

# 1 Related Work

## 1.1 Sentiment Analysis Background

Historically Opinion Mining and Sentiment Analysis (OMSA) of a computer based nature arose as a method to gain further insight into “Customer Reviews” in the commercial markets of movies (??) and products (??). As ? states ‘The Web has changed from “read-only” to “read-write” and what came of this change was an increasing amount of user engagement with websites, blogs, forums and social media all of which lend themselves to being a rich source of public opinion.

### Approaches to Sentiment Extraction

In literature, the common underpinnings of research towards sentiment extraction follow a rule based or statistical approach with some scholars combining the two. Early rule based approaches employed the creation and usage of context-free rules independent of the domain in which they are investigating (?). These approaches used prior polarity to determine semantic orientation using generated lexicons (?) as well as the likes of established lexicons; SentiWordNet (?), the General Inquirer (?) and Maryland (?) to name just a few. The weaknesses as highlighted by ? is that the “coverage” of these lexicons are too generic and fail to aid in sentiment extraction of texts for which the sentiment lies in the domain from where they are sourced. For instance, the word ‘long’ may have positive sentiment when discussing the life-span of product i.e. ‘this watch is well built and will last long’ however when applied to the context of a movie i.e. ‘the movie was long and tedious’ the same phrase would carry a negative sentiment. This contrast of sentiment between domains leads to poor generalization when applying context-free lexicons like the aforementioned to a domain text. This therefore prompted further research into corpus based lexicons; ? found the reduction of a lexicon’s size and reclassifications of it’ contained sentiment polarities with respect to domain led to an overall improvement in accuracy of sentiment classification. Having derived that a lexical approach towards sentiment extraction of a domain specific corpus performs poorly and how ? hypothesizes ‘with regards to topic and temporal dependency, it is still more effective to use supervised machine learning and accept some loss in performance when analysing data of a different topic or time-period’ and thus the domain and temporal research as outlined by this paper extends on the supervised statistical approach towards sentiment extraction.

## 1.2 Sentiment Analysis on Twitter

More recently in the field of OMSA, research has been led into the user engagement of Social Media more specifically Twitter with a total of 11.06% of recent research papers

(2000-2015) containing twitter datasets (?). Twitter as a platform originally intended for user conversations has since developed into a social hub of subjective opinions due to the paradigm ‘shift from traditional communication tools (such as traditional blogs or mailing lists) to microblogging services’ (?). With the rise of ubiquity in microblogging among society and the ever-growing channels of discussion on Twitter, a wealth of opinion has been conceived thus presenting opportunism for new insights into a plethora of topical public opinion.

## **Areas of Focus**

The nature of previous OMSA twitter studies have varied as Twitter has evolved to become a leading online platform for short-informal communication. Again, the same early commercial motives of sentiment extraction can be seen in Twitter OMSA; studies into consumer brand opinion extraction have used lexicon’s comprised of uni-grams and bi-grams with each having an associated positive and negative polarity distribution. ? were able to use such a lexicon to classify a tweet on a five-point scale using a multinomial statistical Bayes model approach. However, there are numerous nonprofit Twitter sentiment extraction studies focused on past and present topics. These studies typically aim to provide clarity on investigatory topics as well as building effective models in predicting future outcomes using twitter encapsulated entities as data to explore these aims. One such noteworthy study by ? investigated the correlation of collective mood changes on Twitter with social, economic and political events. The study uses previous work of extracting sentiment using an adapted psychometric profile of mood states test (?) to model the time series standard deviations on a six-point-mood-scale towards socio economic phenomena. Interestingly the results found by ? were ‘events in the social, political, cultural and economical sphere do have a significant, immediate and highly specific effect on the various dimensions of public mood’ with public mood deviations varying the most from ‘short-term events such as the news cycle, elections, national holidays, and other events’ compared that of long term economic events where public mood ‘changes seem to have a more gradual and cumulative effect’. The investigation into twitter discourse regarding socio economic phenomena therefore has substance in extracting directly affected public opinion. And it is this observation of linkage between socio economic events and twitter sentiment which is pertinent to this papers investigatory aims of extracting public opinion regarding Britain leaving the European Union using tweets from the eventful temporal window after the ‘eu-referendum results’.

## **Approaches to Sentiment Extraction**

Twitter Sentiment Analysis studies have seen significant work on classifying tweets or other twitter encapsulated entities with respect to binary and multinomial classifications such as the determination of a tweet’s polarity. The structure of supervised Twitter OMSA work have commonality in the following; the data is typically preprocessed, the classification al-

gorithms are refined through experiments using labelled training samples and the evaluation metrics of the final model are reported. The limitations of such a supervised model can be seen to be somewhat linked to the limitations of the available dataset. This assumption has been evidenced in Twitter such that the limitation in size of a dataset has been revealed to directly impact the performance accuracy of sentiment classifiers (?), however it is not the case that the size of dataset and performance accuracy are proportionally linked as the performance gain tapers off after the dataset is large enough. The maxima of performance gain from the size of data is arguably the point of vocabulary convergence where unseen samples can be accurately classified from a diverse and representative training set. Despite individuals casting their own opinion, it's argued that opinions of a topical context 'converge' in the language and points for which their opinion contains; (?) and (?) highlight this point of convergence in online discourse for products and political events respectively. Further to size; the correct labelling of a training set is key to producing an optimal classifier. Smaller studies into extracting sentiment through twitter have used manual approaches towards labelling data, however this as a practice is susceptible to introducing bias into the training of a model. Even if bias is removed through cross-labelling techniques the manual nature ceases to be a viable when having to manually annotate large numbers of samples which as highlighted before is needed to produce optimal accuracy. Instead, various studies have taken an automated approach to labelling using distant supervision on large scale training sets. These studies have implemented automated labelling through 'noisy labels' such as the use of hashtags and emojis (?) to act as the label for the semantic orientation of a tweet. Having deduced that an optimal model of sentiment extraction within Twitter is dependent on the representative size and labelling of its training set, the partitioning and labelling of this paper's training set will be carried out using automated labelling influenced by the established techniques of utilizing textual content of a tweet to procure its label.

