4th International Conference on Advanced Research Methods and Analytics (CARMA2022)
Universitat Politècnica de València, València, 2022
DOI: http://dx.doi.org/10.4995/CARMA2022.2022.15040

News versus Corporate Reputation: Measuring through Sentiment and financial analysis

Naiara Pikatza-Gorrotxategi, Izaskun Alvarez-Meaza, Rosa María Río-Belver, Enara Zarrabeitia-Bilbao
Universidad del País Vasco UPV-EHU, Spain.

Abstract

Today's companies cannot overlook their reputation if they want to continue to survive. One way to measure that reputation is through two factors: sentiment analysis of news stories in the press about those companies and the financial data of those companies. In this research, the sentiment analysis of news stories about several Euro Stoxx 50 companies for the years 2016 and 2019 has been carried out. For this purpose, the lexicon-based tools VADER and Hu Liu have been used. Then the trends of the results obtained for this four-year period have been analyzed and compared with the trends in their operating results in the same time period. The results obtained indicate that there is a high correlation between the sentiments reflected in the news and their operating results, i.e., when news sentiment about a company improves, its reputation also improves, and this causes its sales to increase. The same is true in the opposite direction.

Keywords: *Sentiment Analysis; Corporate Reputation; VADER SA tool, Hu Liu SA tool*

1. Introduction

In today's society, companies have increasingly more data at their disposal. This data may contain strategic information for companies, however, it is so voluminous that it is not easy to analyse it in the traditional way, making the use of artificial intelligence and data mining indispensable (Agarwal, 2020). In this context, sentiment analysis is a sub-discipline that falls under the umbrella of data mining and computational semantics. According to Gilbert and Hutto (Hutto & Gilbert, 2014), sentiment analysis, or opinion mining, is an active area of study in the field of Natural Language Processing (NLP) that analyses people's opinions, feelings, evaluations, attitudes and emotions by computationally processing subjectivity in text. It refers to the understanding of collected data obtained from sentiment-rich sources such as news, social media sites, reviews, etc. (Agarwal, 2020). Therefore sentiment analysis is concerned with extracting sentiment, opinions and emotions from text (Ravi & Ravi, 2015) and has applications in a wide range of domains, from customer satisfaction to political opinions (Medhat et al., 2014) (Mäntylä et al., 2018) (Ravi & Ravi, 2015).

Another aspect that companies cannot overlook is their reputation, as it affects, among other factors, consumer satisfaction. (Chun, 2005). According to Raithel in his article "The value-relevance of corporate reputation during the financial crisis" (Raithel et al., 2010), corporate reputation can be measured through 2 indicators: the sympathy felt towards the company, and the competence of that company. Consumers, when deciding on a company's reputation, rely on data received through word of mouth, news, advertising, etc. (Kossovsky, 2012). Therefore, one way to measure this sympathy for the company can be through sentiment analysis of news stories about those corporations. If the consumer perceives that the news has a positive tone about the companies, they will have more sympathy towards them and the company's reputation will increase.

The Python tool, VADER (Valence Aware Dictionary and Sentiment Reasoner), a sentiment analysis framework, uses a lexicon-based approach to determine the sentiment values of a sentence. This is used in conjunction with sentiment values explicitly assigned to keywords commonly found among news headlines, or in individual emails (Agarwal, 2020) (Borg & Boldt, 2020). This sentiment extraction typically results in a score that can be translated into positive, neutral or negative (Hutto & Gilbert, 2014). Another frequently used instrument, the Hu & Liu lexicon, was developed for sentiment analysis of customer reviews. The resulting categories (lexicon-based) are Sentiment (an overall measure of positivity), Positive and Negative (good classification metrics in machine learning tasks). This tool has been chosen because it has almost exclusively been used in studies that do not focus on textual production in the social media. (Mayor & Bietti, 2021)
A substantial number of sentiment analysis approaches rely greatly on an underlying sentiment (or opinion) lexicon. A sentiment lexicon is a list of lexical features (e.g., words) which are generally labeled according to their semantic orientation as either positive or

negative (B. Liu, 2010). With Hu Liu, words are categorized into binary classes (i.e., either positive or negative) according to their context free semantic orientation. (Hutto & Gilbert, 2014). Hu and Liu present a natural language-based approach for providing feature-based summaries of customer reviews. The approach uses a part-of-speech tagger to divide words into lexical categories, as only the semantic orientation of adjectives is considered by the algorithm. The use of different instruments for the automatic coding of the same dataset is essential to assess the robustness of results across tools (Mayor & Bietti, 2021). As there are 2 suitable tools, this research will measure the reputation of companies through sentiment analysis, measured with VADER and Hu Liu. Therefore the aim of the article is to analyse the possible correlation between the sentiment analysis of news about companies and their reputation.

