An Application of Sentiment Analysis Techniques to Determine Public Opinion in Social Media

Andrew Jones
Dept. of Computing & Inf. Systems
Northumbria University
Newcastle upon Tyne, UK

Jeremy Ellman*
Dept. of Computer & Inf. Systems
Northumbria University
Newcastle upon Tyne, UK
jeremy.ellman@northumbria.ac.uk

Nanlin Jin
Dept. of Computer & Inf. Systems
Northumbria University
Newcastle upon Tyne, UK
nanlin.jin@northumbria.ac.uk

***Abstract*—This paper describes a prototype application that gathers textual data from the microblogging platform Twitter and carries out sentiment analysis to determine the polarity and subjectivity in relation to Brexit, the UKś exit from the European Union. The design, implementation and testing of the developed prototype will be discussed and an experimental evaluation of the product described. Specifically we provide insight into how events affect public opinion and how sentiment and public mood may be gathered from textual twitter data and propose this as an alternative to opinion polls. Traditional approaches to opinion polling face growing challenges in capturing the public mood. Small sample response and the time it takes to capture swings in public opinion make it difficult to provide accurate data for the political process. With over 500 million daily messages posted worldwide, the social media platform Twitter is an untapped resource of information. Users post short real time messages views and opinions on many topics, often signed with a '#hashtag' to classify and document the subject matter in discussion. In this paper we apply automated sentiment analysis methods to tweets giving a measure of public support or hostility to a topic ('Brexit'). The data were collected during several periods to determine changes in opinion. Using machine learning techniques we show that changes in opinion were also related to external events. Limitations of the method are that age, location and education are confounding factors where Twitter users over represent a young, urban public. However, the economic advantage of the method over real-time telephone polling are considerable.**

***Keywords*—Twitter, Sentiment Analysis, Opinion Polling Economics.**

I. Introduction

A. Background

In June 23rd, 2016 the United Kingdom held a referendum on whether the country should remain in or leave the European Union. Most traditional pollsters failed to provide a correct prediction of the vote, being unable to ascertain the public's opinion and feelings on the topic. In the modern world many individuals prime source of communication is through social networking platforms and websites such as Twitter. People offer opinions and views on various topics, generated in a digital text format providing a wealth of analytical material to measure public opinion. Artificial Intelligence (AI), Machine Learning and Natural Language Processing (NLP) techniques and methods such as Sentiment Analysis can be applied to effectively gauge what the public opinion is on a topic, and how external factors influence this over time.

The contribution of this paper is to suggest and describe an application that streams content from social media for further data-analysis. Once data has been captured it is analysed to determine of the current political climate in the UK. The application analyses sentiment and opinion mining of the content retrieved to determine the data's subjectivity and polarity. The application concentrates on the social media platform Twitter relating to current, divisive issues in UK politics, specifically Brexit.

This application could benefit many users, particularly in providing a time specific record of public opinion. However, the most important benefit is offering an alternative to traditional opinion polling methods. This would be available by only utilising a fraction of the resources normally required, while potentially offering a more accurate 'snapshot' of the current engagement of an issue. Both the voting public and political parties and institutions would be able to utilise this data, with the option to display findings publicly as traditional polling methods do, or as a tool for political parties to assess the effectiveness of their campaign and policy announcements.

Traditional opinion polling methods often include contacting a sample of participants, chosen to reflect the electorate, and then apply weighting to the sample to correct random fluctuations in the daily responses. Such samples are chosen to be statistically independent and identically distributed [1] ('iids'), but this ideal is challenging and difficult to achieve in practice.

In recent elections (such as the Brexit referendum and the past two British General Elections) we can see these traditional approaches face growing trials at capturing the public mood [2] Small sample response and the time it takes to capture swings in public opinion make it difficult for traditional polling to provide accurate data for the political process. For example, shortly before the Brexit referendum most polls agree the result would be largely remain. While the polls conducted may not be wrong, they failed to capture rapid changes in public opinion due to occurring events. There is also the aspect of truthfulness as the public will often say what they think is the politically correct answer then do the opposite i.e. head against heart. Twitter responses are generally truthful. In short, political polls carried out in real-time are rarely used and the electorate (as well as political institutions) must rely on data which expresses opinions which are no longer current or correct. With over 500 million daily messages posted worldwide, the social media platform Twitter is an extremely useful information resource. Users post real time messages (known as 'tweets') of typically around 34 characters (and at most 280, [3]) expressing views and opinions on a number of topics, often signed with a '#hashtag' as a way to classify and document the subject matter in discussion. By using a script developed in Python and several modules, these tweets can be streamed in real time by the application, providing large quantities of up-to date data without the delay and cost of traditional methods.

