

2013

Controversy trend detection in social media

Rajshekhar Vishwanath Chimmalgi
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Engineering Science and Materials Commons](#)

Recommended Citation

Chimmalgi, Rajshekhar Vishwanath, "Controversy trend detection in social media" (2013). *LSU Master's Theses*. 2000.

https://digitalcommons.lsu.edu/gradschool_theses/2000

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

CONTROVERSY TREND DETECTION IN SOCIAL MEDIA

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Interdepartmental Program in
Engineering Science

by

Rajshekhar V. Chimmalg
B.S., Southern University, 2010
May 2013

To my parents Shobha and Vishwanath Chimmalgi.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Dr. Gerald M. Knapp for his patience, guidance, encouragement, and support. I am grateful for having him as an advisor and having faith in me.

I would also like to thank Dr. Andrea Houston and Dr. Jianhua Chen for serving as members of my committee.

I would like to thank Dr. Saleem Hasan, for his guidance and encouragement, for pushing me to pursue graduate degree in Engineering Science under Dr. Knapp.

I would like to thank my supervisor, Dr. Suzan Gaston, for her encouragement and support. I also thank my good friends, Devesh Lamichhane, Pawan Poudel, Sukirti Nepal, and Gokarna Sharma for their input and supporting me through tough times. I would also like to thank all the annotators for their time and effort to help me annotate the corpus for this research. And thanks to my parents, brother, and sister for their love and support.

Table of Contents

Acknowledgments.....	iii
List of Tables	v
List of Figures	vi
Abstract	vii
Chapter 1: Introduction	1
1.1 Problem Statement	1
1.2 Objectives.....	3
Chapter 2: Literature Review.....	4
2.1 Motivation to Participate in Social Media.....	4
2.2 Trend Detection.....	4
2.3 Sentiment Analysis.....	9
2.4 Controversy Detection.....	10
2.5 Corpora.....	11
Chapter 3: Methodology	13
3.1 Data Collection.....	13
3.1.1 Annotation	16
3.1.2 Preprocessing.....	17
3.2 Controversy Trend Detection	18
Chapter 4: Results and Analysis	20
4.1 Controversy Corpus.....	20
4.2 Controversy Trend Detection	22
4.2.1 Feature Analysis	22
4.2.2 Classification Model Performance	23
4.2.3 Discussion	30
Chapter 5: Conclusion and Future Research.....	32
References	33
Appendix A: Wikipedia List of Controversial Issues	37
Appendix B: IRB Forms	44
Appendix C: Annotator's Algorithm	46
Vita.....	47

List of Tables

Table 1: Data elements returned by the Disqus API.....	15
Table 2: Interpretation of Kappa.....	21
Table 3: Feature ranking	22
Table 4: Performance comparison between different classifiers at six hours.....	23
Table 5: Classification after equal class sample sizes	24
Table 6: Confusion matrix of classification after equal class sample sizes	25
Table 7: Classification	26
Table 8: Confusion Matrix.....	27
Table 9: Classification after equal class sample sizes	29
Table 10: Confusion matrix of classification after equal class sample sizes	30

List of Figures

Figure 1: Data collection.....	14
Figure 2: Example of comments returned by the Disqus API	14
Figure 3: Database ERD	15
Figure 4: Definition of controversy	16
Figure 5: Example of the annotation process.....	17
Figure 6: Pseudocode for calculating controversy score and post rate	19
Figure 7: F-Score calculation.....	20
Figure 8: Distribution of comments	22

Abstract

In this research, we focus on the early prediction of whether topics are likely to generate significant controversy (in the form of social media such as comments, blogs, etc.). Controversy trend detection is important to companies, governments, national security agencies, and marketing groups because it can be used to identify which issues the public is having problems with and develop strategies to remedy them. For example, companies can monitor their press release to find out how the public is reacting and to decide if any additional public relations action is required, social media moderators can moderate discussions if the discussions start becoming abusive and getting out of control, and governmental agencies can monitor their public policies and make adjustments to the policies to address any public concerns.

An algorithm was developed to predict controversy trends by taking into account sentiment expressed in comments, burstiness of comments, and controversy score. To train and test the algorithm, an annotated corpus was developed consisting of 728 news articles and over 500,000 comments on these articles made by viewers from CNN.com. This study achieved an average F-score of 71.3% across all time spans in detection of controversial versus non-controversial topics. The results suggest that it is possible for early prediction of controversy trends leveraging social media.

Chapter 1: Introduction

Millions of bloggers participate in blogs by posting entries as well as writing comments expressing their opinions on various subjects, such as reviews on consumer products and movies, news, politics, etc. on social media website such as Twitter and Facebook, essentially providing a real-time view of opinions, intentions, activities, and trends of individuals and groups across the globe (Gloor et al., 2009). Recent surveys reveal that 32% of the nearly 250 million bloggers worldwide regularly give opinions on products and brands, 71% of active Internet users read blogs, and 70% of consumers trust opinions posted online by other consumers (Glass & Colbaugh, 2010).

Contents created by bloggers may enable early detection of emerging issues, topics, and trends in areas of interest even before they are recognized by the main stream media (Colbaugh & Glass, 2011). Detecting emerging trends is of interest to businesses, journalists, and politicians, who want to extract useful information on a particular time series and to make it possible to forecast future events (Mahdavi et al., 2009). These trends, however, are buried in massive amounts of unstructured text contents, and can be difficult to extract using automation.

For example, social media has been used in health care to estimate the spread of diseases. One such research conducted by Signorini et al. (2011) used Twitter to track rapidly-evolving public sentiment with respect to H1N1 or swine flu and to track and measure actual disease activity. They were able to estimate the influenza activity one to two week before Centers for Disease Control and Prevention (CDC). Emerging Trend Detection can assist CDC and other public health authorities in surveillance for emerging infectious diseases and public concerns.

An emerging trend is a topic area that is growing in interest and utility over time in social media sites such as Twitter, Facebook, blogs, etc. The task of emerging trend detection (ETD) is to identify topics which were previously unseen and are growing in importance from a larger collection of textual data within a specific span of time (Kontostathis et al., 2003). A controversial trend is a popular topic which invokes conflicting sentiment or views (Choi et al., 2010). Controversy trend detection is important to companies, governments, national security agencies, and marketing groups so that they can identify which issues the public is having problems with and to develop strategies to remedy them.

1.1 Problem Statement

For trend detection in social media, researchers have looked at burstiness of terms mentioned within a certain time span (Alvanaki et al., 2012; Cataldi et al., 2010; Heum et al., 2011; Jeonghee, 2005; Mathioudakis & Koudas, 2010; Reed et al., 2011). Cataldi et al. also considered authority of users in identifying trends. Others have looked at how a topic spreads in clusters of connected users (Budak et al., 2011; Glass & Colbaugh, 2010; Takahashi et al., 2011).

Sentiment analysis determines the sentimental attitude of a speaker or writer. Researchers have studied sentiment analysis using data from product reviews, blogs, and news articles, obtaining reasonable performances in identifying subjective sentences, determining their polarity values, and finding the holder of the sentiment found in a sentence (Kim & Hovy, 2006; Ku et al., 2006; Zhuang et al., 2006). Researchers have used sentiment lexicons consisting of words with positive and negative polarity in detection of controversy (Choi et al., 2010; Pennacchiotti & Popescu, 2010).

Little research has been done specifically in detecting controversial trends in social media. Pennacchiotti and Popescu (2010) have considered disagreement about an entity (i.e. proper nouns) and the presence of explicit controversy terms in tweets to detect controversy in Twitter resulting in an average precision of 66%. Choi et al. (2010) defines a controversial issue as a concept that invokes conflicting sentiment or views. They focus on detecting potential controversial issues from a set of news articles about a topic by using a probabilistic method called known-item query generation method and determine if the detected phrase is controversial by checking if the sum of the magnitude of positive and negative sentiments is greater than an specified threshold and difference between them is less than a specified threshold value. They evaluate their methodology on a dataset consisting of 350 articles for 10 topics by selecting top 10 issue phrases for each topic and asking three users if the phrase is appropriate for a controversial issue or not achieving a precision of 83%. While Vuong et al. (2008) used disputes among Wikipedia contributors, where an article is constantly being edited in a circular manner by different contributors expressing their personal viewpoint getting a precision of 15%.

A topic becomes popular if it is something that the public cares about or impacts them personally for example: cure for AIDS or Cancer, tax break for the middle class, tax the rich, free education, etc. (Deci & Ryan, 1987).

For a controversial topic to become popular to the public, it should exhibit the following characteristics:

- People like or dislike the topic and express extreme emotions – either they are for it or against it (Popescu & Pennacchiotti, 2010).
- Most people consider the topic to be controversial (Popescu & Pennacchiotti, 2010).
- People will share a topic which they strongly agree or disagree with (Takahashi et al., 2011).
- Public opinion is usually close to an even split for or against the topic (Choi et al., 2010; Popescu & Pennacchiotti, 2010).

This research focuses on the early prediction of whether topics are likely to generate significant controversy (in the form of social media such as comments, blogs, etc.). An algorithm is developed to predict controversial trends by taking into account sentiment expressed in comments, burstiness of comments, and controversy score. To train and test the algorithm, an annotated corpus was developed consisting of 728 news articles and comments made on those articles by viewers of CNN.com. The methodology predicts which articles are controversial or non-controversial and how early they can be predicted.

1.2 Objectives

Following are the objectives of this research:

- Develop an improved algorithm to detect controversial trends incorporating new features such as number of ‘Likes’, number of threaded comments, positive sentiment count, negative sentiment count, and controversy score.
- Create an annotated corpus for training/testing of the algorithm.
- Implement the model in an application.
- Analyze the models performance.

Chapter 2: Literature Review

2.1 Motivation to Participate in Social Media

Based on the theory of reasoned action, Hsu and Lin (2008) developed a model involving technology acceptance, knowledge sharing and social influence. They indicated that ease of use and enjoyment and knowledge sharing (altruism and reputation) were positively related to the attitude toward blogging. Social factors such as community identification and attitude towards blogging significantly influenced a blog participant's intention to continue to use blogs.

Blogging provides an easy way for a person to publish material on any topic they wish to discuss on the web. Blogging is an act of sharing, a new form of socialization. With a popular issue, a blog can attract tremendous attention and exert great influence on society, for example blogs describing the firsthand accounts of human rights violation and persecution of the Syrian people by the Assad regime.

Deci and Ryan (1987) have done research incorporating intrinsic motivation such as perceived enjoyment involving the pleasure and satisfaction derived from performing a behavior, while extrinsic motivation emphasizing performing a behavior to achieve specific goals or rewards has been done by Vellerand (1997).

The theory of reasoned action (Fishbein & Ajzen, 1975) advocates that a person's behavior is predicted by their intentions and that the intentions are determined by the person's attitude and subjective norm concerning his or her behavior. Social psychologists consider that knowledge sharing motivation has two complementary aspects – egoistic and altruistic. The egoistic motive was based on economic and social exchange theory. It includes economic rewards (Deci, 1975). Bock and Kim (2002) combined the theory of reasoned action and economic and social exchange theory to propose expected rewards, expected social associations, and expected contribution as the major determinants of an individual's knowledge sharing attitudes. The altruistic motive, assumes that an individual is willing to increase the welfare of others and has no expectation of any personal returns resembling organization citizenship behavior. While bloggers provide knowledge, they expect others' feedback, thus obtaining mutual benefit. Reputation, expected relationships, and trust are also likely to provide social rewards.

2.2 Trend Detection

Many researchers have used burstiness of terms within a certain timespan in order to detect trending topics. Alvanaki et al. (2012) created EnBlogue system, which monitors web 2.0 streams such as blog postings, tweets, and RSS news feeds to detect sudden increase in popularity of a tags. Research at the University of Toronto have created TwitterMonitor, a system which detects trends over Twitter stream by first identifying keywords which suddenly appear in tweets at an unusually high rate and groups them into trends based on their co-occurrences (Mathioudakis & Koudas, 2010). Jeonghee (2005) utilizes temporal information

associated with documents in the streams and discover emerging issues and topics of interest and their change by detecting buzzwords in the documents. A candidate term is considered a buzzword, if its degree of concentration is higher than a given threshold. Goorha and Ungar (2010) describes a system that monitors news articles, blog posts, review sites, and tweets for mentions of items (i.e. product or company) of interest, extract 100 words around the items of interest them and determine which phrases are bursty. A phrase is determined to be bursting if the phrase has occurred more than a specified minimum number of times today or recently occurred more than a specified number of times, and increased by more than a specified percentage over its recent occurrence rate. A phrase is determined to be significant if it is mentioned frequently and is relatively unique to the product with which it is associated.

Budak et al. (2011) in their study, propose two structure-based trend definitions. They identify coordinated trends as trends where the trendiness of a topic is characterized by the number of connected pairs of users discussing it and uncoordinated trends as trends where the score of a topic is based on the number of unrelated people interested in it. To aid in coordinated trend detection, they give a high score to topics that are discussed heavily in a cluster of tightly connected nodes by weighing the count for each node by the sum of counts for all its neighbors. To detect uncoordinated trends, they give high scores to topics that are discussed heavily by unconnected nodes by counting the number of pairs of mentions by unconnected nodes. They considered twitter hash tags as topics, getting 2,960,495 unique topics. Performing their experiment on Twitter data set of 41.7 million nodes and 417 million posts, they achieved an average precision of 0.93 with a sampling probability of 0.005.

