An Analysis of Slant in Tweets: Case Study

Nagaraju Vadranam
Department of Computer Science
Oklahoma State University
Stillwater, OK 74078, USA
nvadran@okstate.edu

K. M. George
Department of Computer Science
Oklahoma State University
Stillwater, OK 74078, USA
kmg@cs.okstate.edu

Starla Marier Demings†
Department of Mathematics
East Central University
Ada, OK 74820, USA
stamdem@email.ecok.edu

ABSTRACT

Determination of quality and reliability of information found in social media have been subjects of study by sever researchers. One set of solution may not work in all cases. This paper presents a method to estimate the slant of tweets related to a topic. The general approach followed is to construct labeled data from tweets and use supervised learning to build predictive models. Results obtained from two datasets are compared against OTC model and a CNN based model.

CCS CONCEPTS

• General and reference~Empirical studies • General and reference~Experimentation • General and reference~Evaluation

KEYWORDS

bias, veracity, tweets, sentiment analysis, clustering, neural network, prediction.

ACM Reference format:

Nagaraju Vadranam, K.M. George and Starla Marier Demings. 2019. An Analysis of Slant in Tweets: Case Study. In *Proceedings of IEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT '19), December 2-5, 2019, Auckland, New Zealand.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/1234567.

1. INTRODUCTION

As of the first quarter of 2019, Twitter alone averaged 330 million monthly active users [12]. So, an accurate assessment (veracity) of information found or propagated through social media is a major concern. News stories about misuse of Facebook for political purposes during the 2016 US Presidential election [13] underscores the problem. Veracity of tweets is subject to different interpretations such as objectivity, truthfulness and credibility [8]. In this paper, we propose a method to determine veracity of tweets. For this purpose, we view veracity of a tweet as its fidelity to a given topic. We characterize this fidelity by bias or slant/leaning within the context of given topic. (In the reminder of the paper, unless stated, veracity and slant will have same meaning).Then, we may approach the problem as a 3-way classification problem relative to a topic. For this purpose, we treat tweets as points of a space with a set of features associated to the topic forming a basis. This paper outlines one approach to analyze and classify slant of tweets related to a topic (defined by a set of words) as positive, negative, or neutral.

We implemented and tested two classification schemes (crisp and fuzzy). The two implementations are: 1) rate the tweet slant as positive (0), negative (2), or neutral (1); and 2) assign to tweets a score in the interval [0, 1]. In the first approach, we combine clustering and supervised learning algorithms. In the second approach, we use fuzzy c-means clustering [1] and nonnegative matrix factorization (NMF) [4][7]. We have also tried fuzzy clustering to build labeled data for supervised learning. Extensive tests show that a combination of K-means clustering and ANN based on five features of tweets (described later) are effective in identifying a tweet's slant as positive, negative, or neutral as related to the topic of analysis.

2. RELATED WORK

This section provides a minimal list of available related work. The research in [11] provides a survey of papers that have studied rumors and provides a rumor classification system architecture with four components, namely rumor detection, rumor tracking, stance classification, and veracity classification. According to Goel and Uzuner [5], detecting fraud is a complex problem and no one set of predictors will be always successful in fraud detection. Kwon et. al. [6] claim as one of the first papers on analysis of rumor propagation in social media. Their approach identifies rumors based on "temporal, structural, and linguistic properties of rumor propagation". They built classifiers based on decision tree, random forest, and SVM to classify a topic as rumor or non-rumor by analyzing related tweets. Giasemidis et.al. [3] approach rumor identification as a supervised binary classification problem. In [9], Nguyen et.al. present a method for early stage rumor detection. They employ Convolutional Neural Networks (CNN) to learn hidden representations of rumor related tweets. Several papers study veracity of tweets. Lukoianova and Rubin [8] study veracity in big data across three main dimensions: objectivity/subjectivity, truthfulness/deception, and credibility/implausibility. The three dimensions form a veracity index. Chen et.al. [2] use CNN for rumor verification.

