

# A Social Opinion Gold Standard for the Malta Government Budget 2018

**Keith Cortis**

Maynooth University, Maynooth, Ireland  
Malta Information Technology Agency, Malta  
keith.cortis.2019@mumail.ie

**Brian Davis**

ADAPT Centre,  
Dublin City University,  
Glasnevin, Dublin 9, Ireland  
brian.davis@adaptcentre.ie

## Abstract

We present a gold standard of annotated social opinion for the Malta Government Budget 2018. It consists of over 500 online posts in English and/or the Maltese less-resourced language, gathered from social media platforms, specifically, social networking services and newswires, which have been annotated with information about opinions expressed by the general public and other entities, in terms of sentiment polarity, emotion, sarcasm/irony, and negation. This dataset is a resource for opinion mining based on social data, within the context of politics. It is the first opinion annotated social dataset from Malta, which has very limited language resources available.

## 1 Introduction

European usage trends show that Malta is the second highest user of social media, with around 90% of the adult population being online and active on social media (Eurostat, 2017), whereas around 80% of users read news online (Caruana, 2018). In terms of social media, this is not only used by individuals, but is also increasingly being used by enterprises (Eurostat, 2018). In fact, governments and businesses are spreading their news via social media and moving away from newswires (Grech, 2019). This has increased the importance of social opinions and the need to refine data mining techniques that are able to identify and classify opinions related to a particular aspect, e.g., entity or topic, which can be beneficial.

This paper presents a dataset of opinion-annotated social online posts targeting the Malta Government Budget for 2018<sup>1</sup> presented on 9th October 2017 by the Honourable Minister for Finance, Edward Scicluna. It contains the opinions and reactions (in terms of sentiment, emotions,

<sup>1</sup><https://mfin.gov.mt/en/The-Budget/Pages/The-Budget-2018.aspx>

etc.) of the public and professionals towards the mentioned budget as expressed over various social channels, specifically, social networking services and newswires, during and after the event. In addition, it has the potential of identifying commendations, regrets and other reactions concerning any presented measure, such as tax matters, industry specific initiatives, strategic initiatives and social measures. This dataset can support government initiatives for the development of opinion mining tools to better capture the public perception towards an upcoming/current/past budget presented to the House of Representatives. Such valuable insights can be taken in consideration within the upcoming budgets and/or any bills presented and discussed in Parliament.

## 2 Related Work

The Politics domain is one of the most popular application areas in the social media-based opinion mining domain, with such techniques being applied on election, debate, referendum and other political events' (such as uprisings and protests) datasets. However, applying such techniques on government budgets is not common.

Kalampokis et al. (2011) proposed a method that integrates government and social data (from social media platforms, such as Twitter and Facebook) to enable decision makers to understand public opinion and be able to predict public reactions on certain decisions. The methodology discussed by Hubert et al. (2018), uses emotion analysis to study government-citizen interactions on Twitter for five Latin American countries that have a mature e-Participation, namely Mexico, Colombia, Chile, Uruguay and Argentina. Similarly, the city of Washington D.C. in the United States, uses sentiment analysis to interpret and examine the comments posted by citizens and businesses over

social media platforms and other public websites (Eggers et al., 2019).

The economic content of government budgets is made publicly available for various countries. The Global Open Data Index<sup>2</sup> provides public open datasets about national government budgets of various countries, but lacks open budget social and transactional data. The OpenBudgets<sup>3</sup> Horizon 2020 project provided an overview of public budget and spending data and tools, in order to support various entities (Musyaffa et al., 2018). However, this project targeted public budget and spending data and not the yearly budgets presented by governments. Moreover, it did not use any data from social media platforms and apply any text mining tasks, such as opinion mining.

To the best of our knowledge, the gold standard presented is the first annotated dataset from a social aspect at a European and national level and in the context of Maltese politics.