Gloor et al. (2009) have introduced algorithms for mining the web to identify trends and people launching the trends. As inputs of their method, they take concepts in the form of representative phrases from a particular domain. In the first step geodesic distribution of the concept in its communication network is calculated. The second step adds the social network position of the concept's originator to the metric to include context-specific properties of nodes in the social network. The third step measures the positive or negative sentiment in which the actors use the concepts.

Glass and Colbaugh (2010) proposed a methodology for predicting which memes will propagate widely appearing in hundreds of thousands of blog posts and which will not. They considered a meme to be any text that is enclosed by quotation marks. They identify successful memes by considering the following features: happiness, arousal, dominance, positive, and negative characteristics of the text surrounding the meme, number of posts(t) by time t which mention the meme, post rate(t) by time t , community dispersion(t) by time t , number of k -core blogs(t) (cumulative number of blogs in a network of blogs that contains at least one post mentioning the meme by time t), number of early sensor blogs which mention the meme (early sensor blogs are those which consistently detect successful memes early). They perform their experiment on MemeTracker dataset, selecting 100 successful memes (which attract ≥ 1000 posts during their lifetime) and 100 unsuccessful memes (which attract ≤ 100 posts during their lifetime). Using

Avatar ensembles of decision tree algorithm to classify, they get accuracy of 83.5% within the first 12hrs after the meme is detected and 97.5% accuracy within first 120hr.

Heum et al. (2011) in their study performed (1) extraction of subtopics for topics using feature selection, (2) trend detections and analysis with those subtopics and searching of relevant documents, and (3) seed sentences carrying more specific trend information. Obtained representative features for a given topic using Improved-Gini Index (I-GI) algorithm. For a given topics, retrieved document groups including the topics from the dataset and extracted noun terms, calculated I-GI and used upper 20% features for each topic as subtopics. They evaluated performance of the kNN and SVM classifiers using F-measure, resulting in an F-score of above 95% from both classifiers for the task of retrieving documents for a given topic. They used documents which contained the topic word as true value. Documents containing the subtopics, document frequency and date of the document are used to visualize trends for the subtopics by graphs, tables, and text.

Cataldi et al. (2010) collected Twitter data for a certain timespan and represented the collected Twitter posts as vector of terms weighted by the relative frequency of the terms. They create a directed graph of the users, where an edge from a node A to node B indicates users A follows user B's twitter posts, and weight a given user's posts by their PageRank score. They modeled the life cycle of each term in the Twitter post by subtracting the relative combined weight of the term in the previous time intervals from its combined weight in the given time interval. The emerging terms are determined through clustering based on term life cycle. They use a directed graph of the emerging terms to create a list of emerging topics by co-occurrence measure and select emerging topics by locating strongly connected components in the graph with edge weight above a given threshold.

Cvijikj and Michahelles (2011) in their study propose and evaluate a system for trend detection based on characteristics of the posts shared on Facebook. They propose three categories of trending topics: 'disruptive events' – events that occur at a particular point of time and cause reaction of Facebook users on a global level, such as the earthquake in Japan, etc., 'popular topics' – popular topics might be related to some past event, celebrities or products/brands that remain popular over a longer period of time, such as Michael Jordan, Coca Cola, etc., 'daily routines' – correspond to some common phrases such as "I love you", "Happy Birthday", etc. To detect the topic of a post, they consider a term to be an n-gram with a length between 2 to 5 words belonging to the same sentence within the post. TD-IDF was used to weight the terms, which assigns weight to a term based on the frequency of occurrence of a term within a single document and the number of documents in the corpus which contain the given term resulting in an ordered list of most significant terms in the corpus. Terms which belong to the same topic are clustered together in two steps – (1) clustering by distribution, and (2) clustering by co-occurrence. Their clustering algorithm on average has a precision of 0.71, recall of 0.58, and F-measure of 0.48. The experiments were performed on 2,273,665 posts collected between July 22, 2011 and July 26, 2011 using Facebook Graph API.

Brennan and Greenstadt (2011) focus on identifying which tweets are part of a specific trend. Twitter displays top 10 trends on their homepage and the tweets which consists the trending words. Their system relies on word frequency counts in both the individual tweets and the information provided in the tweet author's profile. The weights of the user's profile word frequencies are reduced by 60%. The dataset created for this research consisted of 40757 tweets from top 10 current trends on Twitter for Jun 2nd through Jun 4th, 2010 and 2471 non-trending tweets from Jun 5th collected from twitter public timeline. The dataset contained 29881 unique users. Profile information was collected for each user and word frequencies were extracted for all words in the user description. The time zone was also pulled for each user to be used the geographic location for the tweets. A separate 'clean' dataset was also created from the original dataset which only included tweets with greater than 15 words, punctuation tokens and included at most one trend keyword. The "clean" dataset was reduced to 23939 tweets from the original dataset containing 43704 tweets. For all data, keywords relating to the trending topics were removed so as not to influence the classification task. The keywords to be removed came directly from the trending topic. They use Transformed Weight-normalized Complement Naïve Bayes classifier (TWCNB). The advantage of using TWCNB is the speed of training a Bayesian classifier while correcting for some of the major weaknesses that a naïve Bayesian approach can have when dealing with data sets that may have incongruous numbers of instances per class. The text modeling corrections TWCNB makes are transforming document frequency to lessen the influence of commonly appearing words and normalizing word counts so long documents don't negatively affect probabilities. The researches leave off the transformed part from TWCNB. They use machine learning techniques to identify which trending topic a tweet is part of without using any trend keywords as a feature.

Rong Lu (2012) proposes a method to predict the trends of topics on Twitter based on Moving Average Convergence-Divergence (MACD). Their goal is to predict in real-time whether a topic in Twitter will be popular in the next few hours or it will die. They monitor some key words of topics on twitter and compute two different timespan's moving averages in real-time, then subtract the longer period moving average from the shorter one to get a trend momentum of a topic. The trend momentum is used to predict the trends of topics in the future. To calculate the trend momentum, moving averages of the topic needs to be calculated. To calculate the moving averages, they divide continuous time into equal time spans and sum the frequency count of the keyword within the time span divided by the time window size. They calculate moving averages for a short time span and for a long time span and subtract the moving average of the shorter time span by the moving average of the longer time span. When the trend momentum of the topic changes from negative to positive, the trend of the topic will rise and vice versa. For their experiment they created two datasets, one consisting of crawled headlines from Associated Press (AP) and tweets consisting of the headlines words, resulting in 1118 headlines and more than 450,000 tweets. For the second dataset they collected about 1% of all public tweets from twitter and twitter trends for the same period, resulting in more than 20 million tweets and 1072 trending topics. Using their methodology, for the keyword 'iPad', they were able to identify that

iPad will be a trending topic 12 hours before Twitter classified it as a trending topic. They discovered that before the topic becomes a twitter trend topic, about 75% of topic's trend momentum value went from negative to positive in the last 16 hours.

Fabian Abel (2012) introduced Twitcident, a framework and a web-based system for filtering, searching and analyzing information about real world incidents or crises. Their framework features automated incident profiling, aggregation, semantic enrichment and filtering functionality. Their system is triggered by an incident detection module that senses for incidents being broadcasted by emergency services. Whenever an incident is detected, Twitcident starts a profiling the incident and aggregating twitter messages. They collect tweets based on keywords from the incident report from the emergency services. The incident profile module is a set of tuples consisting of facet value pair and its weight (i.e. importance of the tuple). A facet value pair characterizes a certain attribute of an incident with a value, for example ((location, Austin), 1). The Named Entity Detection (NER) module detects entities such as person, location, or organization mentioned in tweets. Twitcident classifies the content of the messages into reports about casualties, damages, or risks and also categorizes the type of experience (i.e. feeling, hearing, or smelling something) being reported using a set of rules.

Sakaki et al. (2012) in their research investigate real-time interaction of events such as earthquakes in Twitter and propose an event notification system that monitors tweets and delivers notifications using knowledge acquired from the investigation. First, they crawl tweets including keywords related to target event and use SVM classifier to classify if the tweet is related to the target event based on features – number of words in the tweet, position of the query word within a tweet, words before and after the query word, words in a tweet (every word in a tweet is converted to a word ID); second, they use particle filter to approximate the location of an event. A particle filter is a probabilistic approximation algorithm. The sensors are assumed to be independent and identically distributed. Information diffusion does not occur on earthquakes and typhoons, therefore retweets of the original tweet are filtered out. They have developed an earthquake reporting system that extracts earthquakes from Twitter and sends a message to registered users. They treat Twitter users as sensors to detect target events. Tweets are associated with GPS coordinated if posted using a smartphone or else the user's location on their profile is considered as the tweet location. Their goal is to detect first reports of a target incident, build profile of the incident and estimate the location of the event. For classification of tweets, they prepared 597 positive examples that report earthquake occurrence as training set, using SVM classifier, they get recall of 87.5%, precision of 63.64%, 73.69% f-measure for the query term “earthquake” and for the “shaking” query, they get a recall of 80.56%, 65.91% precision, and 82.5% F-measure. When alarm condition is set to 100 positive tweets within 10 minutes, they were able to detect 80% of the earthquakes stronger than scale 3 and 70% of the alarms were correct. Their alarm notifications were 5 minutes faster than the tradition broadcast medium used the Japan Meteorological Agency (JMA).

Achrekar et al. (2011) focus on predicting flu trends by using Twitter data. Centers for Disease Control and Prevention (CDC) monitors influenza like illness (ILI) cases by collecting data from sentinel medical practices, collating the reports and publishing them on a weekly basis. There is a delay of 1-2 weeks between a patient is diagnosed and the moment that data point become available in CDC report. Their research goal was to predict ILI incidences before CDC. They collected tweets and location information from users who mentioned flu descriptors such as “flu”, “H1N1”, and “Swine Flu” in their tweets. They collected 4.7 million tweets from 1.5 million users for the period from Oct 2009 till Oct 2010. They have 31 weeks of data from CDC for the dataset. They remove all non US tweets, tweets from organization that posts multiple times in a day on flu related activities and retweets, resulting in 450,000 tweets. Tweets are split into 1 week time spans. Their model predicts data collected and published by the CDC, as a percentage of visits to sentinel physicians due to ILI in successive weeks. They get .2367 root mean squared error.

To detecting emerging topics in social streams, Takahashi et al. (2011) focus on social aspect of social networks i.e. links generated dynamically through replies, mentions, and retweets. Emerging topics are detected by calculating mention anomaly score of users. Their assumption is that an emerging topic is something people feel like discussing about, commenting about, or forwarding the information to their friends. Their approach is well suited for micro blogs such as twitter where the posts have very little textual information and in cases where the post is only an image with no textual data.

Budak et al. (2011) in their study, propose two novel structure based trend definition. They identify coordinated trends as trends where the trendiness of a topic is characterized by the number of connected pairs of users discussing it and uncoordinated trends as trends where the score of a topic is based on the number of unrelated people interested in it. They perform their experiments on a Twitter data set of 41.7 million nodes and 417 million posts achieving an average precision of 0.93 with a sampling probability of 0.005.

2.3 Sentiment Analysis

Sentiment analysis determines the sentimental attitude of a speaker or writer, thus it is important for companies, politicians, government, etc. to know how people feel about the products or services there are offering. Sentiment has three polarities – positive, negative, and neutral (Choi et al., 2010). Emotion detection in text is a difficult because of the ambiguity of language. Words, combination of words, special phrases, and grammar all play a role in formulating and conveying emotional information (Calix, 2011).

Osherenko (2008) in his research used the presence or absence of negations and intensifiers as features to train and test an emotion detection model. Tokuhisa et al. (2008) propose a model for detecting the emotional state of user that interacts with a dialog system. They use corpus statistics and supervised learning to detect emotion in text. They implement a two-step approach

where coarse grained emotion detection is performed first followed by fine grained emotion detection. Their work found that word n-gram features are useful for polarity classification.

To select lexical text features, Calix et al. (2010) proposes a methodology to automatically extract emotion relevant words from annotated corpora. The emotion relevant words are used as features in sentence level emotion classification with Support Vector Machines (SVM) and 5 emotion classes plus the neutral class.

Most lexical based sentiment detection models use POS tags (VB, NN, JJ, RB), exclamation points, sentence position in story, thematic role types, sentence length, number of POS tags, WordNet emotion words, positive word features, negative word features, actual words in the text, syntactic parses, etc. (Calix, 2011).

2.4 Controversy Detection

Pennacchiotti and Popescu (2010) focus on detecting controversies involving popular entities (i.e. proper nouns). Their controversy detection method detects controversies involving known entities in Twitter data. They use a sentiment lexicon of 7590 terms and a controversy lexicon composed of 750 terms. The controversy lexicon is composed of terms from the Wikipedia controversial topics list. Wikipedia's list of controversial issues is a list of previously controversial issues among Wikipedia editors. Wikipedia defines a controversial issue as one where its related articles are constantly being re-edited in a circular manner by various contributors expressing their person biases towards a subject (Wikipedia, 2012). Pennacchiotti and Popescu (2010) take a snapshot of tweets which contain a given entity within a time period. The snapshots which have most tweets discussing an entity (buzzy snapshots) are considered as likely to be controversial. They calculate the controversy score for each snapshot by combining historical controversy score and timely controversy score. The historical controversy score estimates the overall controversy level of an entity independent of time, while the timely controversy score estimates the controversy of an entity by analyzing the discussion among Twitter users in a given time period. The timely controversy score is a linear combination of two scores – MixSent(s) and contro(s). MixSent(s) reflects the relative disagreement about the entity in the Twitter data from the snapshot and contro(s) score (i.e. tweets with controversy term/total number of tweets within snapshot) reflects the presence of explicit controversy terms in tweets. Their gold standard contains 800 randomly sampled snapshots labeled by two expert editors of which 475 are non-event snapshots and 325 are event snapshots. Of the 325 event snapshots, 152 are controversial event snapshots, and 173 are non-controversial-event snapshot. Their experiment yields an average precision of 66% with the historical controversy score as baseline.

