

Sentiment analysis of Portuguese Political Parties Communication

Carlos J. Costa

ADVANCE/CSG, ISEG (Lisbon School of Economics & Management), Universidade de Lisboa, Portugal

cjcosta@iseg.ulisboa.pt

Manuela Aparicio

NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal

manuela.aparicio@novaims.unl.pt

Joao Tiago Aparicio

INESC-ID and Instituto Superior Tecnico/Universidade de Lisboa

Dep Transportes LNEC, Portugal

joao.aparicio@tecnico.ulisboa.pt

This is the author accepted manuscript version of the paper published by ACM as:

Costa, C. J., Aparicio, M., & Aparicio, J. (2021). Sentiment Analysis of Portuguese Political Parties Communication. In Proceedings of the 39th ACM International Conference on the Design of Communication (SIGDOC '21) (pp. 63-69). Association for Computing Machinery (ACM): New York, NY, USA. ISBN: 978-1-4503-8628-9. DOI: <https://doi.org/10.1145/3472714.3473624>

Funding:

We are gratefully acknowledged financial support from FCT- Fundação para a Ciência e Tecnologia, I.P., Portugal, national funding through research grants UIDB/04521/2021 - Advance/CSG, ISEG, Universidade de Lisboa and UIDB/04152/2021 - Centro de Investigação em Gestão de Informação (MagIC).

Sentiment analysis of Portuguese Political Parties Communication

Carlos J. Costa

ADVANCE/CSG, ISEG (Lisbon School of Economics & Management), Universidade de Lisboa, Portugal
cjcosta@iseg.ulisboa.pt

Manuela Aparicio

NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal
manuela.aparicio@novaims.unl.pt

Joao Tiago Aparicio

INESC-ID and Instituto Superior Tecnico/Universidade de Lisboa
Dep Transportes LNEC, Portugal
joao.aparicio@tecnico.ulisboa.pt

ABSTRACT

Political communication in social media has gained increasing importance in the last years. In this study, we analyze the political parties' communication on Twitter and understand the sentiment of their communication. First by identifying their communication performance regarding the daily number of tweets, favourite tweets, number of retweets per day and per political party. We present a sentiment analysis by the political party using tweets data. In this study, we propose an explanatory model with the main drivers of retweets. To conduct this study, our approach used data analysis and machine learning techniques methods. Results indicate the main determinants that influence future retweets of political posts globally. Here we present a comparison of the communication content between tweets posts and the political parties' programs available on their institutional websites. We identify the similarities between tweets and formal programs per party and among all parties. Globally contribute to analyze the coherence and effectiveness of the political parties' communication.

CCS CONCEPTS

•General and reference •Cross-computing tools and techniques •Empirical studies; •Information systems • Information retrieval • Retrieval tasks and goals •Sentiment analysis; •Computing methodologies •Artificial intelligence • Natural language processing; •Computing methodologies • Machine learning

KEYWORDS

Twitter; Political Parties; Sentiment Analysis; Document Similarity; Machine Learning

Author accepted paper version, please cite as follows:

Costa, C., Aparicio, M., & Aparicio, J. (2021). Sentiment Analysis of Portuguese Political Parties Communication. *The 39th ACM International Conference on Design of Communication, (SIGDOC '21), October 12-14, 2021*. ACM, New York, NY, USA, p.63–69.
<https://doi.org/10.1145/3472714.3473624>

1. Introduction

Political parties are increasingly using social media to communicate their values and ideas. An example of such behaviour was when Barack Obama's staff successfully used Twitter in the 2008 presidential elections [13]. This practice persisted in Donald Trump's administration in 2016 [14]. However, using social media is not a panacea per se. It is also essential to analyze what is the effectiveness of what is being communicated to the public. The cohesion between parties' political agenda versus their social media communication needs to be addressed. Are parties communicating according to their goals, or is it all part of a hype machine?[16] The usage of social media in a political context has been studied by several authors [6][7][13][14]. In this context, the evolution in natural language processing and sentiment analysis is significant. Nevertheless, there is a research gap in the Portuguese language and its application in a political context as well [15]. The purpose of the work performed in this paper is to analyze political parties' communication, expressed explicitly by the official Twitter accounts of such parties. To reach this main goal, we state four research objectives (RO) as follows: RO1: Identify the performance of each political party on Twitter; RO2: Identify the global sentiment per political party in Twitter communication; RO3: Identify the drivers of retweet behaviour in political parties; RO4: Understand the similarities between social media communication and political program communication.

