature can be identified, as well as features can be mentioned more than once. Therefore, for every possible feature match, a match counter variable which keeps track of the features is assigned to the sentence. Here, the pipeline considers each sentence as independent and it ignores the logical connection between two sentences of the same review.

3.3 Sentiment detection

A very important part of the pipeline is sentiment detection. The algorithm used for this purpose is VADER [3], part of Python packages. VADER is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. Based on the comparisons of 22 sentiment mining tools on specific contexts, VADER is ranked as the best algorithm for comments and the second best for social networks [23]. VADER is able to deal with negation, capital letters, emoticons, punctuation types, mixed sentiment sentences (*but, however*and slangs. The pipeline asks uses VADER to return a sentiment score for every sentence of the review, which would afterwards serve as a discrete score for the identified features in that sentence. Considering that in one sentence, more than 1 feature can be identified or the same feature can

[3]https://pypi.python.org/pypi/vaderSentiment

be identified more than once as mentioned above, a probabilistic model is developed. This model serves for defining the probabilities that a sentiment score would reflect user's opinion on each explicit feature identified within one sentence. For a better understanding, let's have a look at the whole model of data, including the probabilistic sentiment scores.

The whole data corpus consists of 11 053 listings. Each listing in this corpus can be represented by the ID of the listing in the database L_{lis_id}. Every listing from this set has a number of reviews, which varies from one listing to the other and is identified from ID of the review $R_{lis\_id} = \{ R_0, R_1, R_2 ... R_r \}$. For a single review of a certain listing would be $R_{lis\_id,\ rev\_id}$. This review is composed by a number of sentences $S_{lis\_id,\ rev\_id} = \{S_0, S_1, S_2 ... S_s\}$. The pipeline is trained to detect the sentiment of all the sentences, which will be in the same form as the one above. In each of these sentences, the pipeline is trained to identify six accommodation features *F = {accuracy, check-in, cleanliness, communication, location, value}*. When the sentiment of a sentence is detected, it does not particularly refer to a certain feature. In order to calculate the part of the sentiment that belongs to a certain feature identified in the sentence, we use a probabilistic model. According to this model, the probability that the compound sentiment score refers to a certain feature is p_i, where i is the index of the feature from the list F. This probability is uniformly distributed between the identified features and is calculated as 1/k, where k is the number of features mentioned. Likewise, if the feature is not identified in the sentence the probability would be 0.

$$Sentiment = \sum_{i=1}^{n} p_i * sentiment_i \qquad (1)$$

This formula for calculating the sentiment of the features is part of the sentiment detection phase of the pipeline and it is repeated for every single sentence. For instance, considering the reviews of listing 24328, $R_{24328,\ 10146}$ would represent the review: "We had a great stay at Joe's. The handy guide to the house and neighborhood was much appreciated, as were Joe's clear instructions for checking in while he was away. The house was comfortable and eclectic, full of personal character." Thus, $S_{24328,\ 10146,\ 1}$ represents the second sentence of the text above. In this sentence can be

SENTENCE	SENTIMENT	ACCURACY	CHECK-IN	CLEANLINESS	COMMUNICATION	LOCATION	VALUE
The handy guide to the house and neighborhood was much appreciated, as were Joe's clear instructions for checking in while he was away.	0.851896	0	0.28394	0	0.283936943	0.28394	0

Figure 3.2: Example of the probabilistic sentiment results

identified three features *check-in, communication,* and *location*. The compound sentiment score of the sentence above has in this way to refer to one of these features in a probability of 33.3%. Figure 3.2 shows the results of the pipeline for this case. The output of the whole sentiment detection phase is a Comma Separated Vector (.csv) file, consisting of all listings, all listings' reviews, its respective sentences and the sentiment for the identified features on these sentences.

