Deep Self-Taught Learning for Detecting Drug Abuse Risk Behavior in Tweets (Invited to Computational Social Network 2018)

Nhat Hai Phan
New Jersey Institute of Technology
44 PUBLICATIONS **171** CITATIONS
SEE PROFILE

Han Hu
New Jersey Institute of Technology
4 PUBLICATIONS **5** CITATIONS
SEE PROFILE

James Geller
New Jersey Institute of Technology
213 PUBLICATIONS **1,939** CITATIONS
SEE PROFILE

Thang N. Dinh
Virginia Commonwealth University
80 PUBLICATIONS **1,143** CITATIONS
SEE PROFILE

Deep Self-Taught Learning for Detecting Drug Abuse Risk Behavior in Tweets

Han Hu[1], NhatHai Phan[1], James Geller[1], Huy Vo[2], Bhole Manasi[1],
Xueqi Huang[2], Sophie Di Lorio[1], Thang Dinh[3], and Soon Ae Chun[4]

[1] New Jersey Institute of Technology, Newark NJ 07102, USA
[2] The City College of New York, New York NY 10031, USA
[3] Virginia Commonwealth University, Richmond VA 23284, USA
[4] City University of New York, Staten Island NY 10314, USA

Abstract. Drug abuse continues to accelerate toward becoming the most severe public health problem in the United States. The ability to detect drug abuse risk behavior at a population scale, such as among the population of Twitter users, can help us to monitor the trend of drug-abuse incidents. Unfortunately, traditional methods do not effectively detect drug abuse risk behavior, given tweets. This is because: (1) Tweets usually are noisy and sparse; and (2) The availability of labeled data is limited. To address these challenging problems, we proposed a deep self-taught learning system to detect and monitor drug abuse risk behaviors in the Twitter sphere, by leveraging a large amount of unlabeled data. Our models automatically augment annotated data: (i) To improve the classification performance, and (ii) To capture the evolving picture of drug abuse on online social media. Our extensive experiment has been conducted on 3 million drug abuse-related tweets with geolocation information. Results show that our approach is highly effective in detecting drug abuse risk behaviors.

Keywords: Deep learning · Self-taught learning · Drug Abuse · Tweets.

1 Introduction

Abuse of prescription drugs and of illicit drugs has been declared a "national emergency" [12, 17]. This crisis includes the misuse and abuse of cannabinoids, opioids, tranquilizers, stimulants, inhalants, and other types of psychoactive drugs, which statistical analysis documents as a rising trend in the United States. The most recent reports from the National Survey on Drug Use and Health (NSDUH) [27] estimate that 10.6% of the total population of people ages 12 years and older (i.e., about 28.6 million people) misuse illicit drugs in 2016, which represents an increase of 0.5% since 2015 [26]. According to the Centers for Disease Control and Prevention (CDC), opioid drugs were involved in 42,249 known deaths in 2016 nationwide [10]. In addition, the number of heroin-involved deaths has been increasing sharply for 5 years, and surpassed the number of firearm homicides in 2015 [22].

In April 2017, the Department of Health and Human Services announced their "Opioid Strategy" to battle the country's drug abuse crisis [12, 17]. In the Opioid Strategy, one of the major aims is to strengthen public health data collection, in order to inform a real-time public health response, and to improve the timeliness, as the epidemic evolves. Given its 100 million daily active users and 500 million daily tweets [11], Twitter has been used as a sufficient and reliable data source for many detection tasks, including epidemiology [30] and public health [1, 4, 5, 15, 21, 28], at the population scale, in a real-time manner. Motivated by these facts and the urgent needs, our goal in this paper is to develop a large-scale computational system to detect drug abuse risk behaviors via Twitter sphere.

Several studies [5–7,18,28,29] have explored the detecting of prescription drug abuse on Twitter. However, the current state-of-the-art approaches and systems are limited in terms of scales and accuracy. They applied strictly keyword-based approaches to collect tweets explicitly mentioning specific drug names, such as Adderall, oxycodone, quetiapine, metformin, cocaine, marijuana, weed, meth, tranquilizer, etc. [5,7,28,29]. That may not reflect the actual distribution of drug abuse risk behaviors on online social media, since: (1) The expressions of drug abuse are often vague, in comparison to common topics, i.e., a lot of slang is used; and (2) Strictly keyword-based approaches are susceptible to lexical ambiguity in natural language [21]. In addition, the drug abuse-related Twitter data usually is very imbalanced, i.e., dominated by non-drug abuse tweets, such as reports, advertisements, etc. The limited availability of annotated drug abuse-related tweets makes it even more challenging to distinguish drug abuse risk behaviors from drug-related advertisements, social discussions, reports, and news. However, existing approaches [5–7, 18, 28, 29] have not been designed to address these challenging issues for drug abuse detection on online social media.