We used data analysis and machine learning techniques described in section 3 of this paper for each of these research objectives. This study contributes to a better understanding of how political parties communication in Twitter can be analyzed in terms of the sentiment of posts, and globally we contribute to analyze the coherence and effectiveness of the political parties communication. We also used also conducted our approach following a data science perspective [23][24].

2. Literature Review

Social media analysis has been a study in the context of design of communication [22][25][26]. In this context, many approaches were used[22][25]. Social media usage has been studied to identify the performance of each political party in Twitter [17]. Twitter is also a relevant data source to identify the global sentiment per political party in Twitter communication [18]. Sentiment analysis refers to using several approaches, like natural language processing, text analysis, computational linguistics, and biometrics, to systematically identify, extract, quantify, and study affective states and subjective information. Emotions can be reactions to internal stimuli (such as thoughts or memories) or events in our environment. To perform sentiment classification, we may use a previously defined lexicon. It involves typically a very extensive work of manual classification and validation. It is also possible to use model training. In this case, it is needed to have texts previously labelled. To analyze emotions, some studies proposed lexicons [1][2][3][4][27]. Hutto & Gilbert [5] proposed a rule-based model for sentiment analysis named VADER, using a combination of qualitative and quantitative measures to classify positive, negative, and neutral emotions extracted from text.

Prediction of retweet behaviour is especially relevant. Specifically, it may substantially impact the diffusion of political ideology and programs in the context of political communication. Several researchers identified possible strategies to select features [19][20] Some researcher even analyzed the relation between sentiment analysis and tweet and retweet [21].

Tweets of different political parties may have high variability in what concerns the sentiment expressed [27]. However, the sentiments may concern different subjects. So, a question may arise: are the political parties communicating about the same subjects? Along with this, it is important to understand the cohesion of what it is communicated and the actual proposals of the parties. This is why it is important to understand the similarities between social media communication and political programs.

To measure text similarity with different characteristics, it is important to apply different natural language processing methods before applying the similarity measure. This includes removing stop words, stemming, lemmatization, emoji removal, among others[27]. According to previous research, to calculate the similarity of documents, embedding-based and text-based similarity measures have a higher performance [8][9]. However, methods that use bag-of-words such as the TF IDF has had better performance than context-sensitive methods such as Word2Vec and BERT.

3. Methodology

To accomplish our research objectives, we conducted different methods, which are summarized in the following table (Table 1).

Table 1. Methodological approach

Research objective	Method/ Technique
RO1: Identify the performance of each political party on Twitter	Descriptive statistics
RO2: Identify the global sentiment per political party in Twitter communication	NLP (Natural Language Processing) Sentiment analysis
RO3: Identify the drivers of retweet behaviour in political parties	ML (Machine Learning) Supervised Learning; Regression analysis
RO4: Understand the similarities between social media communication and political program communication	NLP (Natural Language Processing) Document similarity

We collected tweets from the official accounts of political parties. Twitter corresponds to the last 3200 tweets of each political party corresponding to the pandemic value.

We started by translating the tweets from Portuguese to English. Then, we used the Vader library from NLTK [5]. Vader allows classifying the tweets into positive and negative. Each tweet has a value for positive and a value for negative. We obtained the average of each day's tweets for each political party to compare the political parties. We also analyzed the average and standard deviation of sentiment analysis for each tweeter account.

We also performed a machine learning approach, supervised learning, using regression analysis to understand (and predicting) the retweeting behaviour. Several algorithms were tested in this context.

To understand the cohesion between the content being communicated and the effective goals of each party, we extracted the electoral programs from each party official website. Then extracted the text in those documents. Note that the corpora can be defined as both the text in the documents and the tweets in the context of this study. We also performed a text processing and applied document similarity to compare the contents of the electoral programs with communication using the tweeter. For this analysis, stop words along with emoji were removed from the corpora. We took this liberty since the emoji content is not relevant for policy proposals available in the programs. The whole corpora were tokenized and lemmatized. The similarity of the documents is calculated based on pairwise similarity using one of the most used bag-of-words similarity metric, tf-idf., which can be defined as:

$$tfidf(t,d)=tf(t,d)\cdot idf(t), \text{ with } idf(t)=\log n/(df(t)+1)$$

4. Study Context

Democratic participation is generally associated with ancient Greeks. In this period, the citizens participated directly in the life of the polis. But assemblies of representatives are rooted in the Middle Ages. In Portugal, those parliaments were called Cortes. In the beginning, it only included representatives of the nobility and clergy. The first assemblies of representatives are the legendary Cortes of Lamego (1143). In 1211, the Cortes of Coimbra established the first general laws (*Leis Gerais do Reino*). In 1254, the Cortes of Leiria also included representatives of the municipalities. The Cortes have met on an irregular basis. Typically, Cortes have a more consultative than deliberative function. Cortes were also crucial in the confirmation of the king. In the XIX century, the power of Parliament was increased and the level of representativity. It was Census suffrage, but already with several political parties.