Choi et al. (2010) proposes a controversial issue detection method which considers the magnitude of sentiment information and the difference between the amounts of two different polarities. They perform their experiment using the MPQA corpus which contains manually

annotated sentiments for 10 topics consisting of 355 news articles. They measure the controversy of a phrase by its topical importance and sentiment gap it incurs. They first compute the score for positive and negative sentiment for a phrase and then determine if it is sufficiently controversial by checking if the sum of the magnitude of positive and negative sentiments is greater than a specified threshold value and also the difference between them is less than a specified threshold value. The precision of the proposed methodology is 83%.

In their research, Vuong et al. (2008) proposes three models to identify controversial articles in Wikipedia – the Basic model and two Controversy Rank models. The basic model only considers the amount of disputes within an article while the Controversy Rank (CR) models also consider the relationships between articles and the contributors. They thought a dispute in an article is more controversial, so the model utilizes the controversy level of disputes which can be derived from the articles' edit histories. The CR models define the article controversy score and the contributor controversy score. An article is controversial when it has lots of disputes among less contributors and a contributor is controversial when they are engaged in lots of disputes in less articles. They conduct their experiments on a dataset of 19,456 Wikipedia articles achieving an precision of 15%. This model can only be applied to Wikipedia since it is the only source in which contributors can edit others' work and the history of it is kept.

2.5 Corpora

MemeTracker (Leskovec et al., 2009) phrase cluster dataset contains clusters of memes. For each phrase cluster the data contains all the phrases in the cluster and a list of URLs where the phrases appeared. The MemeTracker dataset has been used in the tasks of meme tracking and information diffusion.

The Blog Authorship Corpus (Schier et al., 2006) consists of 681,288 posts collected from 19,320 bloggers in August 2004. Each blog is identified with the blogger's id, gender, age, industry and astrological sign. The Blog Authorship Corpus has been used in Data Mining and Sentiment Analysis.

TREC Tweets2011 dataset consists of identifiers, provided by Twitter, for approximately 16 million tweets sampled between January 23 and February 8, 2011 (Tweets2011, 2011).

ICWSM 2011 Spinn3r dataset consists of over 386 million blog posts, news articles, classified, forum posts and social media content between January 13th and February 14th, 2011. The content includes the syndicated text, its original HTML, annotations and metadata (Burton et al., 2011).

Reuters-21578 (Lewis) dataset contains 21578 documents which appeared on Reuters newswire in 1987. The documents are annotated with topics and entities. The Reuters-21578 dataset is used in information retrieval, machine learning, and other corpus-based research.

MPQA Opinion Corpus contains news articles from a wide variety of news sources manually annotated for opinions and other private state such as beliefs, emotions, sentiments, speculations, etc. (MPQA)

Chapter 3: Methodology

The methodology consisted of three major phases. In the first phase, articles and comments posted on them were collected (Section 3.1) and annotated (Section 3.1.1, page 17) to create an annotated corpus. Pre-processing was also performed to remove URLs and stop words from the data. In the second phase (Section 3.2, page 19) a machine learning model was developed to detect controversial trends, including identification, calculation, analysis, and extraction of features including sentiment and controversy scores. The third phase was analysis and improvement of performance of the model, discussed in Chapter 4, page 21.

3.1 Data Collection

This research involved development of a new controversy corpus. The corpus consists of comments made by viewers on 728 articles published by Cable News Network (CNN) on its online news portal¹. CNN is a broadcast news company based in the U.S. offering world news on its cable T.V. channel as well as on its website. CNN.com utilizes the Disqus Plugin to permit readers to post comment and provide feedback on their news stories. Disqus is an online discussion and commenting service provider for websites. It allows users to comment by logging in using their existing accounts with other social media websites (i.e. Twitter, Facebook, Google+, etc.) without having to create a new user account (Disqus, 2012).

To collect data, an application was created using VB.NET. A screenshot of the application is shown in Figure 1. The application collected a list of news articles posted on CNN.com using Disqus API². Using the list of news articles, CNN.com was crawled to gather the articles' text. Comments posted on the articles and information about the users who commented was collected by making calls to the Disqus API.

The comments were accessed for an article via the Disqus API using [http://disqus.com/api/3.0/threads/listPosts.json?api_key=\[api_key\]&forum=cnn&limit=100&thread=link:\[article_url\]](http://disqus.com/api/3.0/threads/listPosts.json?api_key=[api_key]&forum=cnn&limit=100&thread=link:[article_url]). Disqus API works over the HTTP protocol as a REST web service. When a GET request is sent, the API returns data in JSON format. There is a limit of 1000 request per day with each request containing a maximum of 100 objects. In the returned JSON a cursor is provided for the next set of posts. An example of comments data returned for the article titled *"Rescuers search for missing after deadly Hong Kong ferry crash"* is displayed in Figure 2. Data elements returned by the Disqus API are displayed in Table 1. All information retrieved was stored in an SQL database. The Entity Relationship Diagram (ERD) of the database is displayed in Figure 3.

¹ <http://cnn.com>

² <http://disqus.com/api/docs/>

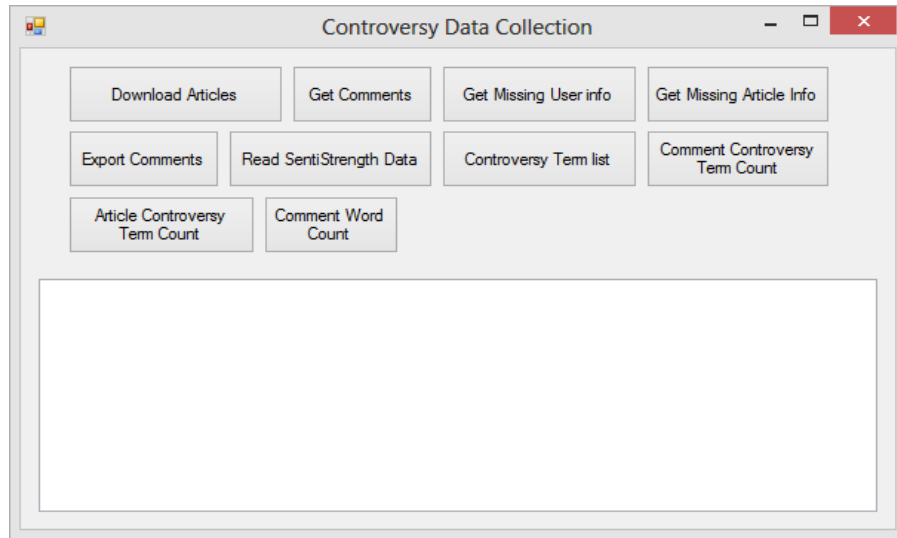


Figure 1: Data collection

```
{
  "isJuliaFlagged": true,
  "isFlagged": false,
  "forum": "cnn",
  "parent": 668477936,
  "author": {
    "username": "cnn-9af9ffd566cd9d85bf86e7ee42",
    "about": "",
    "name": "cernventure",
    "url": "#",
    "joinedAt": "2012-08-27T17:34:00",
    "rep": 1.236586,
    "profileUrl": "http://disqus.com/cnn-9af9ffd566cd9d85bf86e7ee42/",
    "emailHash": "60ac121c9303a83d2e1830abff849079",
    "reputation": 1.236586,
    "location": "",
    "isPrimary": true,
    "isAnonymous": false,
    "id": "31217376",
    "avatar": {
      "small": {
        "permalink": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar32.jpg?1327959337",
        "cache": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar32.jpg?1327959337"
      },
      "large": {
        "permalink": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar92.jpg?1327959337",
        "cache": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar92.jpg?1327959337"
      },
      "permalink": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar92.jpg?1327959337",
      "cache": "http://mediacdndisqus.com/uploads/forums/19/3157/avatar92.jpg?1327959337"
    }
  },
  "media": [
  ],
  "isDeleted": false,
  "isApproved": true,
  "dislikes": 0,
  "raw_message": "\u00a0sorry for such an accident in this blazing time in the world .condolence for the families ",
  "createdAt": "2012-10-04T06:07:53",
  "id": "670620865",
  "thread": "867614535",
  "numReports": 0,
  "likes": 0,
  "isEdited": false,
  "points": 0,
  "message": "<p>\u00a0sorry for such an accident in this blazing time in the world .condolence for the families </p>",
  "isSpam": false,
  "isHighlighted": false,
  "userScore": 0
},
```

Figure 2: Example of comments returned by the Disqus API

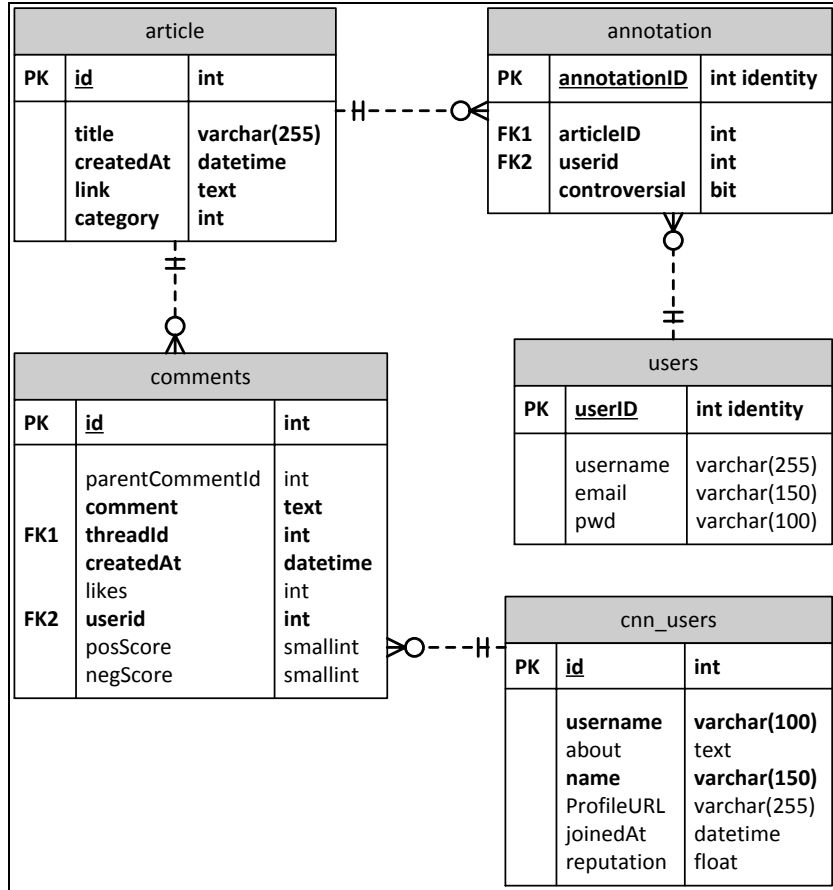


Figure 3: Database ERD

Table 1: Data elements returned by the Disqus API

Author	
username	user name of the author
id	user id of the author
name	name of the author
about	text about the author
url	URL to the author's profile page
reputation	reputation score of the author
Comment	
message	comment text
id	id of the comment
parent	id of the original comment if the comment is a reply to another comment
likes	number of likes

3.1.1 Annotation

To aid in annotation, a web application was created. Each article was annotated by at least 3 annotators. There were a total of 20 annotator from various educational backgrounds – 2 from business, 2 from education, 3 from engineering, 6 from humanities, 3 from sciences, 4 from social sciences. The annotators were given the definition of controversy (see Figure 4) and were instructed to identify which articles they think are controversial. The annotators were displayed articles along with their respective comments (see Figure 5). For each article, the annotators classified whether the article is controversial or not. When there was a conflict between annotators in classifying an article, then a voting scheme was be used where the class with the majority votes won. Inter-annotator agreement statistics are discussed in Chapter 4, page 19.