3.4 Data analysis

The last step of the pipeline and the most interesting one for service providers and business is data analysis. Up to here, we saw the pipeline processing all the text data step by step and transform it into meaningful sentiment scores for each sentence and each accommodation feature of the reviews. This scores are saved locally in .csv format and are analyzed with Pandas[4], an open source library providing high-performance, easy-to-use data structures and data analysis tools for the Python. The workspace of analysis and visualization of results is Jupyter Notebook, a web application that allows users to create and share documents that contain interactive code, equations, visualizations and explanatory text. The analysis includes calculation and visualization of sentiment scores per reviews and listings, frequency of mentioned features, ranking of listings based on overall sentiment or based on sentiment of one or more specific features, ranking of features for one listing based on their sentiment, identification of listings with the biggest number of reviews and identification on listings where the host have canceled once or more the reservation, computation and comparisons of sentiment scores from the pipeline with the ratings of Airbnb and so on. An example of the analysis, Figure 3.3 shows all the listings with low rating (less than 4 stars since the lowest rating is 3 stars) and visualizes the differences between the sentiment scores generated from the pipeline and the ratings found on Airbnb. The analysis can of course be extended and customized to match the specific interest of the service providers or the customers. The main results of this analysis with be explored in the results section, followed by the implementation of pipeline according to its usefulness for service providers and customers respectively.

```
In [31]: # Compare the ratings and sentiment scores from pipeline for listings which have less than 4 stars

         low_rating = comparison[comparison['Sentimentscore']<4]
         low_rating[['Overall','Sentimentscore']].plot(kind='bar')

Out[31]: <matplotlib.axes._subplots.AxesSubplot at 0x761dd50>
```

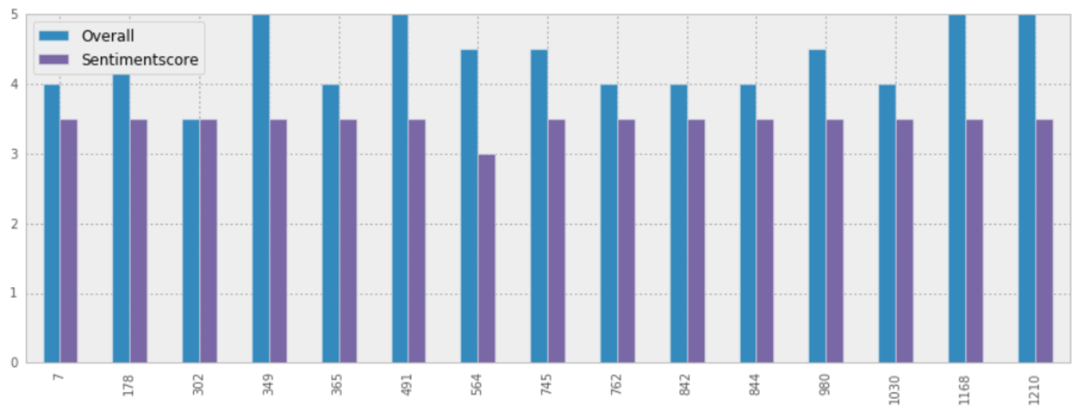

Figure 3.3: Example of data analysis in Jupyter Notebook with IPython & Pandas Library

[4]http://pandas.pydata.org/

Chapter 4

Methodology

In order to build the proposed solution, this research is based on two sources of data: text reviews and ontology. The next section aim to describe the methodology used for data collection and building the accommodation ontology. (+ pipeline coding?)

4.1 Accommodation ontology

In knowledge management and Semantic Web research areas, ontologies are considered essential in order to describe various concepts and relationships between them. Ontologies are formal specifications of a shared conceptualization of a domain. For the accommodation domain and related domains a few ontologies have been proposed, which cover different aspects in this domain. However these ontologies are often found as sub-ontologies of the e-tourism domain. In this research, the ontology of accommodation is based on several sources. Firstly HONTOLOGY [4], which is brought in alignment with Accommodation in QALL-ME, Tourist Accommodationsin DBpedia.org and Lodging Business in Schema.org, was initially created by following different scenarios of booking an accommodation and processing reviews in Web services. Secondly ACCO Accommodation Ontology [11], which is an extension of GoodRelations ontology [10], is based on Owl Ontology Language (OWL) and is supported by Google and Yahoo for the e-commerce accommodation offers. Besides these ontologies, further covers features mentioned in the Web based accommodation services of AirBnb.com, Booking.com and TripAdvisor.com.

4.2 Dataset

This subsection will describe the characteristics of data that will be used for research. In details how each property of the data scraped will be used, name of databases and so on.