In what concerns the house of Parliament (*Palacio de São Bento*), it is in the place of a Benedictine monastery. In 1820, there was a Liberal Revolution. As a consequence, in 1835, the religious orders were suppressed, and the monks were expelled from the monastery. Portuguese Parliament was installed in the building, then called Palácio das Cortes. In 1910, the republic was established. The Parliament typically elected the government. The representativity was also enlarged. However, universal suffrage was established only after the 1974 revolution.

Today, the main political parties represented at the Portuguese Parliament are PS, PSD, BE, CDU, PAN, CDS, IL, Chega and Livre (Figure 1). PS (*Partido Socialista*) is a socialist and social democrat party and belongs to the Party of European Socialists. PSD (*Partido Social Democrata*) belongs to the European People's Party. Those two parties are ideologically very similar. For example, the PSD includes an extensive range of political sensibilities, from social democrats to liberal conservatives and Christian democrats. BE and CDU belong to the European United Left. CDU is a collision of the PCP (*Partido Comunista Português*) and Verdes (Green Party), and they represent the traditional communist party. BE (*Bloco de Esquerda*) merged several minor political parties (mainly Trotskyist and extreme left). PAN (*Pessoas Animais e Natureza*) is a green party belonging to the Green European group. CDS-PP is a Christian Democratic party, and IL is

a liberal party, and it is a member of the Alliance of Liberals and Democrats for Europe Party. Chega is a nationalist and right-wing party.

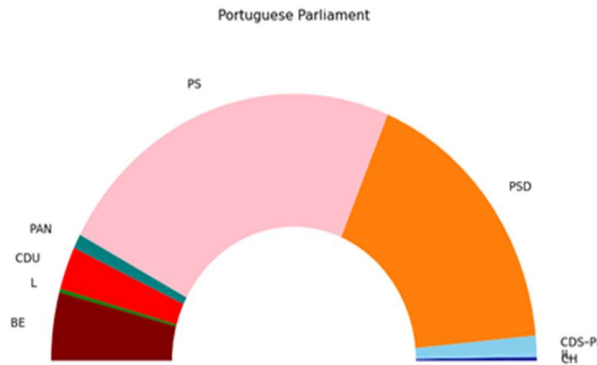


Figure 1 Portuguese Parliament Composition

We analyzed the period between the beginning of October and February. During this period, there are several political moments: presidential elections, the third wave of covid, particularly violent. PS (socialist party) is the government party. PCP and BE. PSD, CDS support PS government, IL and Chega are in the opposition.

5. Results

5.1 Performance of each political party on Twitter

Activity in Twitter shows that BE (*Bloco de Esquerda*) leads to the number of tweets per day. But the impact of IL, PCP and Chega is also significant. As long as having fewer tweets, each tweet has more retweets and favourites by tweet. In the case of IL, the number of retweets and favourites per day is even more prominent, although having fewer tweets.

Each tweeter was classified using Vader. The result presents positive scores, negative scores, neutral scores and compound scores for each tweet. Then, an average per day was calculated. Figure 2 represents the political parties performance results in Twitter, concerning the number of tweets per day, number of favorites per day and number of retweets per day.