† supported by REU Site award "Big Data Analytics at Oklahoma State University", CNS- 1659645.

BDCAT '19, December 2–5, 2019, Auckland, New Zealand

ACM ISBN 978-1-4503-7016-5/19/12…$15.00
DOI: https://doi.org/10.1145/3365109.3368770

3. SLANT COMPUTATION FRAMEWORK

In this section, we describe our approach at a design level. We do employ manual validation to label clusters. Other external data are not used in the process. We view veracity of a tweet as bias/slant in relation to a topic of interest rather than truthfulness, accuracy, or correctness in relation to a ground truth. We assume that a set of k features (depended on the topic) can be associated to each tweet and these feature values map tweets into a k-dimensional space. The features collectively are used as indicator of the direction of the slant.

We programmed two approaches for slant detection. The first approach consists of two parts. The first part constructs a slant estimation model and the second part predicts the slant direction from the features input. Common to both parts is feature vector construction module. To construct the slant estimation model, first we build a labeled set of feature vectors by classifying them into three groups by a classification algorithm (K-means). The labeled groups are 'positive', 'negative' and "neutral". This labeling is a manual process. An ANN module trains a prediction model from the labeled data.

The second approach that we followed is fuzzy clustering. The motivation is to avoid errors that could occur in assigning definitive statements of positive, neutral, or negative that could be incorrect labels. In this approach, we assign a weight measure between 0 and 1 to tweets. We have tested this approach with c-means clustering and nonnegative matrix factorization as clustering methods. The fuzzy membership value is used as ranking measure.

The idea behind the fuzzy clustering approach is to classify the feature vectors (training set) using fuzzy clustering algorithms into three fuzzy clusters. Observe that every vector (by implication the associated tweet) will be in the three fuzzy clusters with different membership values. The sum of the membership values is 1. Choose the fuzzy cluster C that is closely associated to positive slant tweets (currently this is a manual processes). Choose the fuzzy membership value in C of tweets as measures of slant. This membership value is a number in the interval [0, 1]. The process so far is similar to model building in the first approach. Unlike the first approach, the second one provides a ranking of tweets. This ranked set of tweets is considered as the model. To compute the slant of a new tweet, construct its feature vector and determine the closest one above and below in the model using the chosen similarity measure used for clustering. The slant measure of the new tweet is chosen as the average of the measures of the two closest tweets in the ranking above and below. To implement this approach, we have used fuzzy c-means algorithm [1] and nonnegative matrix factorization (NMF) algorithm [4][7]. To implement fuzzy clustering with NMF we used Algorithm 1.

Algorithm 1: Fuzzy clustering
INPUT: An n-by-m matrix X (columns are features)
OUTPUT: Fuzzy membership functions μ_1, μ_2, μ_3

Step 1: Compute the NMF of X, $X \approx WH$, where W is n-by-3 matrix and H is a 3-by-m matrix. Each row of W represents a tweet. [MatLab can be used for NMF or one of the available algorithms can be implemented]

Step 2: Let each column of W be a cluster.

Step 3: Normalize the rows of W so that the sum of the elements is 1.

Step 4. The columns of the normalized W be the three clusters, and the row values are the membership functions μ_1, μ_2, μ_3.

4. EXPERIMENTS

In this section, we report experiments using two tweet topic areas: 1) the North Korean peace talk and 2) National Rifle Association (NRA). The first topic was current during the data collection period. The REU student chose the second topic out of personal interest. We consider only original tweets as retweets do not change the slant of a tweet. Peace talk data were collected from May 31 to July 27, 2018. Of the 8990152 tweets collected, 525478 were identified as original tweets and considered for our experiments. NRA dataset was collected during the period June 7 to August 3, 2018. Of the 1942518 tweets collected, 98920 were original tweets and used for our experiments. We associated five features with the tweets (Table 1). Feature determination was a manual process. The considerations in choosing the features are based on our own observation of tweets and information from the literature.