Contributions: To address these challenges, our main contributions are to propose: **(1)** A large-scale drug abuse-related tweets collection mechanism based on supervised machine learning and data crowd-sourcing techniques; and **(2)** A deep self-taught learning algorithm for drug abuse-related tweet detection.

We first collect tweets by filtering tweets through a filter, in which a variety of drug names, colloquialisms and slang terms, and abuse-indicating terms (e.g., overdose, addiction, high, abuse, and even death) are combined together. We manually annotate a small number of tweets as seed tweets, which are used to train machine learning classifiers. Then, the classifiers are applied on unlabeled data to produce machine-labeled tweets. The machine-labeled tweets are verified again by humans on Mechanical Turk, i.e., a crowd-sourcing platform, with good accuracy but at a much lower cost. The new labeled tweets and the seed tweets are combined to form a sufficient and reliable labeled data set for drug abuse risk behavior detection, by applying deep learning models, i.e., convolution neural networks (CNN) [16], long-short term memory (LSTM) models [14], etc.

However, there is still a large amount of unlabeled data, which can be leveraged to significantly improve our models in terms of classification accuracy. Therefore, we further propose a self-taught learning algorithm, in which the training data of our deep learning models will be recursively augmented with a set of new machine-labeled tweets. These new machine-labeled tweets are generated by applying the previously trained deep learning models on a random sample of a huge number of unlabeled tweets, i.e., the 3 million tweets, in our dataset. Note that the set of new machine-labeled tweets possibly has a different distribution from the original training and testing datasets. An extensive experiment conducted on 3 million drug-abuse related tweets with geolocation information shown that our approach is highly effective in detecting drug abuse risk behaviors.

2 Background and Related Work

On the one hand, the traditional studies, such as NSDUH [23], CDC [10], Monitoring the Future [20], the Drug Abuse Warning Network (DAWN) [31], and the MedWatch program [33], are trustworthy sources for getting the general picture of the drug abuse epidemic. On the other hand, many studies that are based on modern online social media, such as Twitter, have shown promising results in drug abuse detection and related topics [1, 4–7, 15, 18, 21, 28, 29]. Butler et al., [18] measured online endorsement of prescription opioid abuse by developing an integrative metric through the lens of Internet communities. Hanson et al. [7] conducted a quantitative analysis on 213,633 tweets discussing "Adderall", a prescription stimulant commonly abused among college students. Hanson et al. [6, 7] focused on how possible drug-abusers interact with and influence others in online social circles. Furthermore, Shutler et al. [29] performed a qualitative analysis of prescription opioid related tweets and found that indication of positive abuse was common.

Our previous work [24] showed the potential of applying machine learning models in drug abuse monitoring system to detect drug abuse-related tweets. Several other works also utilized machine learning methods in detecting and analyzing drug related posts on Twitter. For instance, Sarker et al. [28] proposed a supervised classification model, in which different features such as n-grams, abuse-indicating terms, slang terms, synonyms, etc., are extracted from manually annotated tweets. Then, these features are used to train traditional machine learning models to classify drug abuse tweets and non-abuse tweets. Chary et al. [5] discussed how to use AI models to extract content useful for purposes of toxicovigilance from social media, such as Facebook, Twitter, and Google+. Recently, Coloma et al. [8] illustrated the potential of social media in drug safety surveillance. Furthermore, Twitter and social media have been shown to be reliable sources in analyzing drug abuse and public health-related topics, such as cigarette smoking [1, 21], alcohol use [15], and even cardiac arrest [4].

Although existing studies have shown promising approaches toward detecting drug abuse-related risk behavior and information in Twitter, their performance, in terms of accuracy and scales, is still limited. In this paper, we propose a deep self-taught learning system to leverage a huge number of unlabeled tweets. Self-taught learning [25] is a method that integrates concepts of semi-supervised and multi-task learning, in which the model can exploit examples that are unlabeled and possibly come from a distribution different from the target distribution. It has already been shown that deep neural networks can take advantage of unsupervised learning and unlabeled examples [2,34]. Different from other approaches mainly designed for image processing and object detection [3,9,13,35], our deep self-learning model shows the ability to detect drug abuse risk behavior given noisy and sparse Twitter data with a limited availability of annotated tweets.

3 Deep Self-Taught Learning System for Drug Abuse Risk Behavior Detection

In this section, we present the definition of the drug abuse risk behavior detection problem, our system for collecting tweets, labeling tweets, and our deep self-taught learning approach (Figure 1).