The algorithm deals with a huge set of reviews, which are retrieved from the feedback system of Airbnb and they serve as the input of the pipeline. The whole corpus of reviews consists of 3.4GB, including reviews of the listings from the Netherlands, which is filtered to be focused only in Amsterdam. The data set is stored in cloud, in the Neo4J graph database, where the nodes can represent a listing, a guest or host, review or response. (Appendix figure?)

Chapter 5

Evaluation

For checking the efficiency of the proposed pipeline, its performance is compared to human logic for two main purposes. First, the evaluation aims to find out how well VADER, the selected sentiment detection algorithm, performs in the chosen dataset and secondly, how well can the pipeline identify the features of a certain sentence of review. To answer these two questions, a sample of 100 text reviews are randomly chosen from the dataset and are given to humans for evaluation. Each respondent is required to first read the sentences, then to estimate a score of sentiment for each of them on a scale from -1 to +1 and finally to mark the accommodation features, which the sentence in question refers to. These tasks are similar to what the pipeline is programmed to do. The evaluation form is completed by 5 humans, who come from different educational backgrounds, are geographically diverse and are Airbnb users. Their answers are analyzed and compared to the pipeline output using SPSS. Considering the diversity between users and in order to understand the common human logic, their evaluations are firstly compared to each other using correlation values between samples and distribution statistics. This analysis showed that three of the respondents had very similar answers, consisting in a correlation varying between 72% to 83 % between each other and a very similar distribution curve. Based on these results, the chosen sample to represent the average human logic are the average rates of the three "unbiased" respondents.

5.1 Evaluation of VADER algorithm

VADER is considered to be one of the best algorithms for sentiment mining with a very high accuracy [13]. However, when VADER was tested to perform in 22 different datasets, it was noticed an overall accuracy of 78.7%, which varies significantly from one dataset to the other [23].

To clear all the doubts of how VADER would perform on reviews and specifically in the chosen dataset of this research, its scores are compared to the average human rating. The correlation between the two sets based on cases results to be 70.8%. Although, this value is lower than the average case of VADER's accuracy, it indicates a satisfactory scale of accuracy for the Airbnb dataset. From Figure 5.1 can be noticed that the two datasets share the same mean (4.125 compared to 4.13) and have similar standard deviations. The differences between VADER scores

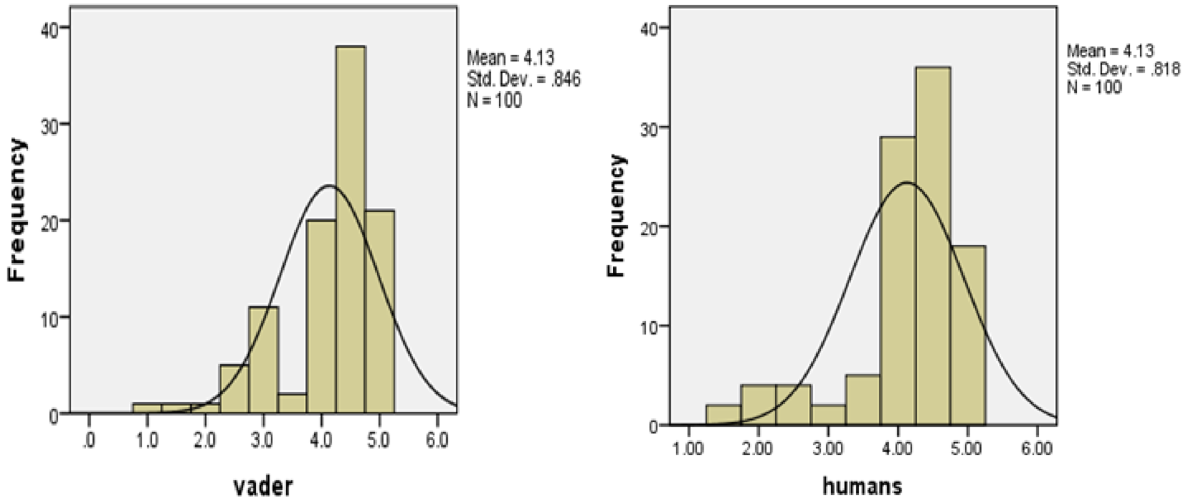

Figure 5.1: Distribution of VADER and humans' sentiment scores

and human evaluation are also calculated and they result in 53% of cases when VADER has detected the exact score as human logic, 24% cases where the difference is just half a star, 14% cases with one star difference and the 9% cases with more than one star difference. The last ones form the group of "errors" in VADER sentiment scores. The highest value of error is 2.5 stars, in only 3% cases, which means that VADER will never consider a very positive sentence as a very negative one and the other way round. However, it may consider these kind of sentences as neutrals or may be confused of the sign of sentences with slight sentiment.

