Effective method for detecting drunk texting

Marcos A. Grzeça[1], **Karin Becker**[1], **Renata. Galante**[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{magrzeca,karin.becker,galante}@inf.ufrgs.br

Nível: Mestrado
Data de ingresso no mestrado: Março de 2017
Data prevista para conclusão do mestrado: Março de 2019
Etapas concluídas: revisão da literatura; identificação do baseline; definição e implementação do modelo de enriquecimento semântico.
Etapas futuras: integração entre word embeddings e enriquecimento semântico; experimentação; avaliação.

Abstract. *Data from social networks is used by researchers to analyze social problems due to the ease in obtaining up-to-date data. A current social problem is excessive alcohol consumption. Social networks are key factors to understand the reasons that lead to excessive consumption, since the use of social networks and alcohol consumption are behaviors linked to young people. Classifying drunk texting from social networking data is complex because texts are short, sparse and written with diverse vocabulary. The aim of this work is to provide an efficient method to identify drunk texting through a framework that combines semantic enrichment and word embeddings. Developing a framework combining these techniques is important to analysis of short texts, because semantic enrichment provides features that add context to the texts, but adding features to word embeddings without interfering with context is still an open problem.*

Resumo. *Dados de redes sociais são utilizados por pesquisadores para analisar problemas sociais devido a facilidade em obter dados atualizados. Um problema social atual é o consumo excessivo de álcool. Redes sociais são fatores chaves para entender os motivos que levam ao consumo excessivo, visto que a utilização de redes sociais e o consumo de álcool são comportamentos vinculados a jovens. Classificar drunk texting a partir de dados de redes sociais é complexo, pois os textos são curtos, esparsos e escritos com um vocabulário diversificado. O objetivo deste trabalho é fornecer um método eficaz capaz de identificar drunk texting através de um framework que combine enriquecimento semântico e word embeddings. Desenvolver um framework combinando essas técnicas é importante para analisar textos curtos, pois o enriquecimento semântico fornece features que adicionam contexto aos textos, mas adicionar as features ao word embeddings sem interferir no contexto ainda é um problema em aberto.*

1. Introduction

Researchers have investigated the use of social networks and big data to discover information regarding topics related to health, social problems, and events [Kershaw et al. 2014]. Social networks are often used as basis of these studies due to the data being generated in large volumes, up-to-date, easy to obtain [Culotta 2013].

Excessive alcohol consumption is a major problem that affects our society. Alcohol caused around 3.3 million deaths worldwide in 2012 and is responsible for 15% of deaths resulting from traffic accidents [Organization and of Substance Abuse Unit 2014]. Young people are more likely to end up in fatal traffic accidents linked to alcohol consumption [de Informações sobre Saúde e Álcool 2014]. It is important to understand and monitor the factors that cause high alcohol consumption to assist in public health policies.

Social networks are key elements to this purpose, since alcohol consumption and the use of social networks are behaviors related to youngsters. Young people who use social networks are more likely to use tobacco, alcohol, and marijuana [Johnson et al. 2011] — and repeated exposure to posts involving drugs on social networks may encourage others, as this behavior may then be perceived as normal [West et al. 2012]. Texting while under the influence of alcohol is popularly regarded as drunk texting[1]. In Twitter, tweets written under the influence of alcohol are 'drunk tweets', and the opposite are 'sober tweets' [Joshi et al. 2015].

Official data on alcohol consumption and data extracted from drunk tweets are strongly correlated [Culotta 2013, Kershaw et al. 2014], confirming the accuracy of inferences extracted from social networks. Moreover, about 500 million tweets are posted daily, and 37% of registered users are between 18 and 29 years old. Thus, Twitter can be a valuable source of information to this end [West et al. 2012] because it provides messages, mostly posted by young people, which may be related to alcohol consumption and can be obtained in real time.

The automatic identification of drunk texting allows detailed studies on alcohol consumption based on a large volume of data and can warn family or friends about some danger (e.g. driver is driving while drunk). Such studies can provide public officials with information related to the identification of factors that cause excessive consumption of alcohol, helping in its prevention and control. Techniques suitable for drunk texting classification can be adapted to identify other types of drug abuse.

The classification of drunk texting is a complex task, because texts are written in natural language and different vocabulary with a great morphological variation. Another issue is the scarcity of categorized databases that can be manipulated for this goal. Related work has developed drunk tweets classifiers based on features extracted using traditional natural language processing (NLP) techniques (e.g. n-grams, stemming), sentiment analysis, and the morphology of the sentence to identify alcohol consumption [Jauch et al. 2013, Aphinyanaphongs et al. 2014, Joshi et al. 2015, Hossain et al. 2016]. None of these works have attempted to improve the categorization text by providing contextual meaning through semantic enrichment.