The paper proceeds as follows. Firstly we discuss some problems and limitations of traditional opinion polling, and describe the the area of sentiment analysis. Next the system design and implementation are discussed. We then move on to describe the data collection methodology and present an initial analysis. Finally we conclude with a discussion of limitations and propose future work.

II. Background and Related Work

It has been acknowledged through previous research, traditional methods at gauging public opinion on a topic are becoming increasingly unreliable and more difficult to accurately predict. Recently, events such as the 2016 US Presidential Election and the 2016 EU Referendum ("Brexit") are both examples of established polling companies providing inaccurate predictions; right up to the day of the voting the polls indicated Hillary Clinton would achieve victory and the Brexit vote would achieve a majority to "Remain" within the European Union. When the results were announced, many media sources were surprised when the predictions were proven false. To exacerbate the problem, the past two general elections in the UK both included polls providing inaccurate information indicating a clear problem. Whilst the polls used may have sound methodology, research suggests they struggled to capture a swing in public opinion due to real-world events, such as the murder of British MP Jo Cox or accusations regarding Clinton's handling of sensitive information on an unsecure e-mail server. Traditional methods include contacting a sample of participants, chosen to reflect and represent the wider electorate, over several days. As such, real-time polling is a costly and expensive and therefore rarely done. Decisions makers, businesses and policy experts must rely on out-of-date information which does not reflect the current mood of the public. A primary problem is in determining what public opinion is for a topic, and how real-world events influence this. Research has determined an automated-means of sentiment analysis could alleviate this. An investigation in the challenges, techniques and methods will be required.

The use of Social Media as a source of information and data well-documented. With a micro-blogging platform such as Twitter, users post a large variety of opinions, thoughts and views on topics that interest them. Topics are often grouped by #hashtag as a means to classify and document the subject being discussed; an ideal feature in establishing public opinion. However, issues with incorrect spelling and grammar, use of slang and emoticon symbols means it is difficult for researchers to filter relevant data. A secondary problem is the viability of using social media to analyse public opinion and obtaining an accurate, well presented data set from social media for further data analysis. Hence, an investigation into the theories and practices of opinion and data-mining will also be required.

A. Why Twitter

Here we consider whether social media is a useful tool in discussing the relationship between politics, the media and public opinion. Other forms of media have long been considered as a viable avenue to explore, with [4] discussing the relationship between Television (TV) content and the role of social media, such as the micro-blogging platform Twitter, encouraging debate. They explore the recent appearance of reality TV that focused on welfare recipients, which has led to concerns that main stream media are exacerbating misconceptions around major societal issues such as welfare reform and poverty. Public reaction to TV can be measured by performing an analysis on the social media "backchannel" that takes place over the broadcast. The quality of discussion in the backchannel is reflective of the viewers reaction to, and opinion of, the political stories and their framing that is evident within TV editing and production. This assumption is based on observations that the casual practice of engaging with a twitter #hashtag whilst simultaneously watching the content unfold, is a common and natural occurrence on the platform. [4] argue that platforms such as Twitter act as a 'second screening' practice, offering users the opportunity to engage with issues of political concern. The researchers developed two smartphone apps Spotting Guide and Moral Compass to create an environment which encourages users to identify, categorise, tag and filter the various tropes and patterns that appear during these broadcasts.