### 3 Method

A variety of Web social media data covering the local Maltese political domain was taken in consideration for this study, namely traditional media published by newswires, and social media published through social networking services.

#### 3.1 Data Collection

The following data sources were selected to collect the dataset: i) Newswires (News): Times of Malta<sup>4</sup>, MaltaToday<sup>5</sup>, The Malta Independent<sup>6</sup>; and ii) Social networking services (SNS): Facebook<sup>7</sup>, Twitter<sup>8</sup>. The selection of the data sources were based on their popularity and usage with the Maltese citizens. In fact, Facebook and Twitter are two social media platforms that are highly accessed (TMI, 2018), with the Times of Malta, MaltaToday and The Malta Independent being amongst the top news portals accessed in Malta<sup>9</sup> for both reading and social interaction purposes. Table 1 presents details about the social dataset collected on the Malta Budget 2018.

Three different kinds of online news articles –in

<sup>2</sup><https://index.okfn.org/dataset/budget/>

<sup>3</sup><http://openbudgets.eu/>

<sup>4</sup><https://www.timesofmalta.com/>

<sup>5</sup><https://www.maltatoday.com.mt/>

<sup>6</sup><http://www.independent.com.mt/>

<sup>7</sup><https://www.facebook.com/>

<sup>8</sup><https://twitter.com/>

<sup>9</sup><https://www.alexa.com/topsites/countries/MT>

Source Type & Name	Query strings/ Articles	Online Posts
SNS - Twitter	1	38
SNS - Facebook	1	28
SNS - Twitter-The Malta Independent	1	12
News - Times of Malta	4	249
News - MaltaToday	4	175
News - The Malta Independent	4	45

Table 1: Data sources used for the consolidated dataset

terms of content published– were selected for each newswire mentioned:

- Overview of the upcoming budget, published on the budget day;
- Near to real-time live updates in commentary format, on the budget measures being presented for the upcoming year;
- Overview of the presented budget, published after the budget finishes, on the same day and/or the following day.

The aforementioned news articles above allow users to post social comments, which in nature are similar to online posts published on social networking services. These comments were extracted for our dataset, given that the annotation of opinions from user-generated social data is the main objective of this work. In addition, for diversity purposes, four online articles for each newswire were chosen to gather as much online posts as possible from the general public. This ensures that the different opinions expressed throughout on both the budget at large and specific budget topics, are captured.

With regards to the online posts from social networking services, a small sample was extracted, specifically the ones that contained the “malta budget 2018” search terms (as keywords and/or hashtags) that were posted on 9th and 10th October 2017. The criteria for the chosen keywords were based on the manual identification of common keywords associated with content relevant to the Malta Budget. The necessary filters were applied to exclude any spam and irrelevant content, whereas any references to non-political people were anonymised.

#### 3.2 Annotation

A total of 555 online posts were presented to **two** raters. Both were proficient in Malta’s two official

languages - Maltese (Malti) –a Semitic language written in the Latin script that is the national language of Malta– and English, which are equally important<sup>10</sup>. Moreover, the raters worked in the technology domain, were given a lecture about opinion mining and provided with relevant reading material and terminology for reference purposes. The following metadata and annotation types (#6-13) were created for each online post:

1. **Online Post Identifier:** unique numerical identifier for the online post;
2. **Related Online Post Identifier:** numerical identifier for the parent online post (if any);
3. **Source Identifier:** numerical identifier referring to the actual data source (e.g., website) of the online post;
4. **Source Name:** origin of the online post (e.g., Twitter, MaltaToday);
5. **Online Post:** textual string of the online post;
6. **Sentiment Polarity:** a categorical value (3-levels) for the sentiment polarity of the online post (negative, neutral, positive);
7. **Sentiment Polarity Intensity:** a categorical value (5-levels) for the sentiment polarity intensity of the online post (very negative, negative, neutral, positive, very positive);
8. **Emotion (6-levels):** a categorical value for the emotion of the online post based on Ekman’s (Ekman, 1992) six basic emotions (anger, disgust, fear, happiness, sadness, surprise);
9. **Emotion (8-levels):** a categorical value for the emotion of the online post based on Plutchik’s (Plutchik, 1980) eight primary emotions (joy, sadness, fear, anger, anticipation, surprise, disgust, trust);
10. **Sarcasm/Irony:** binary value, with 1 referring to sarcasm and irony in online posts<sup>11</sup>;
11. **Negation:** binary value, with 1 referring to negated online posts<sup>12</sup>;