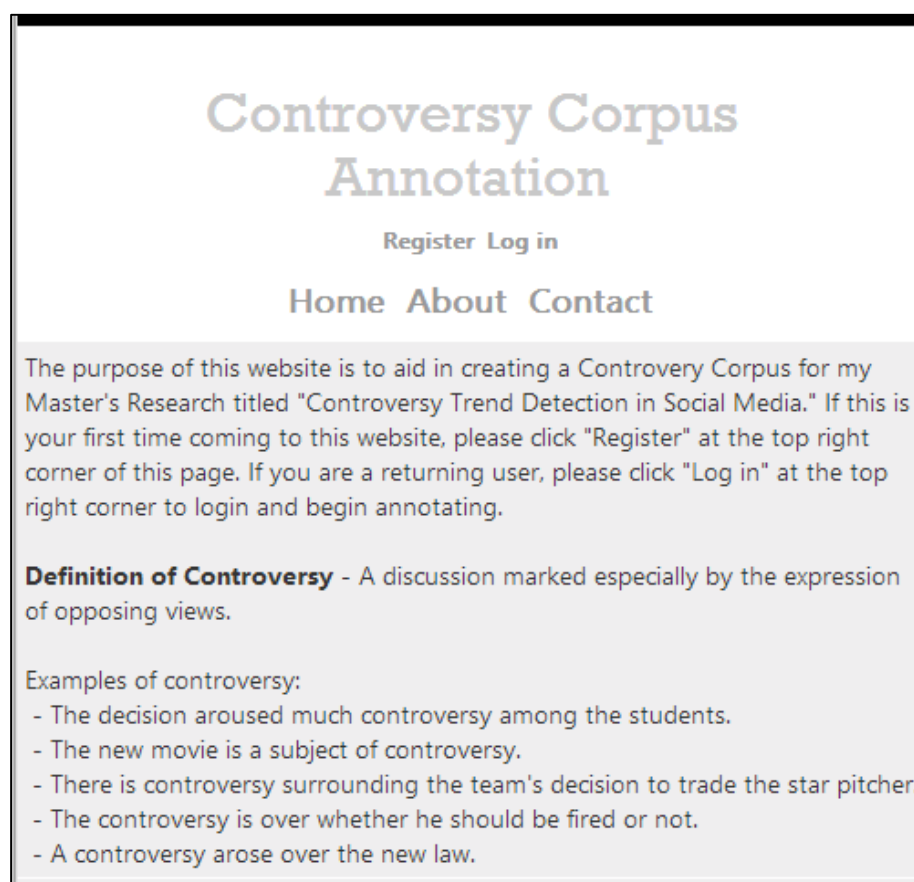


Figure 4: Definition of controversy

The annotations were stored in the "annotation" table with the userid of the annotator, articleID of the article being annotated, and classification made by the annotator as shown in Figure 3. When the annotator classified an article as controversial then 1 was stored in the "controversial" column, otherwise a 0 was stored. An Institutional Review Board (IRB) exemption from institutional oversight was obtained (see Appendix B, page 44).

Controversy Corpus Annotation
Home
AL

Dozens dead after passenger vessels collide off Hong Kong

At least 36 people have died after two passenger boats collided off Hong Kong's Lamma Island on Monday evening, throwing more than 100 people into the sea.

The crash happened on a night when Hong Kong's busy waters were even more crowded than usual, as the city celebrated China's National Day.

One of the vessels, owned by The Hong Kong Electric Company, was carrying company employees and their families to watch the scheduled fireworks display when it was struck by a passenger ferry traveling from Hong Kong Island to Lamma.

Government officials said the collision occurred off Lamma's coast at around 8:20 p.m. local time.

"I thought we'd hit a rock or a lighthouse," said Chris Head, a school teacher who was on the passenger ferry that crashed into the Hong Kong Electric vessel. He said the force of the impact threw him out of his seat at the back of the ferry, which was not very full of people.

Comments:

username	comment	Posted on
cnn-5c6cef9521a227ebce6216f392	I am sorry about those who lost their lives in this terrible accident. Obviously one of the captains is to blame for this, but the other also holds blame. Even if you have right of way, it is incumbent to AVOID COLLISIONS. It's not rocket science. Obviously one or both were not paying attention to their duties.	10/2/12 7:31 AM
cnn-94f19d5e46f57ecfd89fb8ed44	Such a tragic loss of life. My heart goes out to all those affected. I have lots of friends in Hong Kong. I hope none of them were among the victims.	10/2/12 8:04 AM

Annotation:
☐ Controversial
☐ Non-Controversial
Next

Figure 5: Example of the annotation process

3.1.2 Preprocessing

After the data was collected, URLs from the comments were removed. Some articles containing only an image gallery were removed since there was no textual information in the article. A total of 72 articles were removed from the original 800 collected bringing the total number of articles in the corpus to 728. Since stop words do not provide any information and introduce noise, they were removed using a stop word list. The preprocessing application was developed in the windows environment in VB.NET.

3.2 Controversy Trend Detection

To detect which articles are controversial, the sentiment of the comment text was analyzed using SentiStrength (Thelwall et al., 2012) to classify whether the viewer is expressing positive, negative, or neutral. After the sentiment classification was done, a controversy score was calculated and other features - number of posts, post rate, number of posts with positive sentiment, number of posts with negative sentiment, number of 'Likes', number of posts with responses from other viewers, controversy term count in article text, sum of controversy term count in comments text, average comment word count, sum of user reputation score, total number of users, total number of new users since article post were extracted. The controversy score was calculated by dividing total number of negative comments the ratio of negative sentiment count to positive sentiment count. Pseudocode for calculating the controversy score and the post rate of an article is shown in Figure 6. The presence of controversy terms in comments and articles' texts were done by creating a controversy term list from "Wikipedia: List of controversial issues"³.

To predict how soon a controversial article can be detected, the comments for their respective articles were divided into time spans of 6hr, 12hr, 18hr, 24hr, 30hr, 36hr, 42hr and 48hr. For each of the time spans, features were extracted and a Decision Table classifier was trained and tested to see how well the classifier performs using the features extracted from comments belonging to a specific time interval. All the features were normalized between 0 and 1.

Features used for this research are listed below with features unique to this research marked with an asterisk:

- Comment count(t) – total number of comments by time t
- Comment post rate(t) – post rate of comments per hour by time t
- Likes(t)* – number of 'Likes' by time t
- Threaded comments count(t)* – number of comments which have responses from other users by time t
- Positive sentiment count(t) *– total number of positive comments by time t
- Negative sentiment count(t)* – total number of negative comments by time t
- Controversy score(t)* – negative sentiment comment count divided by sum of positive sentiment comments count and negative sentiment comments count by time t
- Reputation(t)* - aggregate reputation scores of users who posted on an article by time t
- Number of users(t)* - total number of users who commented on an article by time t
- New users(t)* - total number of new users who commented on an article by time t
- Article controversy term count* - number of controversy terms that appear in the article text

³ http://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues

- Comment controversy term count(t)* - total number of controversy terms that appear in all the comments posted on an article by time t
- Word count(t)* - average word count of comments posted on an article by time t

The sentiment scores for comments were calculated using an application called SentiStrength, which estimates the strength of positive and negative sentiment in short texts. SentiStrength reports two sentiment strength: -1 (not negative) to -5 (extremely negative) and 1 (not positive) to 5 (extremely positive) (Thelwall et al., 2012). The sentiment strength scores returned by SentiStrength were stored in the comments table as shown in Figure 3, page 16.

```

For each article s:
  PosCount = 0
  NegCount = 0
  For each comment c made on article s:
    If PosScore(c) > |NegScore(c)|
      then PosCount = PosCount + 1
    else if PosScore(c) < |NegScore(c)|
      then NegCount = NegCount + 1

ContScore(s) = NegCount / (PosCount + NegCount)

PostRate(s) = number of comments / total time in hours

```

Figure 6: Pseudocode for calculating controversy score and post rate

Chapter 4: Results and Analysis

To predict how soon an article can be detected as being controversial or not, the data was divided into 6 hour time spans - 6hr, 12hr, 18hr, 24hr, 30hr, 36hr, 42hr and 48hr. For each time span features were extracted from comments posted within the time interval (e.g. in 6hr time span, features were extracted from comments posted within the first six hours of the article being posted). From the controversy corpus containing 728 articles, 664 articles were chosen since 64 articles did not have any comments in the first six hours of the article being posted.

Each of the time spans contained features for 664 articles, of which 365 articles were controversial and 299 were non-controversial. Decision Table classifier was used for training and testing. Ten-fold cross validation was used to minimize the impact of specific case selection on performance results. Performance as a function of time span was measured. F-Measure was used to measure the performance of the methodology. Formulae for F-score calculation is displayed in Figure 7.

$$\begin{aligned} Fscore &= 2 * \frac{Precision * Recall}{Precision + Recall} \\ Precision &= \frac{True\ Positives}{True\ Positives + False\ Positives} \\ Recall &= \frac{True\ positives}{True\ Positives + False\ Negatives} \end{aligned}$$

Figure 7: F-Score calculation

4.1 Controversy Corpus

To measure the quality of the annotation process, inter-annotator agreement Kappa (κ) was calculated. The calculation is based on the difference between how much agreement among annotators is actually present compared to how much agreement would be expected to be present by chance alone. The formula for calculating κ is shown below.

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

\bar{P} denotes the mean value of P_i 's, where $1 \leq i \leq N$, is the extent to which annotators agree for the i^{th} classification. Whereas \bar{P}_e denotes the sum of squares of pj 's, where $1 \leq j \leq k$, is the proportion of all the assignments which were to the i^{th} sample. Therefore, $\bar{P} - \bar{P}_e$ gives the

degree of agreement actually present and $1 - \bar{P}_e$ gives the degree of agreement that is attainable above chance.

A κ value of 0.33 was obtained. The value of κ was compared with the interpretation of Kappa first studied by Landis and Koch (1977) as shown in Table 2. The obtained value of κ can be interpreted as fair agreement between annotators.

Table 2: Interpretation of Kappa

κ	Interpretation
< 0	Poor agreement
0.01 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

In 50% or 363 articles, there was perfect agreement among all annotators. 365 articles did not have perfect agreement among annotators. Of the 365 articles, 136 were related to politics – 78 related to U.S. presidential election and 58 were related to foreign politics. Rest of 229 articles which had imperfect agreement were related to terrorism, bombings, mass shootings, sexual abuse, environment, and sports. The disagreement among annotators could be because, they were not given any training as to what to look for in order to determine controversy. Most of the articles with imperfect agreement, had factual news articles which had many comments where comments' authors were expressing conflicting viewpoints against each other. Some annotators were looking for disagreement among commenters to decide if the article is controversial or not while others were gauging whether the viewpoints expressed in the article were controversial or not to them personally (see Appendix C, page 46).

There were 389 articles with 6,547 comments in the first hour and 726 articles with 489,430 comments in the first forty-eight hours after an article was posted as shown in Figure 8. Normal distribution can be fitted on histogram of comments. The parameters are calculated and $N(10196.46, 9168.604)$ fitted on the data. In this figure the intervals are not same but the effect of time and number of articles studied tells us that most of the comments were posted in the twelve hour time interval which proves that normal distribution is appropriate choice for this data.

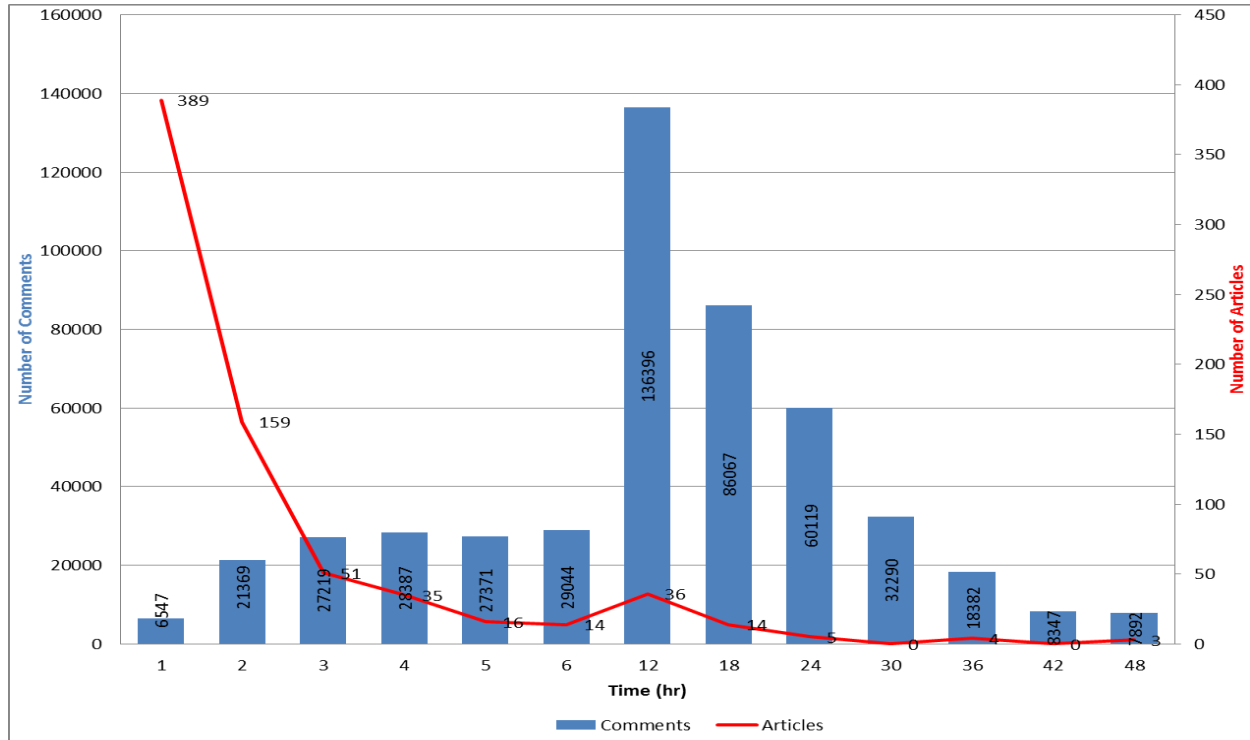


Figure 8: Distribution of Comments

4.2 Controversy Trend Detection

4.2.1 Feature Analysis

Features were evaluated for importance using Weka's Chi-squared Ranking attribute evaluator. It evaluates the worth of an attribute by computing the value of the chi-squared statistic with respect to the class. Controversy score feature had the highest contribution to controversy detection as shown in Table 3. All the features which had a Chi score above zero were used.