### 1.3 Sentiment Analysis of Political Discourse on Twitter

The reactive nature of Twitter sentiment to socio economic phenomena as evidenced above has thus led to a growing number of studies into extracting sentiment on global political affairs.

#### Twitter as a viable source for extracting political sentiment

Public Opinion on political affairs has long been analysed using traditional polling methods. Although traditional methods are manual and costly to carry out, over time they have had significant scientific backing to ensure the results obtained are representative and non-bias. The weakness of polls comes from the restrictiveness of specified questions and ensuring the sampling is completely representative of the nation. Here Twitter differs, the free-form nature provides access to new patterns and the nationwide usage can arguably give a

representative sample of opinion from the nations younger demographic. Twitter sentiment analysis studies have echoed the results of traditional polls; ? found that the share of messages regarding political parties in the 2009 German National parliament election was a close indicator of the true election results; predicting vote share within 1.65 percent mean absolute error across the parties. ? also hypothesized that the percentage of tweets regarding a political entity has some correlation to the percentage of a vote, using the 2012 Iowa elections as a case study they achieved 3.1 percent mean deviation from the true outcome differing by 1.1 percent from that of the 2.2 percent mean deviation of the traditional polls. It could thus be argued that tweet count can be considered an indicator for predicting vote share in upcoming elections however critics have questioned the validity of such studies with respect to the choice of datasets (?). Despite criticisms on validity these earlier political sentiment systems posed the question whether tweets could be used as an indicator of an individuals political sentiment. Divulging away from the simpler models of using tweet count as an indicator of election presence; there have been studies which have used more complex properties of twitter entities to aid in political classifications. ? used features such as; profile information, who a user follows, the textual content of a users tweets and the hashtags included in those tweets to classify a users political affiliation with either the Democrats or Republicans. ? achieved an average success rate over 80 percent across all of the features. Most notably from this case study it seemed that users of the same political affiliation tended to follow the same users, have similar textual content in their tweets and signed their tweets using similar hashtags.

## Sentiment Analysis of Brexit

Sentiment Analysis Classifiers in general benefit from clear cut orientations and political discourse is of no difference, whether it be the choice of an electoral candidate, the support towards a political party or an opinion on a political event by providing sentiments which are distinguishable i.e. the semantic orientations are different enough to warrant individual orientations, a classification to a higher then random level of accuracy can be made. One political movement which has embodied a strong polarisation in sentiments and is the topic of this paper for that very reason is Britain triggering article 50 to action it's exit from the European union or more commonly known as Brexit. The timeline of Britain leaving the European union can be naively split into three temporal windows; pre eu-referendum, the negotiation period and Britain's exit from the European union.

During the pre eu-referendum period the polling consensus was one of a majority support towards voting for Britain to remain in the European Union. However, ? found that Twitter told a story contrary to this; it was discovered that a surge in pro-leave sympathies came in the summer of 2015 and remained persistently high up until the referendum. This contrast between the polls results and twitter is intriguing with twitter indicating the the

true outcome much earlier it highlights the predictive power of twitter should not be underestimated. In addition, ? also found that the leave side dominated the twitter sphere; finding that the quantity of leave tweets was consistently higher up until election day where the remain side fought to be heard in a last ditch attempt to convince swing voters, that the leave side used more positive language when tweeting and that leave tweets tended to discuss the future with more certainty compared to the remain counterpart which spoke more on the past with a greater tentative tone.

Post eu-referendum, Britain and it's democratic leadership face negotiations to define the terms of Britain's exit from the European Union. Despite Britain having voted toward leaving the European Union the discussion of Brexit still remains an active channel of discussion within the twitter arena but few researchers have chosen to analyse the sentiment in situ of the negotiations. ? had produced a real time sentiment system which analysed the twitter discussion around Brexit in the run up to the referendum but have continued to run their system post eu-referendum. The Edinburgh system extracts sentiment towards the leave and remain stances, tracks trending hashtags and streams in the ongoing discussion as well as highlighting active locations of brexit related tweets all visible through daily and overview granularity. The Edinburgh system however trivially extracts sentiment using 'the counts of human defined pro-Remain and pro-Leave hashtags' to signify sentiment towards the discussed brexit affiliation. Fundamentally, employing the count of a Brexit stance as the sentiment of said stance is a weak indicator of actual sentiment as tweets may critique a Brexit stance by mentioning it and not necessarily support the stance. ? real-time system also tracks Brexit stance sentiment but the Sheffield system differs from the Edinburgh system in that the system aims to analyse the topics discussed, the people involved in the discussion and opinions posted. The Sheffield system offers an improvement on the determination of a tweets Brexit stance; by using the presence of a pre-determined Brexit stance hashtag at the end position of a tweet the system can deduce the political stance of a given tweet.

Being that the systems analysing the temporal window post referendum in situ are scarce and that the collective results of the above Systems leave room to provide incremental improvements in sentiment extraction; the System developed as part of this paper will capitalize on this and produce further insight into public opinion on the matter.