Expressing Sentiments in Game Reviews

Ana Secui[1], Maria-Dorinela Sirbu[1], Mihai Dascalu[1](✉), Scott Crossley[2], Stefan Ruseti[1], and Stefan Trausan-Matu[1]

[1] Computer Science Department, University Politehnica of Bucharest, Bucharest, Romania
{ana.secui,maria.sirbu}@cti.pub.ro, {mihai.dascalu, stefan.ruseti,stefan.trausan}@cs.pub.ro

[2] Applied Linguistics and ESL, Georgia State University, Atlanta, USA
scrossley@gsu.edu

Abstract. Opinion mining and sentiment analysis are important research areas of Natural Language Processing (NLP) tools and have become viable alternatives for automatically extracting the affective information found in texts. Our aim is to build an NLP model to analyze gamers' sentiments and opinions expressed in a corpus of 9750 game reviews. A Principal Component Analysis using sentiment analysis features explained 51.2 % of the variance of the reviews and provides an integrated view of the major sentiment and topic related dimensions expressed in game reviews. A Discriminant Function Analysis based on the emerging components classified game reviews into positive, neutral and negative ratings with a 55 % accuracy.

Keywords: Natural Language Processing · Sentiment analysis · Opinion mining · Lexical analysis

1 Introduction

The domain of opinion mining and sentiment analysis refers to extracting information about feelings, ideas and emotions by analyzing textual productions using Natural Language Processing (NLP) techniques [1, 2]. There is now a lot of interest in this direction due to the tremendous volume of messages, reviews and discussion forum posts on social networks like Facebook or Twitter, and on various web portals (for example, Amazon.com or Youtube.com). Having an application that can identify and extract opinions from this huge amount of data, and provide an estimation of users' preferences about a manufactured good is of great interest to companies. The same interest is also encountered in politics in terms of candidate elections or for a government that wants to introduce new regulations, in order to have a glimpse on people opinions about the organization and their acts.

Natural Language Processing techniques may be applied for extracting sentiments and opinions in two ways: based on lexicons and through machine learning techniques. Both approaches have drawbacks. The first category is based on the polarity of sets of specific words (for example, the word "like" expresses a positive sentiment), but fails when modifiers are used, which may be at a distance in text (for example, negations:

C. Dichev and G. Agre (Eds.): AIMSA 2016, LNAI 9883, pp. 352–355, 2016.
DOI: 10.1007/978-3-319-44748-3_35

"I don't agree that from my previous post you can infer that I like the new phone launched on the market"). For the second category, supervised machine learning approaches (for example, Naïve Bayes, Maximum Entropy and Support Vector Machines) are used to classify texts into positive and negative opinions [3]. Their disadvantage is the need of human annotation for large volumes of training examples [4]. In terms of structure, the paper continues with details on the performed experiment. The third section presents the obtained results and the last section is centered on conclusions.

2 Details of the Experiment

This study was performed in the context of the RAGE H2020 EC project (http://rageproject.eu/), which focuses on serious games for e-learning. Our corpus consists of 9750 game reviews from 44 games, all written in English language and extracted from Amazon.com using crawl4j. The reviews were ranked on a Likert scale from 1 to 5 and were considered relevant if they contained more than 50 content words. The reviews were used to develop component scores from which to determine differences in positive, neutral, and negative game reviews.

Various vectors or word lists covering both general meaning and particular linguistic traits were combined with a Principal Component Analysis in order to determine the latent variables that define the specificities of gamer reviews following the techniques reported in Crossley et al. [5]. To develop our component scores, the following word categories were selected from the *General Inquirer* (GI, http://www.wjh.harvard.edu/~inquirer/homecat.htm) [6]: words referring to role, words indicating overstatement, words reflecting a sociological perspective, words expressing arousal, and general references to humans.

From the *Laswell* dictionary [7], we extracted words talking about skills, respect, power, wealth and gain. From *SenticNet* [8, 9], we selected words used to describe feelings based on four dimensions: attention, sensitivity, aptitude and pleasantness. *GALC* (Geneva Affect Label Coder, [10]) was used for selecting specific word categories about emotions, which were split into: boredom, anger, depression, amusement, admiration, positive, and negative. Word features related to arousal, dominance and affective variables were selected from ANEW (Affective Norms for English Words, [11]).

In addition, word lists that incorporated affective, perceptual, and cognitive processes, as well as personal concerns and relativity were extracted from the *Linguistic Inquiry and Word Count* (LIWC, [12]). The Hu-Liu polarity lists containing 2.000 positive and 4.500 negative words [13], and the Stanford Core NLP [14] sentiment analysis model based on recursive deep networks were also integrated. Only lemmas of content words were considered and multiple indices were computed in order to express the linguistic coverage of each word list for a given review.