The first application Spotting Guide, is designed so users can identify and recognise negative or stereotypical patterns of reality TV broadcasts and note them down for further discussion. In the proceeding workshops, participants reported on how it affected their viewing experiences and think critically about the political significance of the broadcast. [4] summarising "These sentiments were tied into an overall view that the app could promote mindful viewing of TV". The second application Moral Compass allows viewers to view and tag live tweets under different categories. For example, the tweet "now a landlord kicks out 100 people (because of) delayed housing benefit payments" was tagged by users as "disturbing", reflecting how a consensus on public opinion is displayed on-line. [4] conclude that confirms applications such as these can encourage and support critical thinking about TV and its wider context. By urging users to pay attention to the broadcast and summarise what they perceive, users think and reflect on the framing of TV, awareness of content, and how and why it has been presented in that way. The most important aspect to take away is how users are motivated to search and engage with additional content as "Interaction is purposeful and shows how social tagging can evolve into a reflective process".This presents social media, its relationship with public opinion and politics, as a relevant area for academic study and discussion.

An additional research question to consider is what are the real-world benefits in examining social media? [5] present an automated web interface for tracking the prevalence of influenza-like illnesses using Twitter. They state how the monitoring and tracking of epidemic disease, such as seasonal influenza, is an important task and how the health sector uses several methods to detect and track epidemics. [5] argue traditional methods such as counting consultation rates and school or workforce absentee figures, are subject to time delays as well as a lack of infrastructure and how on the web based information provides an additional means to tackle the problem, It has been demonstrated that user queries on web search engines can be used to provide early warning for an epidemicthe social web media have a predictive power on different domains. A comparison to traditional methods is a reoccurring theme in the area of sentimental analysis.

In the light of this, [5] assert that Twitter data mining has a practical use which addresses traditional challenges. An important point is use of the application to complement, rather than replace existing practices.

B. Twitter and Sentiment Analysis

Twitter make their data available through a well-documented Application Programmers Interface (API). There are resource limits, but sufficient free access is available for academic research. Much communication takes places over social networking platforms with people giving their opinions on various topics in digital, text or statistical form, as such the challenges and techniques of processing this data must be discussed. [6] reiterate the practical use of this area, the data collected and classified by context allows predictions to be made and visualised for government, business, marketing and analyst groups to model policies and schemes. [6] aims to classify trending topics using four different categories. Due to the public nature of Twitter, no validation takes place when a user posts a message. As such, much of the text data obtained may contain too much information, of which some may be irrelevant. As these attributes do not contribute any meaningful information, they adding noise to the data affecting the accuracy of a predictive model. Examples of noise in this context may include slang, spelling errors, or spam. Two approaches to dimension reduction (reducing the number of random variables) discussed are Feature Selection (selecting the most relevant features) and Feature Extraction (combining attributes into a record set of features). As tweets collected using the Twitter API contain 80-85 parameters, feature extraction and selection are an important step to reduce this number to a manageable degree in this context features such as tweet text, created time, location are useful ones to consider.

While the viability of using Twitter to classify data has been discussed, its use in measuring public mood is interesting. [7] discuss how mood plays a role in processing information and forming political opinions but traditional methods present several difficulties. [7] argue that social media could overcome these issues, specifically by investigating the sudden changes in public mood.

Reference [7] gathered more than one million tweets over 30 days (01/06/16-30/06/16). Rather than classify the data the occurrence of each token within all tweets collected in an hour is used to produce a time series, representing how often it was present over the 30 days. This enabled a visualisation of key points at a specific time when public opinion can be seen to change. To measure public mood [7] take a lexicon-based approach to measure Positive, Negative, Anger, Anxiousness and Sadness found within the LIWC lexicon [8].

In observing the time-series for each measurement, [7] identified several events or change points. In these moments public mood is characterised by an increase in negativity, anger, anxiety and sadness, with a corresponding drop in positive affect, most significantly as the referendum outcome became the largest negative change in public opinion was seen. While this demonstrates the viability to examine public mood, the most important discussion point is the multitude of competing effects, whether predictable or unknown, that cause fluctuations in collective public mood. Examples of these key events include Football violence in Marseille, a Nightclub shooting in Orlando, the murder of Labour MP Jo Cox and the day of the referendum itself. In terms of political science, this study provides insight on how events and policies affect public opinion.

Any improvement in classification methods affects sentiment analysis. Datasets of tweets have been created that have been sentiment tagged by independent annotators. Sentiment system developers can compare their results on these datasets to indicates their relative accuracy. In particular there have been five Semeval (e.g. [9], [10]) competitions where teams are given training datasets, and then compete on an unseen test. [10] reports that leading systems use sophisticated deep machine learning, or ensemble of classification methods.