Table 3: Feature ranking

Rank	Chi	Feature
1	124.7125	Controversy score
2	79.5735	Comment controversy term count
3	59.6204	Negative comment count
4	58.2435	Word count
5	55.6467	Reputation
6	55.4588	Comment post rate
7	55.4588	Comment count
8	53.0733	Threaded comment count
9	48.382	Number of users
10	41.8239	Likes
11	38.6419	Positive comment count
12	26.8812	Article controversy term count
13	19.4174	New users

4.2.2 Classification Model Performance

Classifiers were trained using Weka (Hall et al., 2009), a machine learning software written in Java developed at the University of Waikato, New Zealand. The extracted features were stored in a text file where each line consisted of a sample. Each sample contained features vectors followed by the sample's classification separated by tab. The text file was used in Weka to train and test the Classifiers. There were two classes – controversial and non-controversial. To compare performance of different classifiers, a dataset with features extracted at six hours from 664 articles consisting of 139,937 comments posted on them was used.

Performance was compared between Naïve Bayes, Support Vector Machine (SVM), Random Forest, and Decision Table classifiers. Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong independence assumption of features. SVM is a supervised learning approach that optimizes the margin that separates the data. Random Forest operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes output by individual trees. Decision Table is a rule based classifier which associates conditions with actions to perform; it does not assume that the attributes are independent.

Decision Table classifier was used for training and testing across all time spans, since it gave the best performance (68.7%) when compared to Naïve Bayes (47.7%), SVM (50.7%), and Random Forest (66.8%). Summary of performance comparison between Naïve Bayes, SVM, Random Forest, and Decision Table classifiers is shown in Table 4.

Table 4: Performance comparison between different classifiers at six hours

	Naïve Bayes	SVM	Random Forest	Decision Table
Non-controversial	0.494	0.746	0.619	0.648
Controversial	0.804	0.579	0.714	0.720
Overall	0.477	0.507	0.668	0.687

Classification at 0 to 5 hours was also done to see the models performance in shorter time frames. Features were extracted at each time span (i.e. time = 0, 1, 2, 3, 4, 5) and Decision Table classifier was used for training and testing. The performance ranged from an F-score of 60.4% at zero hour to an F-score of 68.7% at five hours. Summary of the classification by class is shown in Table 5 and the confusion matrix is shown in Table 6.

Table 5: Classification after equal class sample sizes

Zero Hours (0 comments)		
Class	F-score	Sample Size
Non-controversial	0.591	341
Controversial	0.631	341
Overall	0.604	
One Hour (6,547 comments)		
Class	F-score	Sample Size
Non-controversial	0.578	155
Controversial	0.584	155
Overall	0.58	
Two Hours (27,916 comments)		
Class	F-score	Sample Size
Non-controversial	0.604	224
Controversial	0.637	224
Overall	0.616	
Three Hours (55,135 comments)		
Class	F-score	Sample Size
Non-controversial	0.659	255
Controversial	0.659	255
Overall	0.659	
Four Hours (83,522 comments)		
Class	F-score	Sample Size
Non-controversial	0.685	280
Controversial	0.706	280
Overall	0.694	
Five Hours (110,893 comments)		
Class	F-score	Sample Size
Non-controversial	0.676	289
Controversial	0.699	289
Overall	0.687	

Table 6: Confusion matrix of classification after equal class sample sizes

Zero Hour		
Non-controversial	Controversial	← classified as
238	103	Non-controversial
165	176	Controversial
One Hour		
Non-controversial	Controversial	← classified as
93	62	Non-controversial
68	87	Controversial
Two Hours		
Non-controversial	Controversial	← classified as
154	70	Non-controversial
101	123	Controversial
Three Hours		
Non-controversial	Controversial	← classified as
168	87	Non-controversial
87	168	Controversial
Four Hours		
Non-controversial	Controversial	← classified as
202	78	Non-controversial
93	187	Controversial
Five Hours		
Non-controversial	Controversial	← classified as
207	82	Non-controversial
99	190	Controversial

Classification at Six Hours

Features were extracted for 664 articles from 139,937 comments that were posted on articles in the first six hours. Using the Decision Table classifier yielded an overall F-score of 68.7% F-score was obtained through 10-fold cross validation. 456 samples were correctly classified and 208 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 71.1% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Classification at Twelve Hours

Features were extracted for 664 articles from 275,774 comments that were posted on articles in the first twelve hours. Using the Decision Table classifier yielded an overall F-score of 70.8% F-score was obtained through 10-fold cross validation. 471 samples were correctly classified and 193 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 71.4% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Table 7: Classification

Six Hours (139,937 comments)		
Class	F-score	Sample Size
Non-controversial	0.648	299
Controversial	0.72	365
Overall	0.687	
Twelve Hours (275,774 comments)		
Class	F-score	Sample Size
Non-controversial	0.692	299
Controversial	0.722	365
Overall	0.708	
Eighteen Hours (361,585 comments)		
Class	F-score	Sample Size
Non-controversial	0.681	299
Controversial	0.734	365
Overall	0.711	
Twenty-Four Hours (421,320 comments)		
Class	F-score	Sample Size
Non-controversial	0.668	299
Controversial	0.714	365
Overall	0.694	
Thirty Hours (452,583 comments)		
Class	F-score	Sample Size
Non-controversial	0.681	299
Controversial	0.714	365
Overall	0.699	
Thirty-Six Hours (468,813 comments)		
Class	F-score	Sample Size
Non-controversial	0.712	299
Controversial	0.723	365
Overall	0.718	
Forty-Two Hours (475,680 comments)		
Class	F-score	Sample Size
Non-controversial	0.705	299
Controversial	0.73	365
Overall	0.718	
Forty-Eight Hours (481,998 comments)		
Class	F-score	Sample Size
Non-controversial	0.692	299
Controversial	0.722	365
Overall	0.708	

Classification at Eighteen Hours

Features were extracted for 664 articles from 361,585 comments that were posted on articles in the first eighteen hours. Using the Decision Table classifier yielded an overall F-score of 71.1% F-score was obtained through 10-fold cross validation. 472 samples were correctly classified and 192 were misclassified. Summary of the classification by class is shown in Table 7 and the

confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 71.1% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Table 8: Confusion Matrix

Six Hours		
Non-controversial	Controversial	← classified as
199	100	Non-controversial
108	257	Controversial
Twelve Hours		
Non-controversial	Controversial	← classified as
191	108	Non-controversial
85	280	Controversial
Eighteen Hours		
Non-controversial	Controversial	← classified as
201	98	Non-controversial
94	271	Controversial
Twenty-Four Hours		
Non-controversial	Controversial	← classified as
191	108	Non-controversial
95	270	Controversial
Thirty Hours		
Non-controversial	Controversial	← classified as
188	111	Non-controversial
88	277	Controversial
Thirty-Six Hours		
Non-controversial	Controversial	← classified as
188	111	Non-controversial
76	289	Controversial
Forty-Two Hours		
Non-controversial	Controversial	← classified as
194	105	Non-controversial
81	284	Controversial
Forty-Eight Hours		
Non-controversial	Controversial	← classified as
191	108	Non-controversial
85	280	Controversial

Classification at Twenty-Four Hours

Features were extracted for 664 articles from 421,320 comments that were posted on articles in the first twenty-four hours. Using the Decision Table classifier yielded an overall F-score of 69.4% F-score was obtained through 10-fold cross validation. 461 samples were correctly classified and 203 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 72.4% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Classification at Thirty Hours

Features were extracted for 664 articles from 452,583 comments that were posted on articles in the first thirty hours. Using the Decision Table classifier yielded an overall F-score of 69.9% F-score was obtained through 10-fold cross validation. 465 samples were correctly classified and 199 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 70.4% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Classification at Thirty-Six Hours

Features were extracted for 664 articles from 468,813 comments that were posted on articles in the first thirty-six hours. Using the Decision Table classifier yielded an overall F-score of 71.8% F-score was obtained through 10-fold cross validation. 477 samples were correctly classified and 187 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 71.6% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Classification at Forty-Two Hours

Features were extracted for 664 articles from 475,680 comments that were posted on articles in the first forty-two hours. Using the Decision Table classifier yielded an overall F-score of 71.8% F-score was obtained through 10-fold cross validation. 478 samples were correctly classified and 186 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 71.7% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Classification at Forty-Eight Hours

Features were extracted for 664 articles from 481,998 comments that were posted on articles in the first forty-eight hours. Using the Decision Table classifier yielded an overall F-score of 70.8% F-score was obtained through 10-fold cross validation. 471 samples were correctly classified and 193 were misclassified. Summary of the classification by class is shown in Table 7 and the confusion matrix is shown in Table 8. After removing 66 random samples from the controversial class to have same number of sample sizes per class, an F-score of 70.6% was obtained as shown in Table 9 and the confusion matrix is shown in Table 10.

Table 9: Classification after equal class sample sizes

Six Hours (111,352 comments)		
Class	F-score	Sample Size
Non-controversial	0.714	299
Controversial	0.707	299
Overall	0.711	
Twelve Hours (227,687 comments)		
Class	F-score	Sample Size
Non-controversial	0.719	299
Controversial	0.709	299
Overall	0.714	
Eighteen Hours (301,803 comments)		
Class	F-score	Sample Size
Non-controversial	0.703	299
Controversial	0.719	299
Overall	0.711	
Twenty-Four Hours (354,660 comments)		
Class	F-score	Sample Size
Non-controversial	0.723	299
Controversial	0.725	299
Overall	0.724	
Thirty Hours (381,162 comments)		
Class	F-score	Sample Size
Non-controversial	0.709	299
Controversial	0.699	299
Overall	0.704	
Thirty-Six Hours (395,456 comments)		
Class	F-score	Sample Size
Non-controversial	0.733	299
Controversial	0.701	299
Overall	0.716	
Forty-Two Hours (401,414 comments)		
Class	F-score	Sample Size
Non-controversial	0.718	299
Controversial	0.717	299
Overall	0.717	
Forty-Eight Hours (407,251 comments)		
Class	F-score	Sample Size
Non-controversial	0.714	299
Controversial	0.698	299
Overall	0.706	

Table 10: Confusion matrix of classification after equal class sample sizes

Six Hours		
Non-controversial	Controversial	← classified as
210	89	Non-controversial
84	215	Controversial
Twelve Hours		
Non-controversial	Controversial	← classified as
210	89	Non-controversial
82	217	Controversial
Eighteen Hours		
Non-controversial	Controversial	← classified as
218	81	Non-controversial
92	207	Controversial
Twenty-Four Hours		
Non-controversial	Controversial	← classified as
217	82	Non-controversial
83	216	Controversial
Thirty Hours		
Non-controversial	Controversial	← classified as
207	92	Non-controversial
85	214	Controversial
Thirty-Six Hours		
Non-controversial	Controversial	← classified as
203	96	Non-controversial
74	225	Controversial
Forty-Two Hours		
Non-controversial	Controversial	← classified as
214	85	Non-controversial
84	215	Controversial
Forty-Eight Hours		
Non-controversial	Controversial	← classified as
205	94	Non-controversial
82	217	Controversial

4.2.3 Discussion

The classifier achieved F-score ranging from 68% to 72% across different time spans. Samples were distributed across two classes unevenly. Because of this class imbalance, the overall F-scores across the two classes were thought to be lower than what could be achieved if the each class had equal number of samples. After removing 66 random samples from the controversial class to make the class sample sizes equal, the performance at 6hrs increased from 68.7% to 71.1%. The overall performance across all time spans improved a little after making the two classes to have equal number of samples from 70.5% to 71.3%.

Analysis of misclassified samples was done for the six hour time span which consisted of 664 articles with 139,937 comments posted on them. Full training set was used for training and testing using Decision Tree classifier. There were 178 samples which were misclassified of which 100 non-controversial samples were misclassified as controversial and 78 controversial samples were misclassified as non-controversial. Of the 178 misclassified samples, 109 samples were the ones on which the annotators had disagreement about the classification. Of the 109 samples with annotator disagreement, 48 were non-controversial and 61 were controversial.

By analyzing annotators' algorithms (see Appendix C) for classifying an article as controversial versus non-controversial, a category feature could be used to identify what the general topic of the article is, since 33% of annotators were looking to see whether the topic of the article was related to politics or other controversial issues.

Three different algorithms were used by 20 annotators to detect controversy – annotator's personal views on whether the topic was controversial plus disagreement among commentators (56%), whether the topic is related to politics or controversial issues such as abortion or gay marriage and the number of comments (33%), and whether the article was factual news reporting or presented the author's personal opinion plus disagreement among commentators (11%).

Additional features such as the ratio of comments consisting of controversy terms to the total number of comments, frequency of controversy terms in comments per hour to indicate the rate of controversy terms being used in comments, aggregate of positive comments' sentiment scores to indicate the aggregate strength of emotions expressed in positive comments, and aggregate of negative comments' sentiment scores to indicate the aggregate strength of emotions expressed in negative comments.

The features used in this research did not account for strength of emotions expressed by commentators to indicate how expressive the comments were. While interacting online in social media, people sometimes type text in all capital letter to emphasize their view point which should also be taken into consideration to further improve the classification performance.