2. Methodology

The methodology followed to obtain and analyze the data was as follows:

STEP 1.- Choice of database: The objective is to carry out a sentiment analysis of the news on the 10 highest dividend yielding companies in the Euro Stoxx 50 as of May 2021. This database was chosen because of consistent data for these companies. These companies are: Axa, Eni, Total Energies, Intesa Sanpaolo, ING, Engie, BNP Paribas, Basf, Allianz and Daimler (*El 26% Del Euro Stoxx 50 Paga Una Rentabilidad Por Dividendo Superior Al 4% | Mercados | Cinco Días*, n.d.).

STEP 2.- Data extraction: having selected the companies, their most relevant news items according to different databases were downloaded. To do so, we went to the original source and downloaded the 500 most relevant news items by company and year from the main media. It was decided to analyse the years 2016 and 2019. In the event that a company did not reach 500 news items per year, all of them were downloaded. The total number of news items per company per year is as follows:

Table 1. Extracted news items per year.

Year	2016	2019	TOTAL
News Items	3,994	3,838	**7,832**

For each of the selected companies 1,000 news items (500 per year) have been extracted, with 3 exceptions: *Intesa San Paolo* had only 482 news items in total, ING 250 news items in total, and Total Energies, due to its numerous name changes over the years, produced very little news, so it is not counted. Therefore the total number of news items is 7,832.

STEP 3.- Cleaning and classification of extracted data for SA: Having downloaded all of the news items in txt format, they are imported into the data mining software Vantage Point (W. Liu & Liao, 2017), through which the raw data can be structured for subsequent export in xlx-csv format.

STEP 4.- Conducting Sentiment Analysis: The news items are ready to be exported to Orange, a machine learning and data mining suite for data analysis through Python scripting (Demšar et al., 2013). Now the sentiment analysis of each of the news items will be carried out using the VADER and Hu Liu tools.

STEP 5.- Analyzing the correlation between the Sentiment Analysis of the news and the operating results, by company: The possible correlation between the trend between the Sentiment Analysis with VADER and Hu Liu and the operating profits of each company is analyzed.

3. Results

Once the Sentiment Analysis of the extracted news has been carried out, the results obtained, classified by company and tool used, are as follows:

Table 2. Results of the Sentiment Analysis of the news

	VADER		**HU LIU**	
	2016	**2019**	**2016**	**2019**
AXA	0.5534	0.6067	0.5560	0.9739
ENI	0.4778	0.2920	0.0806	0.0826
INTESA SANPAOLO	0.2188	0.2491	-0.6068	-0.0290
ING	0.7067	0.4582	0.0721	-0.0911
ENGIE	0.6582	0.668075	0.4966	1.063754
BNP Paribas	0.2944	0.226955	-0.2662	-0.5982
BASF	0.7113	0.4003	0.3850	0.1311
ALLIANZ	0.5568	0.5682	0.07709	0.1541
DAIMLER	0.7358	0.5607	1.3256	1.03161

One way to study the data is to analyze their trend over time, and see whether they are improving or worsening. In this way it will be possible to check the trend of the sentiments and opinions reflected in the news about each company, and, since the evolution of sympathy towards these companies is being measured, to analyze whether its reputation could improve or not. An improvement in a company's reputation will, in principle, lead to an increase in sales. Consequently, the evolution in the sentiment analysis has been compared with the evolution in the operating result of each company. The data obtained are as follows in Table 3:

Table 3. Comparison of trends in news Sentiment Analysis and operating profit, by company

	VADER	HU LIU	PROFIT	VADER	HU LIU	PROFIT
AXA				**ENI**		
2016	0.5534	0.5560	7,641	0.4778	0.0806	2,315
2019	0.6067	0.9739	8,427	0.2920	0.0802	8,597
TREND	UP	UP	UP	DOWN	DOWN	UP
INTESA				**ING**		
2016	0.2188	-0.606	8,273	0.7067	0.0721	5,903
2019	0.24914	-0.029	8,760	0.4582	-0.0911	6,834
TREND	UP	UP	UP	DOWN	DOWN	UP
ENGIE				**BNP**		
2016	0.6582	0.4966	9,491	0.2944	-0.2662	10,771
2019	0.6680	1.0637	10,366	0.2269	-0.5982	10,057
TREND	UP	UP	UP	DOWN	DOWN	DOWN
BASF				**ALLIANZ**		
2016	0.7113	0.3851	5,330	0.5568	0.0770	11,056
2019	0.4003	0.1311	4,631	0.5682	0.1541	11,855
TREND	DOWN	DOWN	DOWN	UP	UP	UP
DAIMLER						
2016	0.7358	1.3256	31,963			
2019	0.5607	1.0316	29,165			
TREND	DOWN	DOWN	DOWN			

**Operating profit is shown in millions of euros*

Having obtained the data on the sympathy generated by the companies, it is time to measure the reputation of these companies. In order to measure the reputation trend of companies, two factors should be taken into account, which are also intercorrelated; the likeability that these companies induce and their financial consistency (Raithel et al., 2010). One way to measure this sympathy can be through sentiment analysis of press releases. If those sentiment analyses improve, that will mean an improvement in the company's reputation. This will lead to an improvement in sales and therefore in the operating profit. In turn, an improvement in its financial results will cause the company's reputation to improve, completing the cycle. Figure 1 shows this correlation between the trend in sympathy towards the company and the trend in operating income.