Problem Definition: We use the term "drug abuse risk behaviors" in the wider sense, including misuse and use of Schedule 1 drugs that are illegal; and misuse of Schedule 2 drugs, e.g., Oxycodone, which includes the use thereof for non-medical purposes, and the symptoms and side-effects of misuse. Our task is to develop classification models that can classify a given unlabeled tweet into one of the two classes: a drug abuse risk behavior tweet (**positive**), or a non-risk behavior tweet (**negative**). The main criteria for classifying a tweet as drug-abuse related can be condensed into: *"The existence of abusive activities or endorsements of drugs."* Meanwhile, news, reports, and opinions about drug abuse are the signals of tweets that are not drug abuse-related.

3.1 Collecting and Labeling Tweets

In our crawling system, raw tweets are collected through Twitter APIs. For the collection of focused Twitter data, we use a list of the names of illegal and prescription drug [32] of drugs that have been commonly abused over time, e.g., barbiturates, OxyContin, Ritalin, cocaine, LSD, opiates, heroin, codeine, fentanyl, etc. However, the data is very noisy, since: (1) There was no indication of how to distinguish between drug abuse and legitimate use (of prescription drugs) in collected Tweets, and (2) A lot of slang terms are used in expressing drug abuse-related risk behavior. To address this problem, we added slang terms for drugs and abuse-indicating terms, e.g., "high," "stoned," "blunt," "addicted," etc., into our keyword search library. These slang terms are clearly expressing that the tweets in question were about drug abuse. As a result, most of the collected data is drug abuse-related.

To obtain trustworthy annotated data, there are two integrative steps in labeling tweets. In the first step, 1,794 tweets randomly chosen from collected tweets were manually classified into drug abuse tweets and non-abuse tweets by two professors and three students who have experience in drug abuse behavior study. Several instances of drug abuse tweets and non-abuse tweets are illustrated in Table 1. These labeled tweets are considered seed tweets, which then are used to train traditional binary classifiers, e.g., SVM, Naive Bayes, etc., to predict whether a tweet is a drug abuse-related tweet or not. The trained classifiers are applied on unlabeled tweets to predict their labels, which are called machine labels. In the second step, 5,000 positive machine-labeled tweets with high classification confidence are verified again on Mechanical Turk, which is a well-known crowd-sourcing platform, to improve the trustworthiness and to

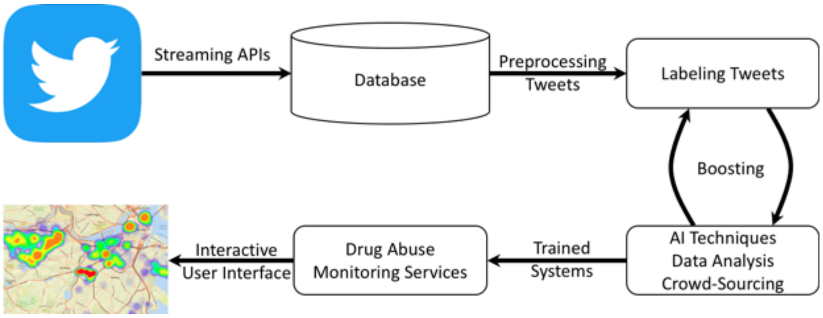


Fig. 1. Drug Abuse Detection System. There are 4 steps as follows: (1) Tweets will be collected through Twitter APIs. (2) Reprocessed tweets will be labeled by humans, AI techniques, and crowd-sourcing techniques. (3) Labeled tweets will be used to augment the training data of our AI models and data analysis tasks to identify drug abuse-related tweets, through a boosting algorithm. And (4) Trained systems will be used in different drug abuse monitoring services and interactive user interfaces.

avoid bias in the annotated data. Our integrative labeling approach results in a reliable and well-balanced annotated data set, with 6,794 labeled tweets. In total, we have collected 3 million drug abuse-related tweets with geo-location, among which 6,794 tweets are labeled.

Tweet Vectorization: Raw tweets need to be first pre-processed, then represented as a vector, before they can be used in training machine learning models. In this study, we choose a commonly used pre-processing pipeline, followed by three different vectorization methods. The pre-processing pipeline consists of following steps:

- The tweets are tokenized and lower-cased. The special entities, i.e., including Emoji, URL, mentions, and hashtags, are removed or replaced with special wordings. The non-word characters, i.e., including HTML symbols, punctuation marks, and foreign characters, are removed. Words with 3 or more repeating characters are reduced to at most 3 successive characters.
- Stop-words are removed according to a custom stop-word list. Stemming is applied using the standard Porter Stemmer.