The objective of this work is to develop a method capable of identifying drunk texting by combining two very active research areas: semantic enrichment ([Romero and Becker 2017]) and word embeddings ([Li et al. 2016]). The research ques-

[1]https://www.urbandictionary.com/define.php?term=drunk+texting

tions that guide our work are: *i*) How to semantically enrich tweets and to identify the most important entities; *ii*) What is the best way to use word embeddings in this context? and *iii*) How can word embeddings and semantic enrichment complement each other?

The combined use of word embeddings and semantic enrichment is important for analyzing short and noisy pieces of text, especially when there is not a large number of records available. Under these conditions, the exclusive use of word embeddings may result in the learning misleading or irrelevant patterns found in the training data [Chollet and Allaire 2018]. On the other hand, the single use of semantic enrichment is not enough to learn all relationships between the tokens present in the texts.

The remainder of this paper is structured as follows: Section 2 presents related works linked to the area. Section 3 introduces the proposed method. Section 4 presents the preliminary results obtained. Section 5 addresses future works.

2. Related work

The detection of drunk texting is inserted in the area of paralinguistics. This area began to receive increased attention from researchers in 2011 with the challenge 'The INTERSPEECH 2011 Speaker State Challenge' ([Schuller et al. 2011]) where the goal was to classify the intoxication level through speech.

Since then, drunk texting classification has been addressed by works such as [Jauch et al. 2013], [Aphinyanaphongs et al. 2014], [Joshi et al. 2015] and [Hossain et al. 2016]. These works use traditional feature extraction techniques for text classification, such as classical pre-processing (e.g. tokenization, removal of stopwords, normalization of mentions/URLs), data cleansing, extraction of n-grams, stylistic features (e.g. number of discourse connectors, number of words, number of capital letters in the tweet), and sentiment analysis. They also experimented with different classification algorithms, such as Random Forest, Support Vector Machine (SVM), Generalized Linear Model (GLM), and Naive Bayes.

The existing papers in this area did not use semantic enrichment nor word embeddings strategies. The developed method allows extracting components that indicate the consumption of alcohol, enabling more detailed studies that help to understand reasons for alcohol abuse.

2.1. Semantic enrichment

To achieve good results in text classification problems, it is necessary that the text be long enough to provide features for the classifier to learn patterns in the data [Li et al. 2016]. To deal with short and sparse texts, such as tweets, related work uses semantic enrichment, because it improves the identification of representative terms or entities related to a tweet [Romero and Becker 2017].

Semantic enrichment uses external knowledge bases (DBPedia, Wikipedia, YAGO, etc...) to add context to short texts and extracts named entities (NER), key words, concepts, and categories present in the text. Entities can generalize texts written with different terms, but with the same semantic meaning.

2.2. Word embeddings

The classification of short texts also suffers from the problem of 'word mismatch', due to its limited context [Li et al. 2016]. To overcome word mismatch, word embeddings can be

used. Word embeddings are dense vector representation of words whose value can be used to determine the similarity between words. However, as pointed out by [Li et al. 2016], 'to get high quality embedding vectors, a large amount of training data is necessary'.

While word embeddings are useful to determine the similarity of words, semantic enrichment contributes by adding context to the text. There is an opportunity to integrate them to analyze short texts because the addition of semantic enrichment allows word embeddings to learn semantic similarity of texts, helping in the classification of short texts written with different words, but with similar meanings. The main challenge in using semantic enrichment and word embeddings is how to add semantic information without interfering in the context of texts, damaging the vector representation of word embeddings.

3. Proposed Framework

The goal is to develop a method that can identify drunk texting in short and noisy pieces of text, such as tweets, by integrating semantic enrichment and word embeddings. Semantic enrichment can provide context to texts, while word embeddings learning the similarity between the words in text, optimizing the classification of short texts.

Initially, semantic enrichment (Section 3.1) and word embeddings (Section 3.2) have been used alone for text classification. We are working to integrate semantic enrichment and word embeddings.

3.1. Semantic enrichment

In this section, we describe the proposed framework for improving the classification of drunk texting in tweets using semantic enrichment, illustrated in Figure 1. This framework is based on the strategies adopted by [Romero and Becker 2017] for event classification.