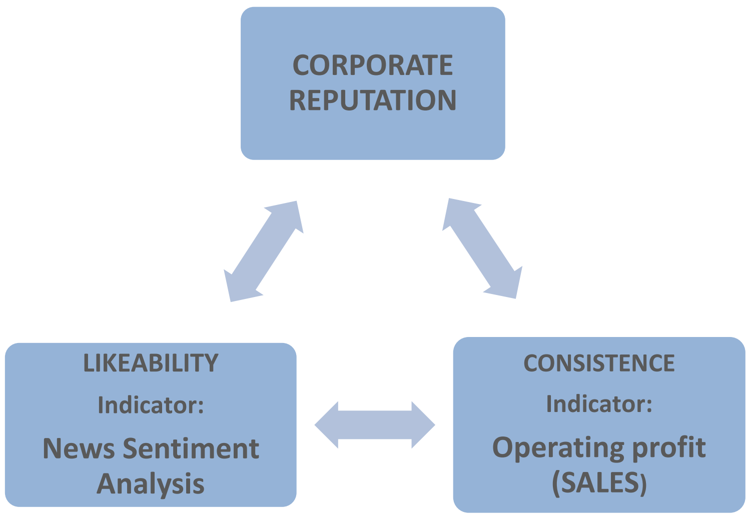

Figure 1: Corporate reputation trend measurement model

4. Conclusions

Based on the obtained data, it can be concluded that the initial thesis is correct. On the one hand, it can be seen that the VADER and Hu Liu data coincide in terms of trend. If we analyze the trend between 2016 and 2019, the trends between these two tools coincide in all cases. In 4 of the cases the trend in the sentiment analysis of the 2 tools is upward with both VADER and Hu Liu (Axa, Intesa, Engie and Allianz), and in the other 5 companies the trend is downward (Eni, ING, Bnp, Basf and Daimler). This data may be an indicator that the 2 tools coincide in their sentiment analysis measurements. If we compare these sentiment analysis trends with the trend in operating results over the same time period, we can see that they also coincide in almost all cases, i.e., companies that have had a positive trend in their news sentiment analysis increase their operating results and vice versa. This occurs in all cases except for Eni and ING, which increase their profits within that period but lower their scores in news sentiment analysis. The reason for this discordance in the data in the case of ING may be the low number of news items analyzed with respect to the other companies (250 news items in the case of ING, and 1,000 news items in the others). In any case, the correlation between sentiment analysis and operating results is positive in 78% of the cases. Therefore, when sentiment analysis shows a positive trend, operating results increase, i.e., sales of that product increase. When the trend is negative, sales decrease.

References

Agarwal, A. (2020). Sentiment Analysis of Financial News. *Proceedings - 2020 12th International Conference on Computational Intelligence and Communication Networks, CICN 2020*, 312–315. https://doi.org/10.1109/CICN49253.2020.9242579

Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, *162*, 113746. https://doi.org/10.1016/J.ESWA.2020.113746

Chun, R. (2005). Corporate reputation: Meaning and measurement. *International Journal of Management Reviews*, *7*(2), 91–109. https://doi.org/10.1111/j.1468-2370.2005.00109.x

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., & Zupan, B. (2013). Orange: Data mining toolbox in python. *Journal of Machine Learning Research*, *14*(August), 2349–2353.

El 26% del Euro Stoxx 50 paga una rentabilidad por dividendo superior al 4% | Mercados | Cinco Días. (n.d.). Retrieved December 2, 2021, from https://cincodias.elpais.com/cincodias/2021/05/18/mercados/1621332461_948065.html

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, 216–225.

Kossovsky, N. (2012). Reputation, stock price, and you: Why the market rewards some companies and punishes others. In *Reputation, Stock Price, and You: Why the Market Rewards Some Companies and Punishes Others* (Vol. 9781430248). https://doi.org/10.1007/978-1-4302-4891-0

Liu, B. (2010). Sentiment analysis and subjectivity. In *Handbook of Natural Language Processing, Second Edition* (pp. 627–666).

Liu, W., & Liao, H. (2017). A Bibliometric Analysis of Fuzzy Decision Research During 1970–2015. *International Journal of Fuzzy Systems*, *19*(1). https://doi.org/10.1007/s40815-016-0272-z

Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, *27*, 16–32. https://doi.org/10.1016/j.cosrev.2017.10.002

Mayor, E., & Bietti, L. M. (2021). Twitter, time and emotions. *Royal Society Open Science*, *8*(5). https://doi.org/10.1098/rsos.201900

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Raithel, S., Wilczynski, P., Schloderer, M. P., & Schwaiger, M. (2010). The value/relevance of corporate reputation during the financial crisis. *Journal of Product & Brand Management*, *19*(6), 389–400. https://doi.org/10.1108/10610421011085703

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, *89*, 14–46. https://doi.org/10.1016/j.knosys.2015.06.015