After the preprocessing steps, common vectorization methods are used to extract features from tweets, including: (1) Term frequency, denoted tf, (2) Tf-idf, and (3) Word2vec [19]. Word2vec is an advanced and effective word embedding method that converts each word to a dense vector of fixed length. We considered two different word2vec models: (i) Custom word2vec which is developed based on our 3 million drug abuse-related tweets, and the model contains 300-dimensional vectors for 1,130,962 words and phrases; and (ii) Google word2vec, which is a well-known pre-trained word2vec vectors built from part of Google News dataset with about 100 billion words, and the model contains 300-dimensional vectors for 3 million words and phrases.

3.2 Deep Self-Taught Learning Approach

By applying both traditional and advanced machine learning models, such as SVM, Naive Bayes, CNN, and LSTM, on the small and static annotated data,

Table 1. Instances of Manually Annotated Drug Abuse Tweets and Non-Abuse Tweets.

	Tweets
Abuse	Ever since my Acid trips like whenever I get super high I just start lightly hallucinating and it's tbh creepy.
	drove like 10 miles on these icy ass roads all to get some weed if imma be locked up in my house for awhile imma need some weed.
	Smoking a blunt at home so much better than going to the woods in Brooksville and puking on yourself Bc you drank too much fireball.
Non-Abuse	Just watched Fear and Loathing in Las Vegas for the first time and I think I should have been on acid to fully understand it.
	today I was asked if I do heroin because I went to Lancaster????
	Morgan told me my Bitmoji looks like a heroin addict?

i.e., 6,794 tweets, we can achieve reasonable classification accuracies of nearly 80%. However, to develop a scalable and trustworthy drug abuse-related risk behavior detection model, we need to: **(1)** Improve classification models to achieve higher accuracy and performance; and **(2)** Leverage the large number of unlabeled tweets, i.e., nearly 3 million tweets related to drug abuse behaviors, to improve the system performance. Therefore, we propose a deep self-taught learning model by repeatedly augmenting the training data with machine-labeled tweets. The pseudo-code of our model is as follows:

Step 1: Initialize labeled data D consisting of 5,794 annotated tweets as the training set. Initialize a testing data T consisting of the remaining 1,000 annotated tweets.

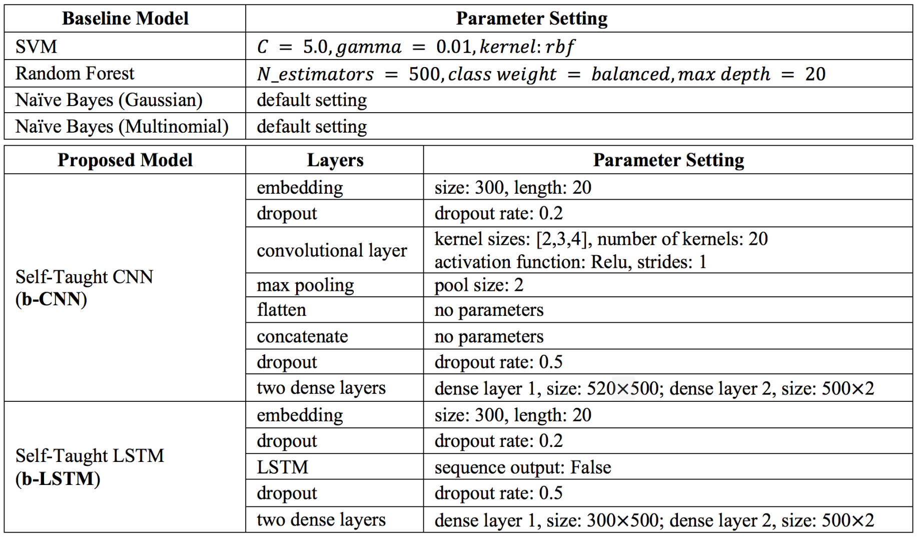
Step 2: Train a binary classification model M using the labeled data D. M could be a CNN model or a LSTM model.

Step 3: Use the model M to label the unlabeled data, which simply consists of 3 million unlabeled tweets. The set of new labeled tweets is denoted as $\overline{D}$, which is also called machine-labeled data.

Step 4: Sample tweets from the machine-labeled data $\overline{D}$ with a high classification confidence, and then add the sampled tweets D^+ into the labeled data D to form a new training dataset: $D = D \cup D^+$. A tweet is considered to have a high classification confidence if it has a classification probability $p \in [0, 1]$ higher than a predefined boosting threshold δ.