5.2 Evaluation of feature identification part

The effectiveness of the proposed technique for feature identification is measured by using precision, recall and accuracy as suggested by [12]. The values are calculated for each of the features as:

$$Precision(p) = \frac{TP}{TP+FP} \qquad (2)$$

$$Recall(r) = \frac{TP}{TP+FN} \qquad (3)$$

$$Accuracy(a) = \frac{TP + TN}{TP + TN + FP + FN} \qquad (4)$$

where TP (true positives) is the number of sentences that the algorithm correctly identifies the right features; FP (false positives) is the number of sentences that the algorithm falsely extract wrong features; FN (false negatives) is the number of sentences that the algorithm fails to identify the right features and TN (true negative) is the number of sentences that the algorithm correctly does not identify any feature. The average human logic identifies a feature in a sentence when it is agreed by the majority of respondents. Ideally the algorithm shall have values of precision and recall close to 1 for each of the features. This evaluation of the pipeline as illustrated in Figure 5.2 shows that the algorithm manages to identify some features better than others. For example the *value* and *cleanliness* features of the listing are identified almost always with high precision and recall, but a low precision for *communication* means that the algorithm retrieves many FP. The opposite happens for *check in* when the precision is very high but many sentences fail to be identified. For the six features overall the algorithm reaches the accuracy and precision 77.8% and recall 79.2%. Future work needs to be done in improving this step of the pipeline.

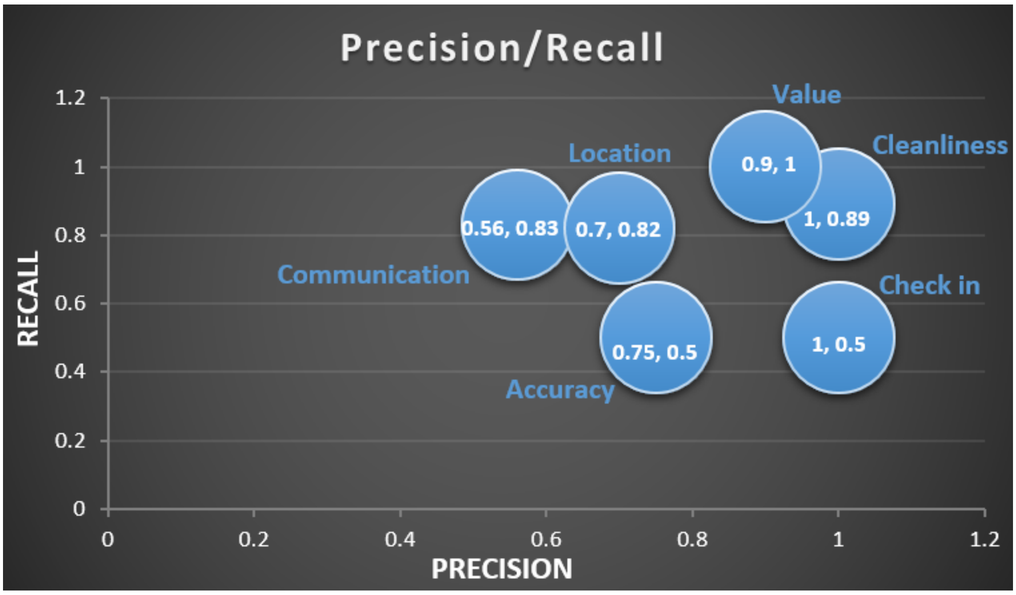

Figure 5.2: Precision/Recall matrix for feature identification

CHAPTER 6

DISCUSSION

Here will go the discussion.