Table 1. Feature Description.

Feature	***Name***	***Rationale***
V1	Non-follower retweet count	Independence of retweets.
V2	Positive sentiment	Positive and negative sentiment of fraudulent tweets are higher than truthful tweets [5]
V3	Negative sentiment	
V4	Word weight	Sum of weights of words occurring in a tweet distinguish user traits [14]
V5	Entropy	Measure of uncertainty implicit in statements.[10].

To compute features, we considered only original tweets. AFINN database was used for sentiment score computation. The five feature computations are done as follows:

1) V1 = if (retweet count > follower count) then rtwwet count – follower count; else 0.
2) V2 = score = get_afinn_scores(Owner_tweet_text); int(score['positive']).
3) V3 = score = get_afinn_scores(Owner_tweet_text); int(score['negative']).
4) V4 = sum of weights of words in the tweet, (Gensim package was used)
5) V5 = entropy computed based on [10]

We have performed several experiments with normalized and un-normalized values of the features V1-V5. If data is not normalized, then the features whose computed values are high seem to dominate the clusters computed by k=means algorithm. Hence, the results presented here are based on normalized data except for nonnegative matrix factorization. Normalization introduces negative coordinates which will affect factorization and thus normalization was not used in the case of fuzzy clustering with NMF.

4. RELATIONSHIP BETWEEN FEATURES

The predictive capability of our method depends on the choice of features selected to capture the topic's sense from tweets. We examine the selected features for dependency between them to avoid having undue bias in representation due to causality. While no representation of topics by features may be 100% accurate, one can posit that if the features are independent, they may represent the topic more accurately and the resulting analysis more reliable. As stated previously, we chose five features that can indicate slant of tweets to the topic of analysis based on review of literature and review of tweets themselves. In order to assess the independence of the five features that we selected for the tweets, we preformed correlation analysis. We analyzed the two datasets mentioned in the previous section. Tables 3 and 4 show very low correlation between the features. The first feature (V1) is depended on followers and retweet counts. Computed the correlation between follower counts and retweet counts is a low value of 0.16 for the peace talk data set and -0.02 for the NRA dataset. This indicates the independence of the retweets of non-followers. It is intuitively obvious that a follower who retweets also supports the original tweet. Therefore, excluding the follower count from the retweet count removes implied bias in the computation of the feature V1.

Table 3. Feature correlation matrix (peace talk data).

Feature Correlations		Features				
		V!	V2	V3	V4	V5
Features	V1	1.00	0.01	0.00	0.01	0.00
	V2		1.00	-0.10	0.30	0.08
	V3			1.00	-0.35	-0.07
	V4				1.00	0.18
	V5					1.00

Table 4. Feature correlation matrix (NRA data).

Feature Correlations		Features				
		V1	V2	V3	V4	V5
Features	V1	1.00	0.13	-0.10	-0.10	0.14
	V2		1.00	-0.10	-0.01	0.33
	V3			1.00	-0.02	-0.37
	V4				1.00	0.19
	V5					1.00

5. COMPARISON WITH OTHER APPROACHES

For comparison purposes, we have implemented OTC (Objectivity, Truthfulness and Credibility) [8] and an adaptation of Rumor verification using Convolution Neural Networks [2]. We chose these models for comparison because they represent different approaches for veracity estimation. These models do not define veracity as slant (the meaning chosen in this paper). However, as there are no exact models for comparison, we chose these two models. We adopt Cohen's kappa (κ) as one measure for comparison and use it to compare OTC method and our proposed method. Cohen' kappa measures the agreement between two or more raters when measurement scale is categorical. It is defined as: $\kappa = 1 - \frac{1-p_o}{1-p_e}$, where p_o is the relative observed agreement among raters (identical to accuracy), and p_e is the hypothetical probability of chance agreement, p_o and p_e can be computed as follows:

Assume that two algorithms classify N items into m groups. Then, their agreement can be given by an mXm contingency table $A = (a_{ij})$. The formulae below are used to compute p_o and p_e.