The results indicate that it is possible for early prediction of controversy trends in social media within the first three hours of an article being posted.

Chapter 5: Conclusion and Future Research

Controversial trend detection is important to companies, governments, national security agencies, journalists, and marketing groups so that they identify which issues the public is having problems with and develop strategies to remedy them. Till date, researchers have done work on detecting controversial topics in Wikipedia and Twitter with precisions ranging from 15% to 83%. This research involved detecting controversial articles using sentiment strength of comments. It also tested the usefulness of new features in the detection of controversial trends using social media. This study achieved an average F-score of 71.3% across all time spans, indicating that it is possible for early prediction of controversy trends in social media. An annotated corpus was created to aid future research in controversy trend detection in social media.

The controversy corpus had 664 samples, with 299 being non-controversial and 365 controversial. The corpus should further be developed to have more samples with samples divided equally across the two classes. A separate training and testing sets should be created. The training set should be used to train a model and the testing set for testing the model to gauge how the algorithm performs in the real world.

Additional features such as the ratio of comments consisting controversy terms to the total number of comments to consider how many controversial comments are made compared to non-controversial comments, frequency of controversy terms in comments per hour to take into consideration the burstiness of controversial comments, aggregate of positive comments' sentiment scores, and aggregate of negative comments' sentiment scores to account for the strength of positive and negative emotions expressed in comments could be used to improve the classification performance.

Furthermore, trolls on the internet who posts inflammatory, extraneous, or off-topic repeated messages should also be taken into consideration to avoid misclassification. Trolls can be detected by checking to see if there is more than one message which contains the same exact text.

SentiStrength's emotion dictionary weights can be further optimized by using the six datasets provided on SentiStrenght's website⁴. The six datasets include at least 1000 human coded texts each on comments posted on MySpace, BBC, Digg, Runners World, Twitter, and Youtube. Each text is coded by three independent coders. The provided datasets will best optimize the sentiment strength scores of texts posted in social media since the texts used in the dataset are short informal text collected from comments posted by users in online discussion forums, news articles, Twitter, and Youtube.

⁴ <http://sentistrength.wlv.ac.uk/documentation/6humanCodedDataSets.zip>

References

- Achrekar, H., Gandhe, A., Lazarus, R., Ssu-Hsin, Y., & Benyuan, L. (2011). *Predicting Flu Trends using Twitter data*. In 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pages 702-707.
- Alvanaki, F., Michel, S., Ramamritham, K., & Weikum, G. (2012). *See What's enBlogue - Real-time Emergent Topic Identification in Social Media*. In 15th International Conference on Extending Database Technology (EDBT 2012), Berlin, Germany, pages 336-347.
- Bock, G., & Kim, Y. (2002). Breaking the Myths of Rewards: An Exploratory Study of Attitudes about Knowledge Sharing, pages 14-21.
- Brennan, M., & Greenstadt, R. (2011). *Coalescing Twitter Trends: The Under-Utilization of Machine Learning in Social Media*. In 2011 IEEE Third International Conference on Social Computing (socialcom), pages 641-646.
- Budak, C., Agrawal, D., & Abbadi, A. E. (2011). *Structural trend analysis for online social networks*. *Proc. VLDB Endow.*, 4(10), pages 646-656.
- Burton, K., Kasch, N., & Soboroff, I. (2011). *The ICWSM 2011 Spinn3r Dataset*. In Fifth Annual Conference on Weblogs and Social Media (ICWSM 2011), Barcelona, Spain.
- Calix, R. (2011). *Automated Semantic Understanding of Human Emotions in Writing and Speech*. PhD, Louisiana State University Baton rouge, LA.
- Calix, R., Mallepudi, S. A., Bin, C., & Knapp, G. M. (2010). *Emotion Recognition in Text for 3-D Facial Expression Rendering*. *IEEE Transactions on Multimedia*, 12(6), pages 544-551. doi: 10.1109/tmm.2010.2052026
- Cataldi, M., Caro, L. D., & Schifanella, C. (2010). *Emerging topic detection on Twitter based on temporal and social terms evaluation*. In Proceedings of the Tenth International Workshop on Multimedia Data Mining, Washington, D.C., pages 1-10.
- Choi, Y., Jung, Y., & Myaeng, S.-H. (2010). *Identifying Controversial Issues and Their Sub-topics in News Articles*. In the Proceedings of the 2010 Pacific Asia conference on Intelligence and Security Informatics, pages 140-153.
- Colbaugh, R., & Glass, K. (2011). *Detecting emerging topics and trends via predictive analysis of 'meme' dynamics*. In 2011 IEEE International Conference on Intelligence and Security Informatics (ISI), pages 192-194.
- Cvijikj, I. P., & Michahelles, F. (2011). *Monitoring Trends on Facebook*. In 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing (DASC), pages 895-902.
- Deci, E. L. (1975). *Intrinsic Motivation*. New York, NY: Plenum Press.
- Deci, E. L., & Ryan, R. M. (1987). *Accessibility and stability of predictors in the theory of planned behavior*. *Journal of Personality and Social Psychology*, 63(5), pages 754-765.

- Disqus. (2012). Disqus Retrieved 7/12/2012, 2012, from <http://disqus.com>
- Fabian Abel, C. H., Geert-Jan Houben, Richard Stronkman, Ke Tao. (2012). *Semantics + Filtering + Search = Twitcident. Exploring Information in Social Web Streams*. In International Conference on Hypertext and Social Media Milwaukee, USA, pages 285-294.
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Glass, K., & Colbaugh, R. (2010). *Toward Emerging Topic Detection for Business Intelligence: Predictive Analysis of 'Meme' Dynamics*. *CoRR*, abs/1012.5994.
- Gloor, P. A., Krauss, J., Nann, S., Fischbach, K., & Schoder, D. (2009). *Web Science 2.0: Identifying Trends through Semantic Social Network Analysis*. In Computational Science and Engineering, 2009. CSE '09. International Conference on, pages 215-222.
- Goorha, S., & Ungar, L. (2010). *Discovery of significant emerging trends*. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, Washington, DC, USA, pages 57-64.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). *The WEKA data mining software: an update*. *SIGKDD Explor. Newsl.*, 11(1), pages 10-18. doi: 10.1145/1656274.1656278
- Heum, P., Eunsun, K., Kuk-Jin, B., Hyuk, H., Tae-Eung, S., & Hyuk-Chul, K. (2011). *Detection and Analysis of Trend Topics for Global Scientific Literature Using Feature Selection Based on Gini-Index*. In 2011 23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI), pages 965-969.
- Hsu, C.-L., & Lin, J. C.-C. (2008). *Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation*. *Information & Management*, 45(1), pages 65-74. doi: 10.1016/j.im.2007.11.001
- Jeonghee, Y. (2005). *Detecting buzz from time-sequenced document streams*. In Proceedings from the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service, pages 347-352.
- Kim, S.-M., & Hovy, E. (2006). *Extracting opinions, opinion holders, and topics expressed in online news media text*. In Proceedings of the Workshop on Sentiment and Subjectivity in Text, Sydney, Australia, pages 1-8.
- Kontostathis, A., Galitsky, L., Pottenger, W. M., Roy, S., & Phelps, D. J. (2003). *A survey of emerging trend detection in textual data mining*. *A Comprehensive Survey of Text Mining*.
- Ku, L. W., Liang, Y. T., & Chen, H. H. (2006). *Opinion extraction, summarization and tracking in news and blog corpora*. In Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs, pages 100-107.

- Landis, J. R., & Koch, G. G. (1977). *MEASUREMENT OF OBSERVER AGREEMENT FOR CATEGORICAL DATA*. [Article]. *Biometrics*, 33(1), pages 159-174. doi: 10.2307/2529310
- Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). *Meme-tracking and the dynamics of the news cycle*. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, Paris, France, pages 497-506.
- Lewis, D. D. Test Collections: Reuters-21578 Retrieved 10/23/2011, 2011, from <http://www.daviddlewis.com/resources/testcollections/reuters21578/>
- Mahdavi, F., Ismail, M. A., & Abdullah, N. (2009). *Semi-Automatic Trend Detection in Scholarly Repository Using Semantic Approach*. In World Academy of Science, Engineering and Technology, pages 224-226.
- Mathioudakis, M., & Koudas, N. (2010). *TwitterMonitor: trend detection over the twitter stream*. In Proceedings of the 2010 international conference on Management of data, Indianapolis, Indiana, USA, pages 1155-1158.
- MPQA. MPQA Opinion Corpus Retrieved 9/1/2012, 2012, from <http://www.cs.pitt.edu/mpqa/>
- Osherenko, A. (2008). *Towards Semantic Affect Sensing in Sentences* In Proceedings of Artificial Intelligence and Simulation of Behavior (AISB).
- Pennacchiotti, M., & Popescu, A.-M. (2010). *Detecting controversies in Twitter: a first study*. In Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, Los Angeles, California, pages 31-32.
- Popescu, A.-M., & Pennacchiotti, M. (2010). *Detecting controversial events from twitter*. In Proceedings of the 19th ACM international conference on Information and knowledge management, Toronto, ON, Canada, pages 1873-1876.
- Reed, C., Elvers, T., & Srinivasan, P. (2011). *What's Trending? Mining Topical Trends in UGC Systems with YouTube as a Case Study*. In The Eleventh International Workshop on Multimedia Data Mining (MDMKDD 2011), San Diego, CA, USA.
- Rong Lu, Z. X., Yang Zhang and Quing Yang. (2012). *Trends Predicting of Topics on Twitter based on MACD*. *IPCSIT*, 25, pages 44-49.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2012). *Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development*. *IEEE Transactions on Knowledge and Data Engineering*, PP(99), pages 1-1. doi: 10.1109/tkde.2012.29
- Schier, J., Koppel, M., Argamon, S., & Pennebaker, J. (2006). *Effects of Age and Gender on Blogging*. *Proceedings of the AAAI Spring Symposia on Computational Approaches to Analyzing Weblogs*.
- Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). *The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic*. *PLoS ONE*, 6(5):e19467. doi: 10.1371/journal.pone.0019467

- Takahashi, T., Tomioka, R., & Yamanishi, K. (2011). *Discovering Emerging Topics in Social Streams via Link Anomaly Detection*. In 2011 IEEE 11th International Conference on Data Mining (ICDM), pages 1230-1235.
- Thelwall, M., Buckley, K., & Paltoglou, G. (2012). *Sentiment strength detection for the social web*. *Journal of the American Society for Information Science and Technology*, 63(1), pages 163-173. doi: 10.1002/asi.21662
- Tokuhisa, R., Inui, K., & Matsumoto, Y. (2008). *Emotion classification using massive examples extracted from the web*. In Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1, Manchester, United Kingdom, pages 881-888.
- Tweets2011. (2011). Tweets2011 Retrieved 12/23/2011, 2011, from <http://trec.nist.gov/data/tweets/>
- Vellerand, R. J. (1997). *Toward a hierarchical model of intrinsic and extrinsic motivation*. *Advanced Experiment, Social Psychology*, 29, pages 271-360.
- Vuong, B., Lim, E., Sun, A., Le, M., & Lauw, H. (2008). *On ranking controversies in wikipedia: models and evaluation*. In Proceedings of the international conference on Web search and web data mining, Palo Alto, California, USA, pages 171-182.
- Wikipedia. (2012, 9/7/2012). Wikipedia: List of controversial issues Retrieved 9/10/2012, 2012, from http://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues
- Zhuang, L., Jing, F., & Zhu, X.-Y. (2006). *Movie review mining and summarization*. In Proceedings of the 15th ACM international conference on Information and knowledge management, Arlington, Virginia, USA, pages 43-50.