Reference [11] gives a sentiment annotated dataset in relationship to Brexit. Popular referenda can provide a rich environment for discourse understanding, however the correct application of techniques is required to view the full perspective, a contrast to most political bodies which focus on individuals or parties. The study is separated by three stages, sampling, filtering and annotation; sampling involves streaming 2000 random tweets, tracked in 75 categories by keyword, hashtags and account names which can be carried out via the Twitter Streaming API. Examples of keywords include "#votein" or "#voteleave". Next the sample is filtered; by which any tweet less than 3 characters was discarded and any user which posted more than 100 times per day was classed as spam and ignored. In addition, only content posted between

6am 11pm GMT was included to increase the coverage of postings in the appropriate time-zone. Finally, three annotators assigned a sentiment value (remain or leave), a strength (an integer between 1 and 5) and a contextual dependency (if the interpretation of the tweet depends on external sources or not). [11] state how this method supports a fine-grained view on opinion landscape. .

The gold-standard developed in [12] was built on the context of the SSIX (Social Sentiment Analysis Financial IndeXes). [13] report real-time monitoring of Brexit sentiment on Twitter. The results are presented as a series of rolling metrics known as "X-Scores" that provide raw aggregated sentiment, volumes, averages and a strength index. A deep learning classifier was specifically trained using [11]. Care was taken to improve accuracy so only 2 classes were utilised, "stay" or "leave". [13] concludes that the vote was a close call but not how close it would be between "remain" and "leave. The X-Scores used showed low movement between the sides, however a slow yet constant trend towards 'Remain' was detected before and during the vote "Although the SSIX platform analysedan average of 57.5% of the people would vote stay 48.1% voted stay". In addition, the platform spiked several times in the "leave" area but never stabilised and always returned to remain.

Ref. [13] note that age, location and education as confounding factors. The majority (35%) of Twitter users belong in the age bracket 15-24. As 15-18 year olds in the UK are unable to vote, this produced irrelevant data that was difficult to validate. An education gap was also present, with higher educated citizens consistently voting 'Remain' and those with no formal qualifications voting 'Leave'. [13] advises the age & education bias may be addressed with a correction factor based on an analysis of the social network users profile and their benchmark against the general population. Finally, location should have been considered many of the tweets gathered coming from the London area. As the capital contains a large multi-cultural, young and internet savvy population, this could account for additional bias where London Hammersmith Fulham, voted 70% 'Remain' but Doncaster voted 69% 'Leave'. The location bias may be addressed by considering the geo-location of the tweeter. These findings present a picture of Twitter's popularity with under-50, college-educated users from metropolitan areas skewing results. [13] suggest that opinion mining and sentiment analysis should never be used alone, but alongside other metrics, sources and tools to produce a multiple input analysis.

III. The System Design and Implementation

A. *Using Twitter in Programs*

Twitter support several Application Programmer's interfaces (APIs) that allow the platform to be successfully integrated into multiple applications. For example, the direct message API allows applications to send, receive and publish messages through an external source. A description of othere relevant APIs follows.

The Search API allows users to query content on Twitter. A search can be defined by keywords, #hashtags, tweets by a user and tweets referencing a user. The search API operates a lot like the search function on the Twitter mobile & web app, and although it can be used to search by topic or #hashtag, the API contains rate limits which diminishes its functionality for applications requiring large amounts of data to be acquired. Only tweets from the last 7 days are included in the free use version, however under the paid tiers, searches can obtain content up to 30 days old with even an option to search all tweets posted since 2006 at the top tier.

Unlike the search API which polls data from tweets that have already happened, the Twitter Streaming API instead gathers tweets as they happen in real-time. A set of criteria can be set (such as keywords, usernames, location, #hashtag etc.) and as tweets are posted which match these criteria, they are sent to the user. It is important to establish this is a push of data by Twitter, rather than a user initiating a pull request. There are two main drawbacks to this API. First, a constant HTTP connection must be established and maintained if tweets are streamed, so a constant reliable internet connection is required. Various streaming endpoints can be customised however, so this can be set depending on the use case. The second, most significant drawback of the Streaming API is that only a sample of tweets are provided at a time, anywhere from 1% to 40% of the total tweets real-time are streamed depending on the search criteria provided. For example, the broad criteria of any tweets mentioning the season of summer will stream a substantially larger number of tweets than a more niche, obscure criteria. This is largely because so many messages are posted daily, with Twitter claiming 500 million tweets are posted daily, the infrastructure provided is unable to support this.