3 Results

Eight affective components were identified using a Principal Component Analysis (PCA) which explained 51.22 % of the variance in the selected game reviews. The derived components were related to:

- *Negative Emotions:* the most powerful component, contains words with negative loadings including user frustration or game mechanics that are not working well;
- *Relations and Power:* includes words about interpersonal relationships (including relations between game characters or with other human players in multiplayer sessions), descriptions of actions, gameplay and achievements;
- *Positive Emotions:* reflects positive loadings and emotions, i.e., General Positive Words, GI Positive or GI Virtue;
- *Activities and Skills:* refers to actions within the game, as well as activities and their characteristics, i.e., GI Expressivity and LIWC Leisure activity;
- *Motivation:* reflects the overall impression induced by the game with regards to trust, surprise, attention;
- *Human and Roles:* depicts human functions (e.g., leader or authority) from reviews debating about characters with specific roles (e.g., commanders) or multiplayer modes (e.g., "Call of Duty");
- *Communication:* includes words that present ways and types of communication from GI Human and GI Role lists;
- *Ambiguous and Passive Language:* contains words with no active meaning (e.g., "admire", "passive", or "fell").

These components were used in a multi-variate analysis of variance (MANOVA) and six components yielded significant differences between the three classes of game reviews: *positive emotions* ($F = 1004.72$, $p < .01$, $\eta^2 = .171$), *negative emotions* ($F = 272.10$, $p < .01$, $\eta^2 = .053$), *relations and power* ($F = 39.22$, $p < .01$, $\eta^2 = .008$), *activities and skills* ($F = 32.13$, $p < .01$, $\eta^2 = .007$), *human and roles* ($F = 9.01$, $p < .01$, $\eta^2 = .002$), *ambiguous and passive language* ($F = 3.67$, $p - .025$, $\eta^2 = .001$). A stepwise discriminant function analysis using these six variables retained the first 5 variables and correctly allocated 5,371 of the 9,750 game reviews in the total set, $\chi 2$ ($df = 4$, $n = 9{,}750$) = 51.750, $p < .001$, for an accuracy of 55.1 % (the chance level for this analysis is 33.3 %).

4 Conclusions

Opinion mining and sentiment analysis are of great interest nowadays in many domains of economy, commerce and society. Natural Language Processing techniques can be used to provide useful insights; however, there are limitations. Up to date, only a few studies exist that focus on gaming, despite its huge popularity among people of all ages.

The research described in this paper presents a linguistic analysis centered on extracting language traits used by gamers when expressing opinions about game quality. The PCA analysis explained more than 50 % of the variance in language across all game reviews, while the DFA classification highlighted promising insights into game quality.

Acknowledgement. The work presented in this paper was partially funded by the EC H2020 project RAGE (Realising and Applied Gaming Eco-System) http://www.rageproject.eu/ Grant agreement No 644187.

References

1. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, San Rafael (2012)
2. Jurafsky, D., Martin, J.H.: An Introduction to Natural Language Processing. Computational Linguistics, and Speech Recognition. Pearson Prentice Hall, London (2009)
3. Pak, A., Paroubek, P.: Twitter as a corpus for sentiment analysis and opinion mining. In: LREC 2010, Valletta, Malta (2010)
4. Cheng, O.K.M., Lau, R.Y.K.: Probabilistic language modelling for context-sensitive opinion mining. Sci. J. Inf. Eng. **5**(5), 150–154 (2015)
5. Crossley, S., Kyle, K., McNamara, D.S.: Sentiment Analysis and Social Cognition Engine (SEANCE): An Automatic Tool for Sentiment, Social Cognition, and Social Order Analysis. Behavior Research Methods (in press)
6. Stone, P., Dunphy, D.C., Smith, M.S., Ogilvie, D.M.: The General Inquirer: A Computer Approach to Content Analysis. The MIT Press, Cambridge (1966)
7. Lasswell, H.D., Namenwirth, J.Z.: The Lasswell Value Dictionary. Yale University Press, New Haven (1969)
8. Cambria, E., Grassi, M., Poria, S., Hussain, A.: Sentic computing for social media analysis, representation, and retrieval. In: Ramzan, N., Zwol, R., Lee, J.S., Clüver, K., Hua, X.S. (eds.) Social Media Retrieval, pp. 191–215. Springer, New York (2013)
9. Cambria, E., Schuller, B., Xia, Y.Q., Havasi, C.: New avenues in opinion mining and sentiment analysis. IEEE Intell. Syst. **28**(2), 15–21 (2013)
10. Scherer, K.R.: What are emotions? And how can they be measured? Soc. sci. Inf. **44**(4), 695–729 (2005)
11. Bradley, M.M., Lang, P.J.: Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings. The Center for Research in Psychophysiology, University of Florida, Gainesville (1999)
12. Pennebaker, J.W., Booth, R.J., Francis, M.E.: Linguistic inquiry and word count: LIWC [Computer software] (2007)
13. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004). ACM, Seattle (2004)
14. Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.P.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Conference on Empirical Methods in Natural Language Processing (EMNLP 2013). ACL, Seattle (2013)