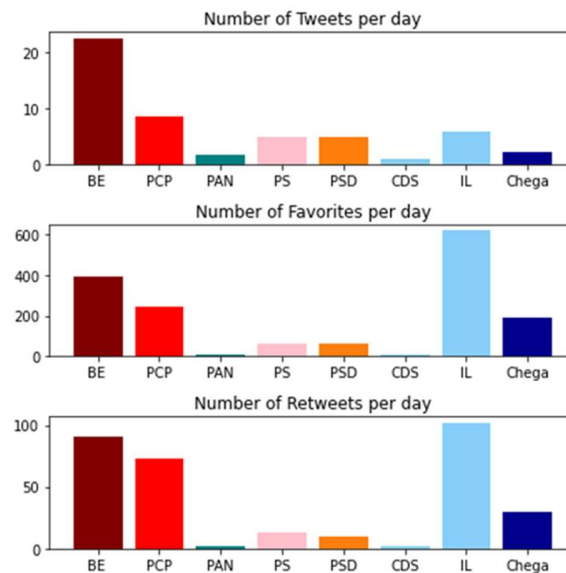


Figure 2: Political parties performance results in Twitter

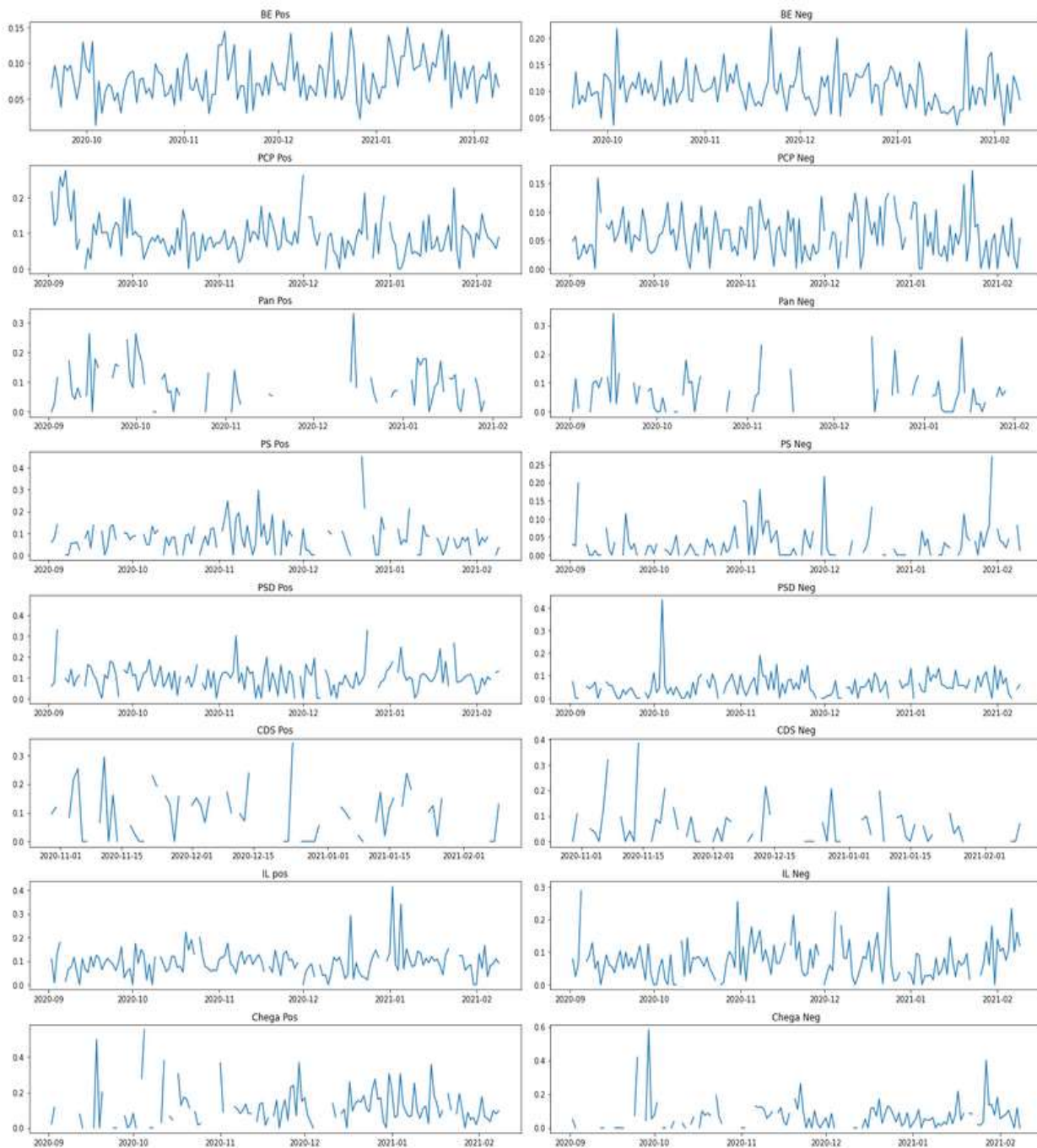


Figure 3: Evolution of positive (Pos) and negative (Neg) sentiment for each political party from September 2020 until February 2021

5.2 Sentiment analysis by political party

Figure 3 **Error! Reference source not found.** shows the average of the positive and negative scores per day for the tweet. The analysis allows to verify that there is a vital variability, but there were no evident trends.