Tweets obtained through the Twitter API are retrieved in JSON format and contain 80-85 data fields. These include obvious attributes ranging from date/time of creation, username, textual data to obscure details such as the background colour on the profile of the user who posted the tweet. When connecting to the streaming API a high-level scripting language is required .

Python is a language widely used in data-mining due to the extensive libraries available for scientific purposes. One open-source library is Tweepy, a package that allows Python to communicate with the Twitter platform and use its API. Tweepy uses OAuth, a method supported by Twitter for secure API connections. A client application must be registered with Twitter via the Twitter developer website. Once a new application is created a consumer token, consumer secret key, access token and access secret keys are generated for the use of authorisation. There is no expiry on the tokens provided so the only time it can go invalid is when a user revokes application privileges. Although documentation for Tweepy is not as complete as other libraries such as twitter-python or Tython, it relies heavily on the Twitter API which has excellent documentation. In addition Tweepy boasts an active development community with features being regularly updated and added.

B. Data Collection System

The first focus will be on obtaining the dataset, using Tweepy to directly stream any tweets being produced. A simple class (known as a stream-listener) is used to allow the application to spot and capture any content by user, #hashtag or terms such as ‘Brexit’ or ‘Election’. Each tweet is saved as an object containing data such as user ID, time sent, location, text content etc.

Once the stream is ’open’ it will continue to pull new data until stopped. Using text processing library such as TextBlob, the messages can be parsed and tokenised, and irrelevant data removed before a subjectivity and polarity score is assigned to the message. These scores are integer values between -1 and +1 each with -1 representing negative and subjective content and the opposite for +1. Finally, the contents of the message as well as the score are saved to a separate file for analysis. By following this method, we can run the application during live events and instantly calculate a ‘snapshot’ of public opinion.

A UML Sequence Diagram for the process is given in Fig. 1, where the authorised retrieval is referred to as ‘scrape’.

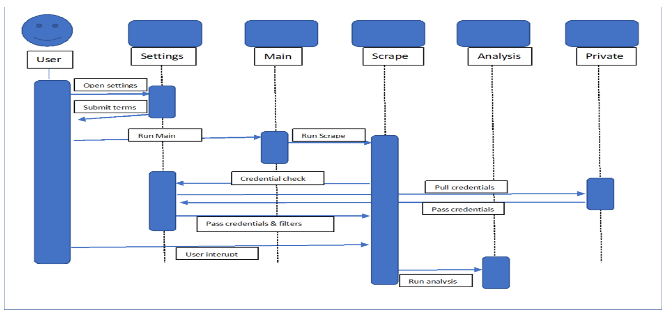

Figure 1: UML Sequence Diagram

C. The Experiment

From the 21st February 2018 to 11 April 2018 tweets were collected every day at 12pm for an hour, with an average of 10,000 - 12,000 tweets gathered in this time frame. Each tweet was cleaned, then a polarity and subjectivity score calculated using TextBlob.

A daily average for the positive, negative and neutral tweets was created. Note that this data does not describe agreement or disagreement with Brexit, or a preference for the ‘Remain’ & ‘Leave’ sides; rather, it represents how users feel and express themselves about the topic of Brexit and provides an accurate snapshot of public opinion.

D. Results and Analysis

The daily averages were added to a spreadsheet to demonstrate a change in sentiment over time, with an average figure calculated. Throughout the period of these experiments 54.7% of tweets posted on the topic of Brexit expressed a negative sentiment. 30.2% were positive, with 15.1% of tweets being neutral. (see Fig. 2)

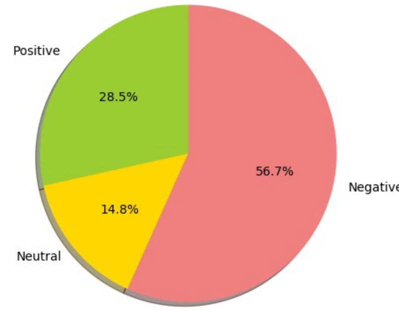

Figure 2: Sentiment pie chart

In addition to the total averages, the positive, negative and neutral daily means were plotted over time to demonstrate a change in public mood. As seen from the chart Fig. 3 , a rise in negative sentiment indicates a drop in positive and negative sentiment.