<sup>10</sup>In Malta both languages are used by the general public, especially English or a mix for transcription purposes, hence why it is important to collect online posts in both

<sup>11</sup>These are treated as one class for this study

<sup>12</sup>A negated post refers to the opposite of what is conveyed due to certain grammatical operations such as ‘not’

12. **Off-topic:** binary value, with 1 referring to off-topic online posts that are political but not related to the budget;
13. **Maltese:** binary value, with 1 referring to online posts (full text or majority of text) in Maltese, and 0 referring to posts in English.

The raters were advised to follow any web links present in their online posts, for example “Budget 2018: #Highlights and #Opportunities can be accessed here - [#https://lnkd.in/eQxeM7G](https://lnkd.in/eQxeM7G) #MaltaBudget18 #KPMG”, when required to reach a decision, especially for determining the sentiment polarity, sentiment polarity intensity and/or emotion.

The online post with textual content “*Tallinja Card b’xejn għal dawk bejn 16 u 20 sena. #maltabudget2018*” (English translation: “free transport card for people aged between 16-20 years”) provides an example of the annotation types created for each collected post:

- **Sentiment Polarity:** Positive;
- **Sentiment Polarity Intensity:** Positive;
- **Emotion (6-levels):** Happiness;
- **Emotion (8-levels):** Joy;
- **Sarcasm/Irony:** 0;
- **Negation:** 0;
- **Off-topic:** 0;
- **Maltese:** 1.

### 3.3 Reliability and Consolidation

Inter-rater reliability, that is, the level of agreement between the raters’ annotations was determined. The percent agreement (basic measure) was primarily calculated on the annotations performed by the two raters. This was followed by the Cohen’s Kappa (Cohen, 1960), a statistical measure that takes chance agreement into consideration, which is commonly used for categorical variables. Moreover, this statistic is calculated when two raters perform annotations on the same categorical values and dataset. Table 2, shows the inter-rater reliability agreement scores for each annotation type.

A fair Kappa agreement was achieved for the sentiment polarity, sentiment polarity intensity and emotion (6-levels) annotations, with a slight

agreement obtained for the emotion (8-levels) annotation<sup>13</sup>. The percent agreement highlights the challenges behind these annotation tasks, especially when an annotation type such as emotion, has several categorical values to choose from and can convey multiple ones, e.g., anger and surprise. These annotation tasks are not trivial, where detecting emotion in text can be difficult for humans due to the personal context of individuals which can influence emotion interpretation, thus resulting in a low level of inter-rater agreement (Canales Zaragoza, 2018). Moreover, words used in different senses can lead to different emotions, hence making emotion annotation more challenging (Mohammad and Turney, 2013). This claim is also supported by Devillers et al. in (Devillers et al., 2005), who mention that categorisation and annotation of real-life emotions is a big challenge given that they are context-dependent and also highly person-dependent, whereas unambiguous emotions are only possible in a small portion of any real corpus. Therefore, the nature of relevant emotion data is too infrequent to provide adequate support for consistent annotation and modelling through fine-grained emotion labels.

Furthermore, a moderate agreement was achieved for sarcasm/irony detection, whereas negation obtained a chance agreement, which underlines how challenging such a task can be. Off-topic annotations achieved a fair level of agreement, whereas detection of Maltese online posts resulted in a near perfect agreement.