The chart of the following figures (Figure 4 and Figure 5) summarize information related to the average and variability that we expected by observing the previous charts.

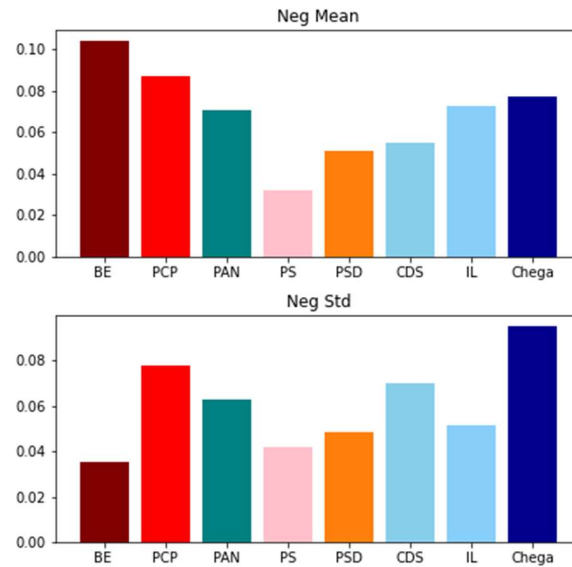


Figure 4: Negative sentiment per political party (Mean and Standard Deviation)

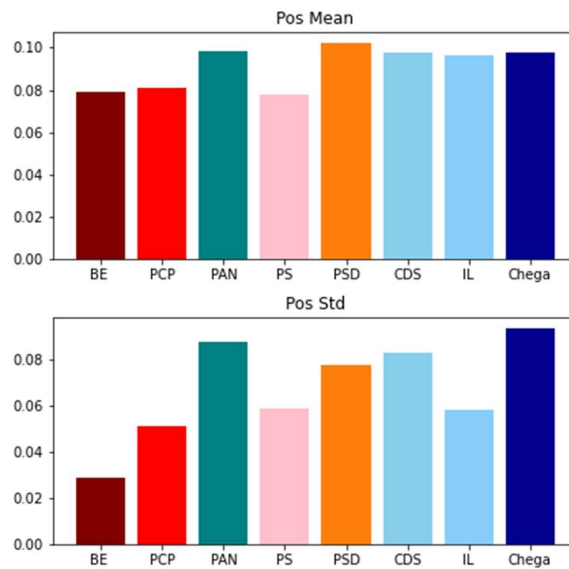


Figure 5: Positive sentiment per political party (Mean and Standard Deviation)

As it is expected, PS has a less negative sentiment. Quite surprisingly, PSD being at the opposition has the most positive sentiment. The right-wing is, on average, more positive than the left-wing. BE is the most consistent in what concern the variability of speech. The standard deviation of the BE speech is low either in the degree of positivity or negativity. Chega is the least consistent in what concern the variability of speech. The standard deviation of the speech is high either in the degree of positivity or negativity.

5.3 Explaining the Retweeting Behaviour

An extended study was also conducted. We need to understand the most influential variables in what concerns political parties retweet from the sentiment analysis data. We made a regression analysis upon which retweet is the dependent variable, and the independent variables are favourites (express the meaning of the same point of view in Twitter), negative sentiment (neg), neutral sentiment (neu), positive sentiment (pos), text length (len Tex); hashtags and mentions. We run the model, and results express that these variables explain 84% of the retweets. Results (Table 2 and Table 3) also clearly indicate that negative sentiment influences greatly next retweets about a political party tweet. Results also indicate that favourites negative sentiment, hashtags, text length, mentions are significantly influencing retweets. In the case of text length, it influences negatively, which means that the shorter the tweet communication, the most likely that tweet can be retweeted to the rest of the network. These results show that if many people favour a tweet, they are more likely to retweet it. If the tweet demonstrates a negative sentiment, a hashtag, or a mention, the readers also tend to retweet. These variables explain 84% of the retweet.