$p_0 = \frac{1}{N}\sum_1^m a_{ii}$ *and* $p_e = \frac{1}{N^2}\sum_1^m A_i B_i$ where A_i is the sum of the i^{th} row of A and B_i is the sum of the i^{th} column of A. If $\kappa = 1$, there is perfect agreement. If $\kappa \leq 0$, perfect disagreement.

The OTC model computes a veracity index for each tweet. To apply Cohen's kappa as comparison measure, we divided the index range (0, 1) into three equal subintervals. The tweets with OTC index in a subinterval are grouped into one cluster to obtain consistent partitioning with other approaches. As K-means based classification gave best results in experiments of our method, it is used in comparison analysis. The model comparisons are given in the following section.

OTC Model

Tatiana Lukoianova and Victoria L. Rubin, in their paper [8], suggest defining veracity across three dimensions: 1) objectivity, 2) truthfulness, and 3) credibility. These dimensions are mutually exclusive. They suggest available tools such as NLP, LWIC, and mutual information to compute the three dimensions. The three main dimensions are normalized to the interval (0, 1) with 1 representing the maximum value in each dimension. They further compute the average of the three values to produce an OTC index. We computed OTC index using Text Blob (python Package) for Objectivity /Subjectivity measure and Empath (python Package) for Truthfulness/Deception measure. We implemented programs to compute mutual information (MI) using the formula for MI, $MI(w_1, w_2) = \log(\frac{p(w_1,w_{2)}}{p(w_1)*p(w_2})$.

To compare OTC with our model, we divided OTC index into high, middle and low. High index is associated to being positive slant, low index is associated to being negative slant and the middle index is associated to being neutral slant in our model. Tables 5, and 6 are comparison contingency tables for the two datasets we analyzed. The table entries are the number of tweets in the intersection of corresponding column and row. The kappa values are computed using the table entries. The values for the two cases are -0.015 and 0.008 which show very low agreement among the outcomes of the two models. This is consistent with our manual examination of results. In Tables 5 and 6, the middle row dominates the other two. This is due to the distribution of OTC index values. Most values fall in the range (0.33, 0.7). There is significant differences in the classification of both methods. Neither method is 100% trustworthy. Two examples are given below where one method is accurate and the other is not.

Tweet :" Democrats Are The Party Of Death Democrats Are The Party Of Death …". While our method classifies it as negative, the OTC index value is 0.75.

Tweet: "Impossible to trust Dems". While OTC index value is 0.33, our method rates it as neutral.

These examples explain the discrepancies and are representative of the results. They explain why the Cohen's kappa values are low.

CNN Model

Our proposed method and the CNN approach provide three clusters each representing positive (0), negative (2) and neutral (1) slant (veracity).

Table 5. NRA data: OTC and our model contingency table.

		Our method		
		Positive (0)	**Neutral (1)**	**Negative (2)**
O	**high index**	1	865	849
T	**Middle index**	9	6130	11420
C	**Low index**	0	28	115

Table 6. Peace talk data: OTC and our model contingency table.

		Our method		
		Positive (0)	**Neutral (1)**	**Negative (2)**
O	**high index**	1193	1066	125
T	**Middle index**	14232	11534	1128
C	**Low index**	130	144	8

We compared the distributions of positive, negative, and neutral slants produced by the two methods. Tables 7 and 8 show results produced by the two datasets we analyzed. CNN method produced less negative slants and more positive slants than our method in both datasets. The proportions of the positive, neutral and negative percentages are similar for both approaches.

Table 7. NRA data: CNN and our model comparison.

	Positive (0)	**Neutral (1)**	**Negative (2)**	**TOTAL count**
Our Method	1.76%	27.85%	70.40%	16863
CNN	8.25	35.09%	56.66%	16487

Table 8. Peace talk data: CNN and our model comparison.