6.1 Limitations

6.2 Implications

6.3 Future work

Chapter 7

Conclusions

Here will go the conclusions.

APPENDIX A

APPENDIX A

Begins an appendix

BIBLIOGRAPHY

[1] G. BOLTON, B. GREINER, AND A. OCKENFELS, *Engineering trust: reciprocity in the production of reputation information*, Management Science, 59 (2013), pp. 265–285.

[2] D. BUHALIS, *eTourism: Information technology for strategic tourism management*, Pearson Education, 2003.

[3] D. BUHALIS AND S. H. JUN, *E-tourism*, Contemporary tourism reviews, (2011), pp. 01–38.

[4] M. CHAVES, L. FREITAS, AND R. VIEIRA, *Hontology: A multilingual ontology for the accommodation sector in the tourism industry*, (2012).

[5] C. DELLAROCAS, *The digitization of word of mouth: Promise and challenges of online feedback mechanisms*, Management science, 49 (2003), pp. 1407–1424.

[6] C. DELLAROCAS AND C. A. WOOD, *The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias*, Management Science, 54 (2008), pp. 460–476.

[7] F. DINI AND G. SPAGNOLO, *Buying reputation on ebay: Do recent changes help?*, International Journal of Electronic Business, 7 (2009), pp. 581–598.

[8] A. FRADKIN, E. GREWAL, D. HOLTZ, AND M. PEARSON, *Bias and reciprocity in online reviews: Evidence from field experiments on airbnb (preliminary-do not cite without permission)*, (2016).

[9] A. GHOSE AND P. G. IPEIROTIS, *Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics*, Knowledge and Data Engineering, IEEE Transactions on, 23 (2011), pp. 1498–1512.

[10] M. HEPP, *Goodrelations: An ontology for describing products and services offers on the web*, in Knowledge Engineering: Practice and Patterns, Springer, 2008, pp. 329–346.

[11] ——, *Accommodation ontology language reference*, tech. rep., Technical report, Hepp Research GmbH, Innsbruck, 2013.

[12] J. HUANG, *Performance measures of machine learning*, University of Western Ontario, 2006.

[13] C. J. HUTTO AND E. GILBERT, *Vader: A parsimonious rule-based model for sentiment analysis of social media text*, in Eighth International AAAI Conference on Weblogs and Social Media, 2014.

[14] B. LIU, *Sentiment analysis and opinion mining*, Synthesis lectures on human language technologies, 5 (2012), pp. 1–167.

[15] G. LOHAN, K. CONBOY, AND M. LANG, *Examining customer focus in it project management: findings from irish and norwegian case studies*, (2011).

[16] M. LUO, R. FENG, AND L. A. CAI, *Information search behavior and tourist characteristics: The internet vis-à-vis other information sources*, Journal of Travel & Tourism Marketing, 17 (2005), pp. 15–25.

[17] G. A. MILLER, *Wordnet: a lexical database for english*, Communications of the ACM, 38 (1995), pp. 39–41.

[18] J. E. MILLS AND R. LAW, *Handbook of consumer behavior, tourism, and the Internet*, Psychology Press, 2004.

[19] M. NIEMANN, M. MOCHOL, AND R. TOLKSDORF, *Enhancing hotel search with semantic web technologies*, Journal of theoretical and applied electronic commerce research, 3 (2008), pp. 82–96.

[20] P. A. PAVLOU, *The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation*, Information Systems Research, 17 (2006), pp. 392–414.

[21] P. A. PAVLOU AND A. DIMOKA, *Institutional feedback technologies in online marketplaces: An investigation of feedback text comments, trust, and price premiums*, Trust, and Price Premiums (February 1, 2006), (2006).

[22] P. RESNICK, R. ZECKHAUSER, J. SWANSON, AND K. LOCKWOOD, *The value of reputation on ebay: A controlled experiment*, Experimental economics, 9 (2006), pp. 79–101.

[23] F. N. RIBEIRO, M. ARAÚJO, P. GONÇALVES, F. BENEVENUTO, AND M. A. GONÇALVES, *A benchmark comparison of state-of-the-practice sentiment analysis methods*, arXiv preprint arXiv:1512.01818, (2015).