The framework is composed of seven steps: *i*) pre-processing, which includes emoticon polarity and posting time identification; *ii*) error handling to deal with typing errors that are a possible side effect of drunkenness; *iii*) extraction of conceptual features using Natural Language Understanding (NLU) to provide meaning to terms; *iv*) use of linked data (from DBPedia) to extract semantic aspects that can be used to generalize the conceptual features; *v*) pruning to select only the discriminant features, since the NLU and linked data stages result in a large number of features, which may degrade the classification performance; *vi*) textual features (1-grams, bi-grams, and hashtags), flags indicating

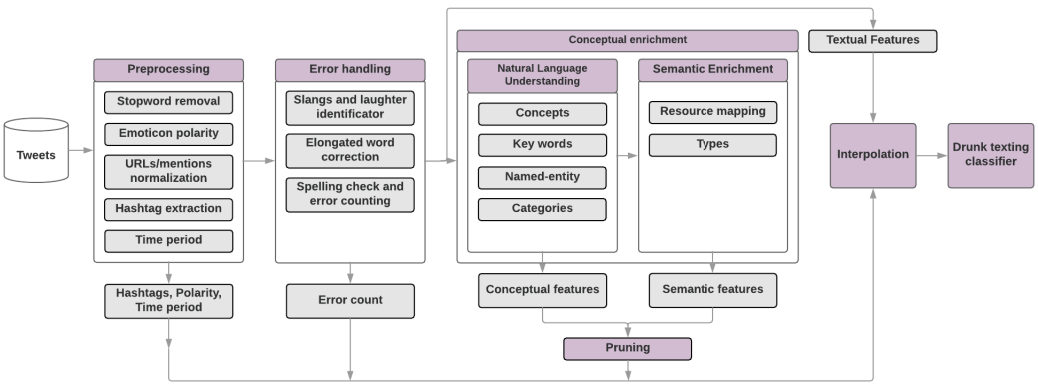

Figure 1. Framework Overview

the presence of errors, posting time and conceptual/semantic features are combined; *vii*) tweets are classified into either: 'drunk texting' or 'sober texting'.

3.2. Word embeddings

We propose to learn the word embeddings from a tweet text using the *word2vec* algorithm. In the current stage of the research, we use these embeddings as one of the layers of a recurrent neural network Long Short-Term Memory (LSTM) to classify drunk tweets. As the next step, we will investigate how to integrate semantic enrichment and word embeddings, and experiment with other LSTM architectures.

4. Experiments and preliminary results

The goal of our experiments is to verify the effectiveness of the proposed framework towards identifying drunk texting. We assess our contribution by comparing metrics of information retrieval (IR) with the results obtained by [Hossain et al. 2016].

4.1. Dataset

The dataset is based on [Hossain et al. 2016]. This article was chosen as baseline because it made available its annotated dataset for tests. We collected 3996 tweets. For each tweet in the dataset, [Hossain et al. 2016] they hired Amazon Mechanical Turk[2] and asked three questions: *Q1*) Tweet mentions the activity (drinking alcohol); *Q2*) Tweet is about the tweeter doing the activity (user drinking alcohol) and *Q3*) Tweet is about user doing the activity when tweeting (user tweeting while drunk).

4.2. Goals

Our goal is to develop one classifier for each label available [Hossain et al. 2016]. With *Q1* we want to evaluate the performance of the framework for classifying tweets that mention alcohol consumption. However, *Q1* may also contain tweets related to news or ads linked to alcohol consumption. Therefore, with *Q2* we want to evaluate the effectiveness of the framework for identifying tweets in which the user is consuming alcohol. Finally, with *Q3* we can classify tweets where the user is tweeting while drunk. We consider *Q2* and *Q3* classifiers more relevant with regard to our goal of drunk texting classification, since *Q1* is annotated for the mere mention to alcohol (including ads).

4.3. Experiments

Experiments with semantic enrichment and word embeddings have been performed exclusive of each other. The results of the experiments with semantic enrichment were submitted to Web Intelligence 2018 for evaluation. On the other hand, experiments with word embeddings are in the initial stage and did not present good results, possibly due to the small dataset.

4.3.1. Experiment 1: Semantic enrichment

We run our experiments using the classification algorithm SVM Poly with cross-validation of 5 folds, reserving 80% of the data for training and 20% for tests. We run each experiment 10 times, and performed a statistical paired t test using $\alpha = 0.05$. Table 1

[2]http://www.mturk.com

Table 1. Results with semantic enrichment

	Recall	Precision	F1	+ pp in Recall	+ pp in Precision	+ pp in F1
Q1	87.517	92.151	89.834	-0.157	3.104	1.474
Q2	96.715	81.398	89.057	7.372	-0.264	3.553
Q3	95.182	80.892	88.037	13.79	4.701	9.249

Table 2. Results with Word Embeddings and LSTM (Q3)

Method	F1	Precision	Recall
Baseline	81.392	81.392	81.392
SVM + Semantic Enrichment	88.037	80.892	95.182
Neural Network LSTM	73.737	75.376	72.699
Neural Network LSTM + Semantic Enrichment	76.125	74.483	78.156

summarizes the results obtained with our semantic enrichment framework. The gains in percentage points (pp) with regard to the baseline are highlighted in the columns with a '+' symbol. The most significant ones are observed for recall in actual drunk texting situations (*Q2* and *Q3*, with 7.372 and 13.79 pp, respectively). The results of *Q2* are statistically significant for all metrics. With regard to *Q3*, the results are statistically significant only for recall and F1-measure. The importance of the proposed features was confirmed, as semantic/conceptual features account for 63-73% of the 30 most relevant features for classification *Q2* and *Q3*.