	Positive (0)	**Neutral (1)**	**Negative (2)**	**TOTAL count**
Our Method	3.14%	53.41%	43.45%	29166
CNN	8.63%	60.93%	30.44%	26396

6. SUMMARY AND CONCLUSIONS

In this paper, we report methods that we designed and implemented in Python language to estimate the veracity of tweets. As determination of ground truth related to any statement is difficult and time consuming, we considered veracity as fidelity to a topic and interpreted veracity of a tweet as the slant in relation to a topic. Our approach combines several ideas including sentiment analysis, clustering, and machine learning. Several features are associated to the topic. We assume that these features form the basis of a space in which tweets are points. From that perspective, determination of features is a critical step. Based on experiments, we conclude that the choice of features to represent the topic will influence results' dependability.

It is obvious that different veracity computation schemes will provide different outcomes. Therefore, metrics for comparison will be useful. We computed Cohen's kappa to compare our method and the OTC model.

We performed extensive runs of implementations of the method with different clustering algorithms. Based on experiments, K-means clustering had the best performance. However, our observation is that fuzzy clustering could be a useful tool to rank tweets in a veracity scale.

REFERENCES

[1] James C. Bezdek, Robert Ehrlich, and William Full, "FCM: The Fuzzy C-Means Clustering Algorithm", Computers & Geosciences, Vol. 10, No. 2-3, 1984, pp. 191-203.

[2] Yi-Chin Chen, Zhao-Yang Liu, Hung-Yu Kao, "IKM at SemEval-2017 Task 8: Convolutional Neural Networks for Stance Detection and Rumor Verification", Proceedings of the 11th International Workshop on Semantic Evaluations (SemEval-2017), Vancouver, Canada, August 3 - 4, 2017, pp. 465–469.

[3] G. Giasemidis, et al. (2016) Determining the Veracity of Rumours on Twitter. In: Spiro E., Ahn YY. (eds) Social Informatics. SocInfo 2016. Lecture Notes in Computer Science, vol 10046. Springer, Cham, pp. 185-205.

[4] Nicolas Gillis, "The Why and How of Nonnegative Matrix Factorization", arXiv:1401.5226v2, Mar 2014.

[5] Sunita Goel and Ozlem Uzuner, "Do Sentiments Matter in Fraud Detection? Estimating Semantic Orientation of Annual Reports", Intelligent Systems in Accounting, Finance and Management, Volume 23, Issue 3, Special Issue in Accounting, Auditing and Finance Applications, July/September 2016, pp. 155-254

[6] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang "Prominent features of rumor propagation in online social media", Proceedings of the 2013 IEEE 13th International Conference on Data Mining, IEEE, 2013, pp. 1103–11083.

[7] Daniel D. Lee and H. Sebastian Seung, "Algorithms for non-negative matrix factorization", Proceedings of the 13th International Conference on Neural Information Processing Systems, Denver, CO, 2000, pp. 535-541.

[8] Tatiana Lukoianova and Victoria L. Rubin, "Veracity Roadmap: Is Big Data Objective, Truthful and Credible?", Advances In Classification Research Online, 24(1). doi:10.7152/acro.v24i1.14671

[9] Tu Ngoc Nguyen, Cheng Li, and Claudia Niederée, "On Early-Stage Debunking Rumors on Twitter: Leveraging the Wisdom of Weak Learners", Proceedings of the Social Informatics: 9th International Conference, SocInfo 2017, Part 2, pp. 141-158.

[10] C. E. Shannon, "A Mathematical Theory of Communication", The Bell System Technical Journal, Volume 27, Issue 3, July 1948, pp. 379-423.

[11] Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter, "Detection and Resolution of Rumours in Social Media: A Survey", ACM Computing Surveys, Vol. 51, No.2, February 2018, pp. 32:1–32:36, https://doi.org/10.1145/3161603.

[12] https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

[13] https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html

[14] Nadeem Ahmad and Jawaid Siddique, "Personality Assessment using Twitter Tweets", Proceedings of the 21st International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2017,6-8 September 2017.