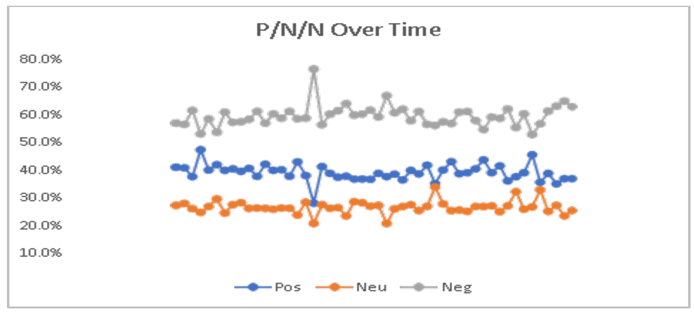

Figure 3: Percentage Vs Swing over Time

The final stage in the experiment involved using the total average as a standard for a sentiment prediction. The percentage of deviation from the predictor was classed as a daily swing, with any swing more than 3% highlighted for interest (see Fig. 4 below).

	A	B	C	D	E	F	G	H	I	J	K
1	Date	Pos	Neu	Neg	Opinion	Predic Pos	Predic Neu	Predic Neg	Pos Swing	Neu Swing	Neg Swing
2	21-Feb	32.5%	15.9%	51.6%	32.92%	30.2%	15.10%	54.70%	2.3%	0.8%	-3.1%
3	22-Feb	32.2%	16.7%	51.1%	35.07%	30.2%	15.10%	54.70%	2.0%	1.6%	-3.6%
4	23-Feb	28.4%	14.5%	57.1%	31.27%	30.2%	15.10%	54.70%	-1.8%	-0.6%	2.4%
5	24-Feb	40.1%	12.9%	47.0%	32.31%	30.2%	15.10%	54.70%	9.9%	-2.2%	-7.7%
6	25-Feb	31.3%	15.4%	53.4%	32.14%	30.2%	15.10%	54.70%	1.1%	0.3%	-1.3%
7	26-Feb	33.6%	18.7%	47.7%	32.95%	30.2%	15.10%	54.70%	3.4%	3.6%	-7.0%
8	27-Feb	31.1%	12.6%	56.3%	30.32%	30.2%	15.10%	54.70%	0.9%	-2.5%	1.6%

Figure 4: Spreadsheet dates and sentiments with % change

Chart Fig. 5 below demonstrates a visualisation of the swing over time. It shows how real-world events have an impact on public opinion. Looking at this data there are several points when a noticeable swing occurs which corresponds to real world events. For example, on February 24th, 2018 a positive swing of 9.92% and a negative swing of -7.70%. On this date the European Council President Donald Tusk gave a speech expressing the need for unity and cooperation (Tusk 2018). This expression for moving away from negativity can be reflected in the daily swing. Another example can be seen during the 10th March 2018.

On this date the leader of the UK’s Labour party, Jeremy Corbyn, gave a speech on his parties’ vision for Brexit, within

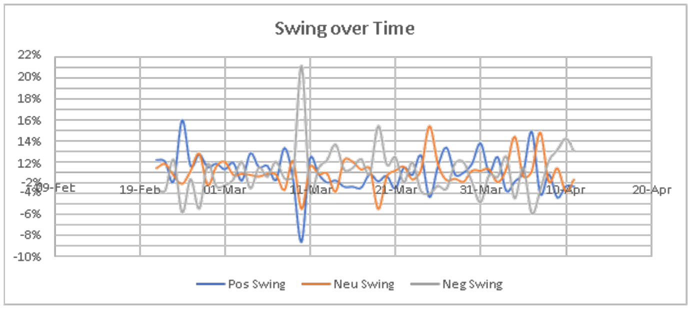

Figure 4: Percentage Vs Swing over Time

which he criticised the party in government, the Conservatives for their handling of Brexit and declared "It would, therefore, be wrong to sign a single market deal without agreement that our final relationship with the EU would be fully compatible with our radical plans to change Britain's economy." The issue of single-market membership is an extremely for the British public. As such the extent of discussion is visualised below with many users expressing frustrations online.