Annotation Type	% Agreement	Cohen's Kappa
Sentiment Polarity	0.6015	0.3703
Sentiment Polarity Intensity	0.4132	0.2182
Emotion (6-levels)	0.3985	0.2394
Emotion (8-levels)	0.2669	0.119
Sarcasm/Irony	0.7843	0.5027
Negation	0.8940	0.0581
Off-topic	0.7148	0.3494
Maltese	0.9854	0.9669

Table 2: Inter-rater reliability measures for each annotation type

A third expert in the domain consolidated the annotations to create a final dataset. In cases where both raters agreed on the annotation this was selected, whereas in cases of non-agreement,

<sup>13</sup>Ekman's 6-levels (Ekman, 1992) and Plutchik's 8-levels (Plutchik, 1980) emotion categories were chosen due to them being the most popular for Emotion Analysis

the third expert selected the most appropriate one to the best of their knowledge.

## 4 Dataset

The gold standard obtained through the method described in Section 3 consists of 547 online posts. This number was achieved after discarding irrelevant posts and ones that consisted of images only. Moreover, some online posts that were originally collected after the budget, were deleted from the original data source at the time of rating, in which case they were also removed. The distribution of the dataset annotations are represented as follows: sentiment polarity in Table 3, sentiment polarity intensity in Table 4 and emotion in Tables 5 and 6, respectively.

Polarity	Online Posts	Percentage
Positive	122	22.3%
Neutral	145	26.5%
Negative	280	51.2%

Table 3: Distribution of sentiment polarity annotations

Polarity Intensity	Online Posts	Percentage
Very Positive	37	6.8%
Positive	85	15.5%
Neutral	145	26.5%
Negative	193	35.3%
Very Negative	87	15.9%

Table 4: Distribution of sentiment polarity intensity annotations

Emotion	Online Posts	Percentage
Anger	131	23.9%
Disgust	159	29.1%
Fear	10	1.8%
Happiness	132	24.1%
Sadness	26	4.8%
Surprise	89	16.3%

Table 5: Distribution of emotion (6-levels) annotations

Emotion	Online Posts	Percentage
Anger	121	22.1%
Anticipation	95	17.4%
Disgust	154	28.2%
Fear	5	0.9%
Joy	50	9.1%
Sadness	23	4.2%
Surprise	39	7.1%
Trust	60	11%

Table 6: Distribution of emotion (8-levels) annotations

The dataset annotation results displayed do not fully reflect the opinions portrayed by the writers,

since a large amount of online posts were off-topic to the budget (34.2%). These are still very relevant for filtering out noisy user-generated posts, which are very common in Malta for such kind of public feedback, especially in newswire comments. Examples of such posts are the ones discussing the topic of smoking and how easy/difficult it is to stop smoking and on the contraband of cigarettes. These emerged as a result of no budget measure being taken towards increasing cigarette prices.

Moreover, certain sentiment polarities, polarity intensities and/or emotions were not targeted at budget measures, but to some previously submitted online post/set of posts. In such cases, the context of the online posts should be considered when determining the opinion, including any related posts<sup>14</sup>. This is a task for aspect-based opinion mining (Hu and Liu, 2004), which classifies a particular opinion type, such as sentiment polarity and/or emotion, for a given entity/aspect, such as a political party or budget measure.

Table 7 presents the distribution of sarcasm/irony, negation, off-topic and Maltese annotations. With regards to the latter, several online posts contained text with Maltese and English terminology. The ones that contained only one term/phrase in a particular language were not considered when annotating the language. The sarcasm and irony annotation was merged given that they convey similar characteristics in the content meaning the opposite of what is being said, where the former has a malicious intention towards the target i.e. person, whereas the latter does not.

Annotation	Online Posts	Percentage
Sarcasm/Irony	126	23.0%
Negation	39	7.1%
Off-topic	187	34.2%
Maltese	177	32.4%

Table 7: Distribution of the sarcasm/irony, negation, off-topic and Maltese language annotations

The dataset has been published<sup>15</sup> for general use under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license<sup>16</sup>.