4.3.2. Experiment 2: Word embeddings

Two experiments have been performed to develop a Q3 classifier. The first one considered only the embeddings of the tweets, while the second experiment used the embeddings of the tweets concatenated the embeddings of the semantic enrichment. We learned the embeddings from tweet text, and generated vectors of 32 dimensions. In our experiments the neural network has two intermediate layers with 16 hidden layers each and uses the activation function *reLU*. The results are shown in Table 2, and were inferior compared to the baseline and semantic enrichment only. We hypothesize that these poor results may be due to the size of the dataset from which the embeddings were extracted. We are currently working on the expansion of this dataset in order to perform new experiments.

5. Next Steps

The semantic enrichment framework enabled us to outperform the baseline, and we are currently developing more experiments with word embeddings to verify whether the poor results can be improved with a larger dataset and alternative neural network architectures.

To address the last research question, we are investigating how to combine word embeddings with semantic enrichment. To this end, we are experimenting different architectures of LSTM and CNN neural networks. Such neural networks have layers of embeddings for the text of tweets and auxiliary inputs for semantic enrichment, so that the embedding vector is concatenated to the auxiliary vector. Through the experiments, it will be possible to verify if these neural networks are adequate to combine with word embeddings and semantic enrichment. The use of these strategies together can provide better results in the analysis of short texts in general, not just the identification of drunk texting.

References

Aphinyanaphongs, Y., Ray, B., Statnikov, A., and Krebs, P. (2014). Text classification for automatic detection of alcohol use-related tweets: A feasibility study. In *Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014)*. IEEE.

Chollet, F. and Allaire, J. J. (2018). *Deep Learning with R*. Manning Publications Company.

Culotta, A. (2013). Lightweight methods to estimate influenza rates and alcohol sales volume from twitter messages. *Language resources and evaluation*, 47(1):217–238.

de Informações sobre Saúde e Álcool, C. (2014). Álcool e trânsito. http://www.cisa.org.br/artigo/4692/alcool-transito.php.

Hossain, N., Hu, T., Feizi, R., White, A. M., Luo, J., and Kautz, H. A. (2016). Precise localization of homes and activities: Detecting drinking-while-tweeting patterns in communities. In *ICWSM*, pages 587–590.

Jauch, A., Jaehne, P., and Suendermann, D. (2013). Using text classification to detect alcohol intoxication in speech. In *Proceedings of the 7th Workshop on Emotion and Computing at the 36th German Conference on Artificial Intelligence*.

Johnson, T., Shapiro, R., and Tourangeau, R. (2011). National survey of american attitudes on substance abuse xvi: Teens and parents. *The National Center on Addiction and Substance Abuse*, 2011.

Joshi, A., Mishra, A., AR, B., Bhattacharyya, P., and Carman, M. J. (2015). A computational approach to automatic prediction of drunk-texting. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Association for Computational Linguistics.

Kershaw, D., Rowe, M., and Stacey, P. (2014). Towards tracking and analysing regional alcohol consumption patterns in the uk through the use of social media. In *Proceedings of the 2014 ACM conference on Web science*, pages 220–228. ACM.

Li, Q., Shah, S., Liu, X., Nourbakhsh, A., and Fang, R. (2016). Tweetsift: Tweet topic classification based on entity knowledge base and topic enhanced word embedding. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 2429–2432. ACM.

Organization, W. H. and of Substance Abuse Unit, W. H. O. M. (2014). *Global status report on alcohol and health, 2014*. World Health Organization.

Romero, S. and Becker, K. (2017). Improving the classification of events in tweets using semantic enrichment. In *Proceedings of the International Conference on Web Intelligence*, pages 581–588. ACM.

Schuller, B., Steidl, S., Batliner, A., Schiel, F., and Krajewski, J. (2011). The interspeech 2011 speaker state challenge. In *Twelfth Annual Conference of the International Speech Communication Association*.

West, J. H., Hall, P. C., Hanson, C. L., Prier, K., Giraud-Carrier, C., Neeley, E. S., and Barnes, M. D. (2012). Temporal variability of problem drinking on twitter. *Open Journal of Preventive Medicine*, 2(01):43.