Appendix A: Wikipedia List of Controversial Issues

2003 invasion Iraq, Abortion, Affirmative action, African American, Alberta secessionism, American Civil Liberties Union , ACLU, American Hunters and Shooters Association, History Jews United States, historical review American Jews, Anarchism, Anarcho capitalism, Anti Americanism, Anti clericalism, Anti Irish racism, Antisemitism, Asian American, Atheism, Austrian School, Osama bin Laden, Black supremacy, Boricua Popular Army , Los Macheteros, British National Party, Capitalism, Capital punishment, Catalonia, Cherokee Americans, China economy, China military, China human rights, Chinese spies, CIA leak scandal, Plame affair, Communism, Communist government, Communist state, Conservatism USA, Controversy, Copyright, Free software, viral license, Copyright infringement software, DMCA, copyright term extension, Child Exploitation, Crime USA, Criticism Walmart, Cuba, Culture war , Cyprus, Deaf American, Demographics Europe, Demographics United States, Domestic violence, Detroit, Economy Japan, Impeach George Bush, European Union, Fascism, Federal Deposit Insurance Corporation , FDIC, Federal Marriage Amendment, Feminism, Fox News, Free trade, Freedom fighter, Gay rights, LGBT rights, Genocide denial, Gun politics, Hamas, Harry's Place, Hate crime, Health care reform United States, Health Services Union expenses affair, Hezbollah, Hitler homosexuality mental illness , Holocaust, Holodomor, Homosexual agenda, Human rights Cuba, Human Rights United States, Human rights Kurdish Turkey, Hyphenated American, Illegal drug trade , drug abuse, Illegal Immigration USA, India economy, India politics, Politics Iran, Ireland Irish anti-british sentiment bias, Isa Gambar, History Israel, Israel Palestine conflict, anti-Italianism, Japan history world power, Jewish Americans lobby, Joe Biden, John Kerry, John McCain, Kashmir, Korean War, Kosovo, Louisiana politics corruption , Chicago-way politics, New Jersey style politics, Lebanon ethnic religious conflict, Liberalism USA, Libertarianism, Macedonia Slavs, Moldovans, Masculism, Mexico economic , Mexico political system, Morality ethics, Multiculturalism, Muslim Brotherhood, NATO , National Democratic Party of Germany bombing Dresden, Native American name controversy, National Anarchism, Nazi Germany, Neo-conservative, Neo-liberalism, New World Order , conspiracy theory, Politics of North Korea, Northern Ireland, Nullum crimen, nulla poena sine praevia lege poenali, Pakkoruotsi, Palestine, Palestine Liberation Organization, Park51, Patriotism USA, PETA, Pioneer Fund, Anti Polish people sentiment, Political correctness, partisan politics, Poverty, Price-Anderson Nuclear Industries Indemnity Act, Public education, Puerto Ricans United States, Quebec Canada, Race relations, Racial quota, Race ethnicity issues, Racism , Recession, Salvador Allende, Same-sex marriage, Sarah Palin, Saskatchewan Party, Laura Schlessinger, School violence, Seamus incident, Republika Srpska, Republic Serbian Krajina, Sedition, Sexual harassment, Separatism, September 11 attack, 9/11, Ariel Sharon, Sinkiang, Silesia, Socialism, Spain autonomous movement, Banking Switzerland, Synarchism, Syrian Social Nationalist Party, Tea Party movement, Tea Party protests, Taliban,

Terrorism, Terrorist groups, Tibet, Tobin tax, Torture, Universal Health care, United Nations, Patriot Act, War on Terror, Valencianism , Valencian Catalan language, Vector Marketing, Western civilization, White Americans, Wiretapping, Paul Wolfowitz, Womyn, Xinjiang, Yugoslavia breakup country, Zaire Democratic Republic of the Congo., Zimbabwe, 2011 Sendai earthquake tsunami, 1953 Iran coup d'état, 1973 coup in Chile, 1992 Los Angeles riots, September 11 terrorist attacks, American Airlines Flight 77, American Revolution, Apartheid, Apollo moon landing hoax, Armenian Genocide, Aryan invasion theory, Atomic bombings Hiroshima Nagasaki, Battle Cuito Cuanavale, Berlin wall, Black Power, Bloody Sunday 1972, Bosnia Herzegovina war, Bromberg Bloody Sunday, Centre Party Germany, Chicano nationalism, Coanda 1910, Cuban missile crisis, Communism history regimes, Confederate States of America, Jim Crow laws , racial segregation, Constitutional law, Crusades holy land, Cyprus dispute, D-Day 1944, East Germany, Ethnicity, Falkland Islands, French Revolution, Genocide, Great Depression, Green Revolution, Heimatvertriebene, Holocaust, Holocaust denial, Ireland history, Irish Potato Famine., Israel, Irredentism, Japanese American internment WWII, Jesus Christ biographical studies, Jews in Ukraine Poland, Antisemitism, Kennedy assassination, Kosovo War, Kurdish genocide, Kuril Islands, Libyan Civil War, Louisiana Purchase 1803, Majestic 12, Maoism, Marijuana legalization, Mexican American War, Middle Ages, Native Americans, Nazism, Nordic race global domination , Aryan race, White race, Okinawa rule 1945 1972, Operation Wetback, mass deportation, Other Losses, Panama canal, US rule Canal Zone, Philippines US rule 1898 1946, Prohibition 1919, Prussia, Puerto Rico annexation, Quebec separatism, Radicalism, Roswell UFO, Russian Polish relations, Sanhedrin, Scientology, Silesia, Slavery, Soviet Union, USSR, Spanish-American war, Spanish civil war, Spanish Inquisition, Stem cell research, Texas Revolution 1836, battle Alamo, Rape Nanking, Tiananmen Square protests 1989, Trail Tears, TWA Flight 800, U.S. Civil War, U.S. War of 1812, presidential election 2000, Unidentified flying objects UFO, Ukrainian Polish relations, Ukrainian Russian relations, Ustaše, Vietnam war, Watergate scandal, White supremacy, Women's rights feminism, causes effects World War, 2012 phenomenon, Mesoamerican Long Count calendar, Adventist, Agnosticism, American Family Association, Anti-clericalism, Anti-Judaism, Apollo Quiboloy, Atheism, Biblical literalism, Catholicism, Christian Coalition, Christian science, Christianity, Christophobia, Conservative Judaism , Creationism, Cult, Deism, Dhimmi, Dorje Shugden controversy, Druze, Eastern Orthodoxy, Evangelical, Falun Gong, Feminism religion, Fundamentalism, Gay Marriage opposition, Gnosticism New Testament, God, Guru structure criticism, Hare Krishna, Hate group , Historicity of Jesus, Homosexuality religion, Iglesia ni Cristo, Imperium Warhammer, Islamophobia, Islam radical, Jainism, Jehovah's witnesses, Judas Iscariot, Kabbala, Kashrut and Kosher, Ku Klux Klan, Last Supper of Jesus, Liturgy, Lutheranism, Makkah Mecca, Menachem Mendel Schneerson, Mennonites, Mormonism, Neocatechumenal Way, New Age, New Kadampa Tradition, Ole Nydahl Diamond Way Buddhism, Opus Dei, Pacifism, Paganism, Paul Gnosticism, Pentecostalism, Politics

religion, Prem Rawat Divine Light Mission, Project Chanology, Protestant Reformation, Reincarnation research, Republican Party religious right, Religion sexuality, Role women religion, Separation Church State, Satanism, Sathya Sai Baba, Scientology, Sharia, Sin, Sufism, Syncretism, Taboo, Temple Mount, United Submitters International, Qur'an alone, Rashad Khalifa peacock words, Universal Life Church, Universalism, Westboro Baptist Church, Abortion, ADD ADHD, AIDS, AIDS denialism, Alcoholism, Allergy causes, Allopathic medicine, Alzheimer disease, Alternative medicine, Anencephaly, Assisted suicide, Aspartame, Aspartame controversy, Asperger syndrome Autism, Astrology science, Bates method, Bioethics, Biology sexual orientation, Birth defect, Black hole information paradox, Blood transfusions, Cancer cures, Cesarean section, Chakra, Chiropractic, Chromosome, Circumcision, Cloning, Cochlear Implant, Cold fusion, Conjoined twins, Creationism, Cryonics, Depleted uranium, Depression, Disability, Drugs, Dyslexia, Ebola virus, Electrical sensitivity, Eugenics, Euthanasia, Evolution, Family Planning, Female genital cutting, Flat Earth Society, Genetic Engineering, Genetically modified foods, Gender differences, Gypsies race, Heredity, Heritability IQ, HIV AIDS, HMO Health, Hodgkin Disease, Homeopathy, Human cloning, Human evolution debate, Human longevity, Huntington disease, H5N1 virus bird flu, Intelligent design, IQ Wealth Nations, Jewish race, Kegels, Lactose intolerance inherited trait, Life extension, Lupus, Master race, Medical torture, Mental illness, Mental retardation, Mind control, Mucoid plaque, Nutrition, Obesity, Organ donor transplants, Overpopulation, Paternity testing, Pharmaceutical industry, Pluto demoted planet, Poppers, Premature birth, Prenatal care, Psychiatry, Race science, Race intelligence, Recapitulation theory, Schizophrenia, Self-mutilation, Sexually transmitted diseases, Smoking and tobacco, Sperm donor Egg, Stem cell research, Subluxation, Surrogate mother child custody, Tay-Sachs disease, Test-tube babies, Therapeutic Touch, Thiomersal, Transgender, Transsexual, BIID persons, Twin paradox, Big Bang theory, Vaccination, Veganism Vegetarianism, female health issues pregnancy menopause breast ovarian cancer, 69ing, Sexuality Abraham Lincoln, Adultery , Age consent, Anal sex, Abasiophilia, Adult grooming child sexual abuse murder online, Asian fetish, Attraction disability, Autogynephilia, BDSM, Child sexuality, Conversion therapy , reparative therapy, Ejaculation, Fetishism, Fingering and fisting, Fornication, G-Spot, Genetic sexual attraction, Genital modification and mutilation, Gerontophilia, Hentai, Hebephilia, Homosexuality, Incest, Lolicon, Love-shyness, Masturbation, Necrophilia, North American Man Boy Love Association, Oral sex, Orgasm, Paraphilia, Pederasty, Pedophilia, Polyamory, Polygamy, Pornography, Pregnancy fetishism, Rape, Sadomasochism, Semen, Sex offender, Sexual abuse, Stalking, Swinging , Transsexual, Women sexuality, Zoophilia, Bullfighting, Bumfight, Entertainment Software Rating Board, Duke Nukem Forever, Gambling, Girls Gone Wild, Hip hop culture, Entertainment restriction, MPAA film rating system, Playboy Magazine, Pop punk, Smooth jazz, Video game controversy, Disc jockeys, best-selling albums worldwide, Psytrance, Deforestation Logging, Dust bowl

intensive farming, Eco-fascism, Eco-terrorism, Energy consumption, Environmental impact hydraulic fracturing, Environmental vegetarianism, Fossil fuels air pollution greenhouse gases, Natural Gas Environmental effects, Coal Mining, Oil spills , Shale oil, Global warming cooling theory, Hydraulic fracturing, Hydrogen Atomic Bomb testing effects, An Inconvenient Truth, Minamata Syndrome, Mining Deep sea, Nuclear power Windfarms, Oil drilling ANWR petroleum reserves Alaska, Overpopulation, Ozone depletion hole, Recycling, Three Gorges Dam, Toxic waste disposal Love Canal, Weather modification, Crime America, Illegal drugs, Judicial system, Jury duty, Law enforcement, Legal status Minors Juveniles, Public education, Mandatory school attendance law, Prisons, Selective Service Military draft, Tax IRS, Ancient Macedonian language, Balto-Slavic languages, Basque language, Bilingualism Canada, Celtic revival, Chinese language policies, Creole languages, Dene-Caucasian languages, Ebonics, French language Canada, Graeco-Armenian, Hate crimes discrimination language speakers, Aryan race theory, Japanese Korean language similarities self-isolates, Ladino, Language revival, Linguistics, Macedonian language, Norwegian language controversies , Occitan Language, Occitania vergonha , linguistic France, Pidgin English, Quebec French official language province English, Riksmål, Sapir-Whorf Hypothesis, Sami language Scandinavia, Spanish United States, Turanism Turkic language, Welsh Language, Yiddish language, Friedrich Nietzsche, Meaning life, Plato, Scientific Method, Socrates, Truth, Carl Jung, Sigmund Freud, Thomas Szasz, Bit Torrent, BitTorrent protocol, Blu-ray Disc, Criticism Microsoft, Criticism Windows Vista, Criticism Apple, Console wars, Digital rights management, Electric car, Electronic voting, Facebook, Factory farming, File sharing, GNU Linux naming controversy, HD DVD, Indent style, MySpace online privacy security issues, Napster, Net neutrality, Norton AntiVirus, Nuclear power, Object-oriented programming, P2P, Removal of Internet Explorer, Smart meter concerns privacy health effects, Solar Power, Standardization Office Open XML, The Pirate Bay, Unmanned combat air vehicle use drones warfare, Windows Vista, Anime shows mature audiences Death Note Ghost in the Shell teens, Beavis and Butthead, Black Entertainment Television, CNN MSNBC Fox News Channel BBC Al-Jazeera, Censorship laws, Chinese television CCTV, Digimon Pokémon., Digital television conversion, Diversity prime-time, Disco music, Drawn Together, Emos, European culture, Family Guy, Federal Communications Commission, Gay culture, Google facebook privacy issue, Gothic subculture, Harry Potter, Heavy metal black death, Hippies, Internet forum message board, Liberal media, Mad TV, Media bias, Netiquette, Pentagon Papers New York Times 1971, Rap hip-hop, Rave culture, Sasha Baron Cohen Ali G Borat Brüno, The Simpsons, Smooth jazz, South Park, Telemundo Univision Spanish language Television, Talk radio limits free speech, Time-Warner Viacom multi-corporate monopolies, Trinity Broadcasting Network televangelists., The Twilight Saga, Twitter social trend, Virgin Records, Youtube content issues, Yuppies, Nero Caligula Roman emperors, Mahmoud Ahmadinejad, President of Iran, Syed Ahmed, Muhammed Ali conversion Islam anti-Vietnam war comments, Kirstie