Table 2: Regression results

OLS Regression Results			
Dep. Variable:	Retweets	R-squared:	0.839
Model:	OLS	Adj. R-squared:	0.839
Method:	Least Squares	F-statistic:	1.763e+04
Date:	Sat, 15 May 2021	Prob (F-statistic):	0.00
Time:	19:38:02	Log-Likelihood:	-86715.
No. Observations:	23769	AIC:	1.734e+05
Df Residuals:	23761	BIC:	1.735e+05
Df Model:	7		
Covariance Type:	nonrobust		

Table 3: Regression parameters

	coef	std err	t	P> t	[0.025	0.975]
const	-1.4423	0.415	-3.473	0.001	-2.256	-0.628
Favorites	0.2296	0.001	349.796	0.000	0.228	0.231
neg	2.0056	0.685	2.930	0.003	0.664	3.347
neu	0.3320	0.435	0.762	0.446	-0.522	1.186
pos	-0.4631	0.594	-0.780	0.436	-1.627	0.701
lenTex	-0.0038	0.001	-4.718	0.000	-0.005	-0.002
Hashtags	0.2503	0.049	5.119	0.000	0.154	0.346
Mentions	0.3500	0.086	4.069	0.000	0.181	0.519
Omnibus:	47844.549	Durbin-Watson:	1.638			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	738581430.277			
Skew:	15.841	Prob(JB):	0.00			
Kurtosis:	865.992	Cond. No.	2.59e+03			

Table 4 demonstrated that except for the constant, the other independent variables do not show multicollinearity problems.

Table 4: Multicollinearity analysis

	Feature	VIF
0	const	47.443872
1	Favorites	1.013429
2	neg	1.521153
3	neu	1.967340
4	pos	1.677445
5	lenTex	1.111620
6	Hashtags	1.038763
7	Mentions	1.029231

Linear regression using OLS is allowed to understand the impact of each variable. But, to understand the model's predicting power, we split the sample into a train (80%) and test (20%). The following models were tested: OLS, Ridge with parameter $\alpha = 0.5$, Lasso with parameter $\alpha = 0.5$, BayesianRidge, Polinomia regression and MLP model. The OLS is the Ordinary least squares Linear Regression. The Ridge model solves a regression model where the loss function is the linear least-squares function, and the l_2 -norm gives regularization. Also known as Ridge Regression or Tikhonov regularization. The Lasso model is the Linear Model trained with l_1 prior as a regularizer. The BayesianRidge model implementation is based on the algorithm [11] where updates of the regularization parameters are done as suggested [10]. The MLP is a Multi-layer Perceptron model that optimizes the squared-loss using stochastic gradient descent, with an optimizer proposed by Kingma and Ba [12]. The MLP model was finetuned, and two hidden layers corresponded to the best solutions. Nevertheless, results (Fig. 6) showed that traditional models overperformed in relation to the neural networks approach.

Train Data				Test Data			
	R2	MAE	MSE		R2	MAE	MSE
OLS	0.852232	3.395771	82.361402	OLS	0.764957	3.444953	104.684972
Ridge	0.852232	3.395747	82.361403	Ridge	0.764956	3.444938	104.685139
Lasso	0.851731	3.371816	82.640604	Lasso	0.765343	3.421643	104.512759
BayesianRidge	0.852195	3.387667	82.381956	BayesianRidge	0.764878	3.438733	104.720064
Polynomial Regression	0.891842	2.669483	60.284191	Polynomial Regression	0.717263	2.814491	125.927009
Neural Network (MLP)	0.881496	2.808005	66.050665	Neural Network (MLP)	0.746654	2.942042	112.836870

Figure 6: Comparison between Models

5.4 Analyzing Similarities Between Social Media Communication and Political Program Communication

We have conducted a throughout the analysis to compare the content of all the political party programs and the tweets posts content of each party (September 2020 until February 2021). This content analysis compares the alignment between the political program and the parties' social media communication on Twitter. The following figure represents the percentage of similarity between the programs' content (be_p; cds_p; chega_p; liberal_p; pan_p; pcp_p; ps_p; psd_p) and the tweets for each party (be_t; cds_t; chega_t; liberal_t; pan_t; pcp_t; ps_t; psd_t) in a heat-map format. This analysis also allows comparing the similarity level among all the programs and tweets between political parties.

Results indicate that in the first place, IL is the one party that has more similar correspondence between the program and the tweets (Figure 7). In the second place, the party with more similarity between the program and the tweets communication is BE. The most similar tweets among all parties are BE, PAN, and Chega. The most different are Chega and PCP. These two parties are completely

opposite from a political point of view. PSD and PS political content program are 75% similar between the two. CDS is 75% similar to PSD and 65% with PS. The political program of BE is more similar to PSD program than with the PS program. The study results indicate that all the political party programs are alike, except one the IL (liberal) program, which is very different from other programs.