IV. Conclusions and Future Work

This paper has argued that current opinion polling methods often yield inaccurate results where public views change quickly over time, and where possible outcomes are close. Although opinion pollsters aspire to accuracy so that their samples represent the population, such samples easily become skewed due to their relatively small size, and high acquisition cost. As such, we have explored the use of social media since it is both economic and readily available.

The use of Twitter as an opinion mining tool has several advantages and disadvantages compared against existing polling techniques. Firstly, it is immediate, and is readily accessed using simple computer APIs. However, results are obviously skewed to twitter users who may not represent the population at large.

The design and development of an application that collects streamed data from the microblogging platform Twitter has been described, with results filtered by the topic 'Brexit'. The data were processed and analysed to determine the sentiment polarity and subjectivity. This data was used in a series of experiments to examine the feasibility of such a product in providing a more accurate picture of gauging the public's opinion on a topic over time, and how it was influenced by external factors. The results demonstrate support for the hypothesis that sentiment analysis of tweets does reflect public opinion.

In future work we are planning to refine our approach taking advantage of improvements in sentiment analysis, and deep learning approaches to Artificial Intelligence.

References

[1] H. A. David and H. N. Nagaraja, "Order statistics," *Encyclopedia of Statistical Sciences*, vol. 9, 2004.

[2] K. Pedley, "Why polling matters: The role of data in our democracy," in *Political Communication in Britain*. Springer, 2019, pp. 265–275.

[3] S. Perez, "Twitters doubling of character count from 140 to 280 had little impact on length of tweets," *Tech Crunch*, 2018, https://techcrunch.com/2018/10/30/twitters-doubling-of-character-count-from-140-to-280-had-little-impact-on-length-of-tweets/, Access Date: 3 Aug, 2019.

[4] T. Feltwell, G. Wood, K. Long, P. Brooker, T. Schofield, I. Petridis, J. Barnett, J. Vines, and S. Lawson, "I've been manipulated!: Designing second screen experiences for critical viewing of reality tv," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 2252–2263.

[5] V. Lampos, T. De Bie, and N. Cristianini, "Flu detector-tracking epidemics on twitter," in *Joint European conference on machine learning and knowledge discovery in databases*. Springer, 2010, pp. 599–602.

[6] P. Sharma, A. Agrawal, A. Lalit, and A. Garg, "Challenges and techniques in preprocessing for twitter data," *International Journal of Engineering Science*, vol. 6611, 2017.

[7] T. Lansdall-Welfare, F. Dzogang, and N. Cristianini, "Change-point analysis of the public mood in uk twitter during the brexit referendum," in *Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on*. IEEE, 2016, pp. 434–439.

[8] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: Liwc 2001," *Mahway: Lawrence Erlbaum Associates*, vol. 71, no. 2001, p. 2001, 2001.

[9] P. Nakov, A. Ritter, S. Rosenthal, F. Sebastiani, and V. Stoyanov, "Semeval-2016 task 4: Sentiment analysis in twitter," in *Proceedings of the 10th international workshop on semantic evaluation (semeval-2016)*, 2016, pp. 1–18.

[10] S. Rosenthal, N. Farra, and P. Nakov, "Semeval-2017 task 4: Sentiment analysis in twitter," in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, pp. 502–518.

[11] M. Hürlimann, B. Davis, K. Cortis, A. Freitas, S. Handschuh, and S. Fernández, "A twitter sentiment gold standard for the brexit referendum," in *Proceedings of the 12th International Conference on Semantic Systems*. ACM, 2016, pp. 193–196.

[12] M. M. Hürlimann, B. Davis, M. Zarrouk, M. A. Poulis, and D. L. PU, "Social sentiment analysis financial indexes."

[13] L. Vasiliu, R. McDermott, M. B. Khaled, T. Daudert, D. Byrne, M. Hürlimann, S. Fernández, A. Cavallini, A. Freitas, F. Carol *et al.*, "In or out? real-time monitoring of brexit sentiment on twitter." in *SEMANTiCS (Posters, Demos, SuCCESS)*, 2016.