Alley, Idi Amin Ugandan dictator, Criss Angel, Yasir Arafat, Pia Kjærsgaard, Joe Arpaio Sheriff Maricopa County, Hans Asperger, Mustafa Kemal Atatürk, Gilad Atzmon, Michelle Bachelet, Brigitte Bardot, Glenn Beck, David Beckham, Joy Behar, Art Bell conspiracy theories, Bruno Bettelheim, Tony Blair, Robert Blake 2002 homicide, Hans Blix, Barry Bonds BALCO drug scandal, Bonnie and Clyde, Ruder Bošković ethnicity, Lorena Bobbitt 1994 mutilated, Leonid Brezhnev, Anita Bryant, Kobe Bryant, Pat Buchanan, George W. Bush George H.W. Bush U.S. President, Jeb Bush, Richard Butler Aryan Nations group, George Carlin, Jimmy Carter President peace negotiator, Fidel Castro, Raúl Castro, Cesar Chavez, Hugo Chávez, Dave Chappelle satirist comedian, Dick Cheney Vice President CEO Halliburton, Cher Sonny Bono Mary Mack politics, Jacques Chirac, Margaret Cho, Cho Virginia Tech University gunman, Chris Christie governor New Jersey, Ward Churchill Colorado State professor comments survivors, Bill Clinton Hillary Rodham board executives Wal-Mart, Kurt Cobain 1994 suicide murder, Christopher Columbus, Sean Combs, Nicolaus Copernicus nationality Polish German, Stephen Colbert, John Corzine governor New Jersey, Ann Coulter, Tom Cruise Scientology, Salvador Dalí, Jeffrey Dahmer, Angela Davis homicide case, Hassan Diab 1980 Paris synagogue bombing, Andrew Dice Clay vulgar performance, The Dixie Chicks anti-Bush comments, Walt Disney antisemitism alcoholism, Phil Donahue, David Duke, Andrea Dworkin, Marc Emery, Eminem Marshall Mathers, Pablo Escobar, Jerry Falwell, Melissa Farley, Gerald Ford Watergate pardon, Henry Ford, Michael Fox Christopher Reeve stem cell research, Al Franken, Kenny G dispute genre, William Gaillard, Janeane Garofalo, Charles de Gaulle, Mel Gibson anti-Semitic homophobic comment DUI arrest, Kathie Lee Gifford sweatshop scandal affairs, Newt Gingrich Speaker House, Girls Aloud, Rudy Giuliani Mayor New York City, Nancy Grace, Wayne Gretzky Janet Jones sports gambling scandal, Joseph Goebbels, Whoopi Goldberg political comments The View, Mikhail Gorbachev, Al Gore Vice President, Temple Grandin autism, Woody Guthrie song lyrics endorse leftism socialism, Axl Rose, Ferenc Gyurcsány, Ted Haggard, Tarja Halonen, Sean Hannity, Phil Hendrie radio comedian., Patty Hearst kidnapping conversion Symbionese Liberation Army, Hideki Tojo leader Japan, Paris Hilton, Anita Hill 1991 Clarence Thomas sexual harassment, Heinrich Himmler, Emperor Hirohito involvement WWII, Adolf Hitler Holocaust exterminate Jews Roma Slavs homosexuals communists, Erich Honecker Communist leader East Germany, Jimmy Hoffa 1975 kidnapping disappearance, Edgar Hoover, Katie Hopkins, John Howard, Ron Hubbard, Mike Huckabee, Rock Hudson, Saddam Hussein, Yusuf Islam Catt Stevens, Andrew Jackson policy American Indians Trail of Tears, George Jackson, Jesse Jackson, Michael Jackson 2005 case, Thomas Jefferson President drafter constitution, Sheila Jeffreys, Elton John, Magic Johnson HIV, Michael Jordan, Jenny Jones, John Kennedy Robert Ted, Jacqueline Kennedy Onassis Greek billionaire, Kim Kardashian, Andy Kaufman, Khomeini, Nikita Khrushchev Soviet leader, Kid Rock, Martin Luther King Jr., Rodney King, Philip Klass, Shosei Koda, Junichiro Koizumi Japanese leader, K.D. Lang homosexuality animal rights activism, Ricki Lake, Lyndon

LaRouche, Lil Wayne, Carlos Latuff controversial cartoonist, Norman Lear producer social satire, Jean-Marie Le Pen, Heath Ledger, Vladimir I. Lenin Soviet leader, John Lennon CIA investigation, Lewis Libby, Rush Limbaugh, Abraham Lincoln President Civil War suspended civil liberties, Jennifer Lopez, Trent Lott, Courtney Love, Martin Luther, Catherine MacKinnon, Madonna, John Major, Nelson Mandela South African civil rights activist jail sentence criminal offense, Charles Manson, Marilyn Manson, Ferdinand Marcos Imelda Philippine, Rachel Marsden, Karl Marx, Jackie Mason, Paul McCartney domestic abuse Heather, Jenny McCarthy activist autism research, Steve McNair murder suicide girlfriend, Timothy McVeigh, Dmitry Anatolyevich Medvedev, Carlos Mencia, Angela Merkel , Freddie Mercury homosexuality AIDS, Harvey Milk, Demi Moore, Michael Moore, Jim Morrison death 1971 Paris, Mother Teresa, Benito Mussolini, Napoleon, Nicole Polizzi Snooki Chilean, Olivia Newton-John environmental activism, Richard Nixon President., Notorious B.I.G., Rosie O'Donnell, Opie & Anthony, Bill O'Reilly., Barack Obama President., Keith Olberman, Ozzy Osbourne, Donny Osmond Marie child stardom depression suicide, Terrell Owens near-death suicide attempt, Sarah Palin Vice Presidential candidate., Nancy Pelosi speaker of the House, Juan Peron Eva Argentine leader, Scott Peterson murder pregnant wife Laci, Augusto Pinochet Chilean dictator, Valerie Plame, Pope Benedict XIV, Pope John Paul II, Pope Pius XII, Elvis Presley death 1977, Prince Charles Princess Diana Camilla, Prince Harry William heirs British crown, Pol Pot Cambodian leader, Vladimir Putin Russian leader, Muammar Gaddafi Libyan leader, JonBenet Ramsey homicide case, Ronald Reagan President, Janet Reno Attorney General, Condoleezza Rice Secretary State, Michael Richards Laugh Factory incident, Yvonne Ridley, Alberto Rivera, Geraldo Rivera, Geshe Michael Roach, Oral Roberts, Pat Robertson, John Rocker baseball pitcher comments racial minorities homosexuals, Kid Rock, Romiõrio, Mitt Romney Presidential candidate, Linda Ronstadt Political activism lesbianism., Franklin Delano Roosevelt Theodore Presidents, Ben Rothlisberger, Pete Rose, Karl Rove, Babe Ruth, Donald Rumsfeld Secretary of Defense, Bobby Sands, Nicolas Sarkozy, Michael Savage, Terri Schiavo, Laura Schlesinger, Gerhard Schroder, Arnold Schwarzenegger governor, Teofilo Vargas Sein, Selena theories murder, William Shakespeare authorship question, Tupac Shakur death, Al Sharpton African American activist, Charlie Sheen, Matthew Shepard hate crime murder, Sifl and Olly, O.J. Simpson double murder, Anna Nicole Smith, Sonia Sotomayor, Snoop Doggy Dogg Calvin Broadus, Britney Spears Kevin Federline, Spice Girls controversies, Jerry Springer, Jon Stossel, Barbra Streisand, Joseph Stalin dictator, Gloria Steinem, Erika Steinbach, Howard Stern radio, Jon Stewart, Jimmy Swaggart, Nikola Tesla nationality, Margaret Thatcher Prime Minister, Pierre Trudeau, Harry Truman President atomic bomb Hiroshima, Donald Trump, Ted Turner, U2 activist, Joran Van der Sloot, Michael Vick dogfighting, Varg Vikernes, Andy Warhol, John Walsh , Kanye West, Ruth Westheimer, Joseph C. Wilson, Woodrow Wilson, Oprah Winfrey, Tiger Woods, Malcolm X, Andrea Yates Mass murderer, Boris Yeltsin Russian leader, Jay-Z, Zinedine Zidane

--

Appendix B: IRB Forms

Application for Exemption from Institutional Oversight

Unless qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/ projects using living humans as subjects, or samples, or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.

– Applicant, Please fill out the application in its entirety and include the completed application as well as parts A-F, listed below, when submitting to the IRB. Once the application is completed, please submit two copies of the completed application to the IRB Office or to a member of the Human Subjects Screening Committee. Members of this committee can be found at <http://research.lsu.edu/CompliancePoliciesProcedures/InstitutionalReviewBoard%28IRB%29/item24737.html>

– A Complete Application Includes All of the Following:

(A) Two copies of this completed form and two copies of parts B thru F.

(B) A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1&2)

(C) Copies of all instruments to be used.

*If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.

(D) The consent form that you will use in the study (see part 3 for more information.)

(E) Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB. Training link: (<http://phrp.nihtraining.com/users/login.php>)

(F) IRB Security of Data Agreement: (<http://research.lsu.edu/files/item26774.pdf>)



Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/irb

1) Principal Investigator: Rajshekhar Chimmalgil

Rank: Graduate Student

Dept: Industrial Engineering

Ph: 225-202-9800

E-mail: rchimm3@lsu.edu

2) Co Investigator(s): please include department, rank, phone and e-mail for each
*If student, please identify and name supervising professor in this space

Dr. Gerald M. Knapp, Industrial Engineering, Designated Associate Professor, 225-578-5374, gknapp@lsu.edu

IRB# E8139 LSU Proposal # _____

☒ Complete Application

☒ Human Subjects Training

3) Project Title: Controversy Trend Detection in Social Media

Study Exempted By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Exemption Expires: 2/13/2016

4) Proposal? (yes or no) ☐ no

If Yes, LSU Proposal Number _____

Also, if YES, either

☐ This application completely matches the scope of work in the grant

OR

☐ More IRB Applications will be filed later

5) Subject pool (e.g. Psychology students) Graduate students

*Circle any "vulnerable populations" to be used: (children <18; the mentally impaired, pregnant women, the aged, other). Projects with incarcerated persons cannot be exempted.

6) PI Signature _____

Date _____

1/25/13

(no per signatures)

** I certify my responses are accurate and complete. If the project scope or design is later changes, I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted. I also understand that it is my responsibility to maintain copies of all consent forms at LSU for three years after completion of the study. If I leave LSU before that time the consent forms should be preserved in the Departmental Office.

Screening Committee Action: Exempted ☒ Not Exempted _____ Category/Paragraph 2

Signed Consent Waived?: Yes ☒ No _____

Reviewer Mathews

Signature _____

Date 2/14/13

Study Exempted By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Exemption Expires: 2/13/2016

Consent Form

Controversy trend detection in social media research requires large amounts of text. The goal of this task is to compile a controversy corpus. We are asking that you read the displayed news article and identify whether you think the article is controversial or not. You will be using a web application to perform this task, which can be accessed from any computer that is connected to the internet. You may annotate the articles at your own pace.

Your participation is voluntary and you may stop at any point. We will make some or all of the data available to the wider research community. No one other than the project staff will have access to any forms you provide us. Your identity will remain confidential. Some general demographics are typically included in the scientific documentation of corpora and in published finding (e.g. age, gender).

To indicate that you wish to participate as outlined above, please complete the following:

I, (please print name).....
have read this form and agree to take part in the research on these terms.

Signature: Date:

Project description

The objective of this project is to annotate a corpus of news stories collected from cnn.com. The human annotators will be instructed to identify which news articles they think are controversial. For each article, the annotators will classify whether the article is controversial or not. A web application is used to perform this task.

For each session, the annotator clicks on "Begin Annotation" to start the annotation process. After clicking "Begin Annotation", the annotator is displayed a news article with the option to select controversial or not controversial. After the annotator select controversial or non-controversial, they will be displayed the next news article to be annotated. The web application is stored on a web server located in Peobody Hall. The annotators can access the website from any computer which has an internet connection. There is no constrained time schedule for the annotators to complete the task. Annotation per article take 10-15 minutes.

Appendix C: Annotator's Algorithm

"I read the article and if I thought it was controversial then I read the comments to see if others thought so too. If there was more back and forth and fighting in the comments then I figured controversial. The sports ones that were strictly just about sports and not doping or cheating, I classified it as non-controversial"

"The main thing I was looking for was whether the article was about anything political - because there seemed to be a lot of articles like that, and politics is something that people never seem to agree on. For everything else, I was looking to see if there was a lot of debate going on in the comments and/or the topic was one that could potentially lend to debate."

"I was looking for strength of opposing comments. Also certain hot topics that already receive a lot of attention in the media, like gay rights or abortion."

"After analyzing the title of an article and reading the introductory paragraphs, I looked mostly for strong keywords to make a decision on classification of that article. Words pertaining to controversial topics like gun control, immigration, Obama, Romney, fiscal policy, national debt, abortion, among others, helped me quickly identify the article without actually reading that article in its entirety. Also, I went through most of the comments which gave a better idea about the article."

"Whether it mentioned a controversial topic or not and whether two sides seemed to be pitted against each other over a certain topic. Looked to see whether people talked about two sides or opposing arguments in comments."

"Whether the topic had a potential to spark some controversies, I was also looking to see whether the news was just 'reporting' or more opinionated."

"Looked at comments, if against article topic or against other commentators then classified it as controversial."

"I was looking for conflict among the commentators, disparate opinions maybe, disagreement on the goodness or badness of the article. I was looking to the commentators for their reactions instead of assessing the article myself."

"If I thought the topic of the article could be controversial to people and if there were a lot of comments then I classified the article as controversial."

Vita

Rajshekhar V. Chimmalgi was born in Gokak, India in 1985. Following his graduation from McKinley Senior High School, Baton Rouge, Louisiana, he obtained his bachelor's degree in computer science from Southern University, Baton Rouge, Louisiana, in 2010. He began working towards the degree of master's of science in engineering science with concentration in information technology and engineering at Louisiana State University in 2011. During his time as a graduate student, he worked as a graduate assistant for Louisiana State Youth Opportunities Unlimited (LSYOU). His degree will be conferred at the Spring Commencement May 2013.