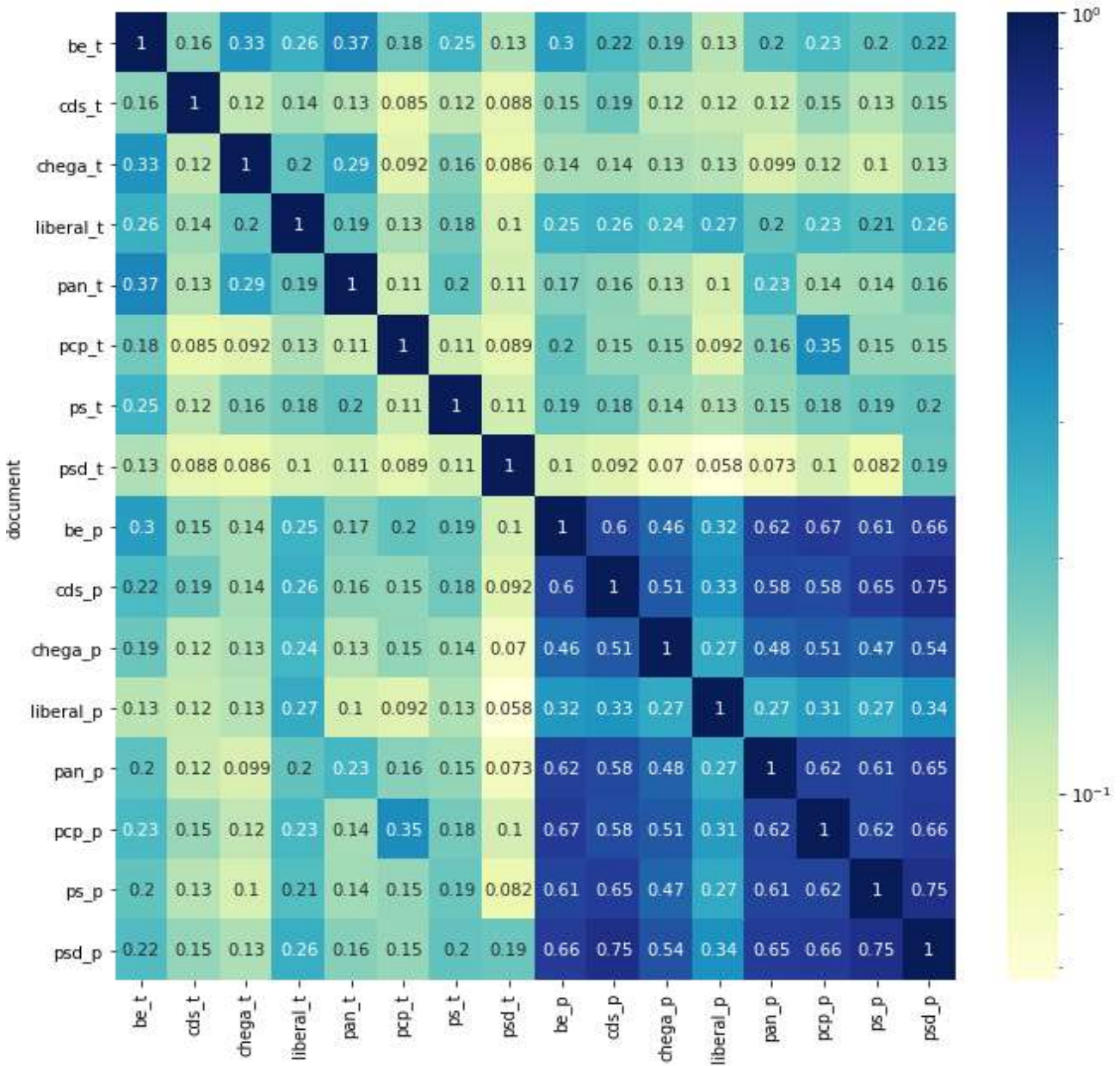


Figure 7: Similarities between political programs' documents, according to the topics

6. Conclusions

The analysis of political parties' communication, expressed explicitly by the official Twitters accounts, allows to verify some differences as expected. BE is the party that was more twitters, while CDS is the party with minor twitter. On the other hand, some strategic approaches may be evidenced. For example, the party that supports the government has a more positive sentiment. It is also essential to identify that favorites, negative sentiments, and hashtags are significant to explain retweet, while positive and neutral sentiments are not significant. Similarities of the programs between some political parties of a very different ideological matrix are interesting to verify. This should be subject to future research. On the other hand, our approach is also relevant as long as it contributes with an innovative way to analyze tweets.

ACKNOWLEDGMENTS

We are gratefully acknowledged financial support from FCT- Fundação para a Ciencia e Tecnologia, I.P., Portugal, national funding through research grants UIDB/04521/2021 - Advance/CSG, ISEG, Universidade de Lisboa and UIDB/04152/2021 - Centro de Investigação em Gestão de Informação (MagIC).