<sup>14</sup>346 online posts were related to at least one another post

<sup>15</sup><https://github.com/kcortis/malta-budget-social-opinion/>

<sup>16</sup><https://creativecommons.org/licenses/by-nc-sa/4.0/>

## 5 Benefits

The following are the benefits of this dataset for the Natural Language Processing community:

- Contains online posts in Malta’s two official languages, Maltese and English;
- Hand-crafted rules using linguistic intuition can be built based on the given data, i.e., a knowledge-based approach, which can be a good start if a rule-based social opinion mining approach is primarily used before evolving towards a hybrid approach (rule and machine learning/deep learning-based) once more data is collected and annotated. The VADER lexicon and rule-based sentiment analysis tool is one such example of a high performing knowledge-based system that implements grammatical and syntactical rules (Hutto and Gilbert, 2014);
- Can be used to bootstrap a semi-automatic annotation process for large-scale machine learning i.e., deep learning models;
- Can encourage more researchers/people working in this domain to add to this dataset which is available for public use;
- Is a representative corpus for computational corpus linguistic analysis for social scientists.

## 6 Conclusions

We have described a novel dataset of social opinions for the Malta Government Budget 2018 which is a valuable resource for developing opinion mining tools that gather political and socio-economic insights from user-generated social data in Malta’s two official languages, Maltese and English. Therefore, it has potential of being used for initiatives by the Maltese Government, such as building intelligence on the Maltese Economy. The novelty of including Maltese and English online posts in this dataset, makes it a valuable resource for Maltese Human Language Technology and for testing Opinion Mining applications in general.

Moreover, it has been annotated with several forms of opinions in sentiment, emotion, sarcasm and irony, making it highly beneficial and a first contribution of its kind for Malta. Finally, this dataset is a work in progress and its volume will be increased per annual budget.

## References

2018. [Social media usage trends in malta in 2018](#). The Malta Independent.
- Lea Canales Zaragoza. 2018. Tackling the challenge of emotion annotation in text.
- Claire Caruana. 2018. [Just 10 per cent of maltese are not on social media](#). Times of Malta.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Laurence Devillers, Laurence Vidrascu, and Lori Lamel. 2005. Challenges in real-life emotion annotation and machine learning based detection. *Neural Networks*, 18(4):407–422.
- William D. Eggers, Neha Malik, and Matt Gracie. 2019. [Using ai to unleash the power of unstructured government data](#). Deloitte Insights.
- Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.
- Eurostat. 2017. [Living online: what the internet is used for](#).
- Eurostat. 2018. [Almost 75 % of enterprises in malta use social media](#).
- Herman Grech. 2019. [Here’s the bad news](#). Times of Malta.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM.
- Rocío B Hubert, Elsa Estevez, Ana Maguitman, and Tomasz Janowski. 2018. Examining government-citizen interactions on twitter using visual and sentiment analysis. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, page 55. ACM.
- Clayton J Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.
- Evangelos Kalampokis, Michael Hausenblas, and Konstantinos Tarabanis. 2011. Combining social and government open data for participatory decision-making. In *International Conference on Electronic Participation*, pages 36–47. Springer.
- Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.
- Fathoni A Musyaffa, Lavdim Halilaj, Yakun Li, Fabrizio Orlandi, Hajira Jabeen, Sören Auer, and Maria-Esther Vidal. 2018. Openbudgets.eu: A platform for semantically representing and analyzing open fiscal data. In *International Conference on Web Engineering*, pages 433–447. Springer.
- Robert Plutchik. 1980. [Chapter 1 - a general psycho-evolutionary theory of emotion](#). In Robert Plutchik and Henry Kellerman, editors, *Theories of Emotion*, pages 3 – 33. Academic Press.