REFERENCES

- [1] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3–4, pp. 169–200, 1992.
- [2] D. A. Sauter, F. Eisner, P. Ekman, and S. K. Scott, "Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations," *Proceedings of the National Academy of Sciences*, vol. 107, no. 6, pp. 2408–2412, 2010.
- [3] R. Plutchik, "A general psychoevolutionary theory of emotion," in *Theories of emotion*, Elsevier, 1980, pp. 3–33.
- [4] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational intelligence*, vol. 29, no. 3, pp. 436–465, 2013.
- [5] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2014, vol. 8, no. 1.
- [6] S. Aral and D. Eckles, "Protecting elections from social media manipulation," *Science*, vol. 365, no. 6456, pp. 858–861, 2019.
- [7] L. M. Kruse, D. R. Norris, and J. R. Flinchum, "Social media as a public sphere? Politics on social media," *The Sociological Quarterly*, vol. 59, no. 1, pp. 62–84, 2018.
- [8] Mandal, K. Ghosh, S. Ghosh, and S. Mandal, "Unsupervised approaches for measuring textual similarity between legal court case reports," *Artificial Intelligence and Law*, pp. 1–35, 2021.
- [9] K. W. Boyack et al., "Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches," *PloS one*, vol. 6, no. 3, p. e18029, 2011.
- [10] D. J. MacKay, "Bayesian interpolation," *Neural computation*, vol. 4, no. 3, pp. 415–447, 1992.
- [11] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244, 2001.
- [12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [13] C. J. Vargo, L. Guo, M. McCombs, and D. L. Shaw, "Network issue agendas on Twitter during the 2012 US presidential election," *Journal of Communication*, vol. 64, no. 2, pp. 296–316, 2014.
- [14] B. Joyce and J. Deng, "Sentiment analysis of tweets for the 2016 US presidential election," in *2017 IEEE MIT Undergraduate Research Technology Conference (URTC)*, 2017, pp. 1–4.
- [15] D. A. Pereira, "A survey of sentiment analysis in the Portuguese language," *Artificial Intelligence Review*, vol. 54, no. 2, pp. 1087–1115, 2021.
- [16] S. Aral, *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health--and How We Must Adapt*. Currency, 2020.
- [17] M. D. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer, "Predicting the political alignment of twitter users," in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, 2011, pp. 192–199.

- [18] D. Hagar, “# vote4me: The impact of Twitter on municipal campaign success,” in Proceedings of the 2015 International Conference on Social Media & Society, 2015, pp. 1–7.
- [19] Q. Zhang, Y. Gong, J. Wu, H. Huang, and X. Huang, “Retweet prediction with attention-based deep neural network,” in Proceedings of the 25th ACM international on conference on information and knowledge management, 2016, pp. 75–84.
- [20] H.-K. Peng, J. Zhu, D. Piao, R. Yan, and Y. Zhang, “Retweet modeling using conditional random fields,” in 2011 IEEE 11th International Conference on Data Mining Workshops, 2011, pp. 336–343.
- [21] J. Chen, M. S. Hossain, and H. Zhang, “Analyzing the sentiment correlation between regular tweets and retweets,” *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–9, 2020.
- [22] M. Aparicio and C. J. Costa, “Collaborative systems: characteristics and features,” in Proceedings of the 30th ACM international conference on Design of communication, 2012, pp. 141–146.
- [23] S. Aparicio, J. T. Aparicio, and C. J. Costa, “Data Science and AI: trends analysis,” in 2019 14th Iberian Conference on Information Systems and Technologies (CISTI), 2019, pp. 1–6.
- [24] C. J. Costa and J. T. Aparicio, “POST-DS: A Methodology to Boost Data Science,” in 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), 2020, pp. 1–6.
- [25] C. J. Costa and M. Aparicio, “Social networks: intentions and usage,” in Proceedings of the 2013 International Conference on Information Systems and Design of Communication, 2013, pp. 101–107.
- [26] C. J. Costa, M. Aparicio, and A. S. Braga, ‘Design of communication: a review of theories and models’, New York, NY, USA, 2012, pp. 15–19. doi: 10.1145/2311917.2311921.
- [27] J. T. Aparicio, J. Sequeira, C. Costa. 2021. Emotion analysis of Portuguese Political Parties Communication over the covid-19Pandemic CISTI’2021 - 16th Iberian Conference on Information Systems and Technologies.