Step 5: Repeat Steps 2-4 after k iterations, which is a user-predefined number. Return the trained model M.

With the self-taught learning method, the training data contains the annotated data D which is automatically augmented with highly confident, machine-labeled tweets, in each iteration. That have a great potential to increasing the classification performance of our model over time. In addition, the unlabeled data can be collected from the Twitter APIs in real time, to capture the evolving of drug abuse-related risk behaviors. In the literature, data augmentation approaches have been applied to improve the accuracy of deep learning models [3]. However, the existing approaches [3, 9, 13, 35] are quite different from our proposed model, since they focused on image classification tasks, instead of drug abuse-related risk behavior detection as in our study. To ensure fairness, testing data T is separated from other data sources during the training process.



Baseline Model	Parameter Setting
SVM	$C = 5.0, gamma = 0.01, kernel: rbf$
Random Forest	$N\_estimators = 500, class\ weight = balanced, max\ depth = 20$
Naïve Bayes (Gaussian)	default setting
Naïve Bayes (Multinomial)	default setting

Proposed Model	Layers	Parameter Setting
Self-Taught CNN **(b-CNN)**	embedding	size: 300, length: 20
	dropout	dropout rate: 0.2
	convolutional layer	kernel sizes: [2,3,4], number of kernels: 20 activation function: Relu, strides: 1
	max pooling	pool size: 2
	flatten	no parameters
	concatenate	no parameters
	dropout	dropout rate: 0.5
	two dense layers	dense layer 1, size: 520×500; dense layer 2, size: 500×2
Self-Taught LSTM **(b-LSTM)**	embedding	size: 300, length: 20
	dropout	dropout rate: 0.2
	LSTM	sequence output: False
	dropout	dropout rate: 0.5
	two dense layers	dense layer 1, size: 300×500; dense layer 2, size: 500×2

Fig. 2. Parameter Settings.

4 Experimental Results

To examine the effectiveness and efficiency of our proposed boosting deep learning approaches, we have carried out a series of experiments using a set of 3 million drug abuse-related tweets collected in the past 4 years. We first elaborate details about our dataset, baseline approaches, measures, and model configurations. Then, we introduce our experimental results.

4.1 Experiment Settings

Dataset: In our first data-collecting session, 71,363 tweets were collected. Among them, 1,794 tweets were manually labeled as a **seed dataset**, by two professors and three students with experience in drug abuse behavior study. 280 drug abuse related tweets and 1,514 tweets not related to drug abuse were identified. The seed dataset was used for building the initial machine learning model (i.e., SVM), which was used to further classify the unlabeled 3 million tweets. These 3 million drug abuse-related tweets with geo-location information cover the entire U.S. We then selected 5,000 tweets labeled by the machine learning model (i.e., SVM) with a high confidence level, and rendered them verified by using Mechanical Turk. In total, the number of manually labeled tweets was 6,794, including 3,102 positive labels and 3,677 negative labels.

Baseline and Deep Learning Models: In our experiments, Random Forest (**RF**), Naive Bayes (**NB**), and **SVM** are employed as baseline approaches in the binary classification task, i.e., to classify whether a tweet is a drug abuse-related tweet or not. Figure 2 shows the parameter settings of baseline approaches and the proposed models. Note that for the Naive Bayes method, we use Gaussian Naive Bayes with word2vec embedding. Meanwhile, we use term frequency (i.e., tf) and tf-idf vectorizations for Multinomial Naive Bayes. This is because: (1) The vectors generated by term frequency-based vectorization had a very high number of dimensions and could be only represented by sparse-matrix, which was not supported by Gaussian Naive Bayes; and (2) The Multinomial Naive Bayes required non-negative inputs, but vectors generated by word2vec embedding had negative values. Regarding our self-taught CNN (**b-CNN**) and self-taught LSTM (**b-LSTM**) models, the Adam Optimizer algorithm with the default learning rate is used for training. The number of iterations k is set to 6. All the experiments have been conducted on a single GPU, i.e., NVIDIA GTX TITAN X, 12 GB with 3,072 CUDA cores.

Measures: Accuracy, recall, and $F1$-value are used to validate the effectiveness of the proposed and baseline approaches. Due to the small size and the imbalanced label distribution, we adopted the Monte Carlo Cross-Validation technique. In each run, a fixed number of data instances is sampled (i.e., without replacement) as the testing dataset, and the rest of the data as the training dataset. Multiple runs (i.e., 3 times) are generated for each model in each set of parameters and experimental configurations. We report the average of these runs as result.

4.2 Validation of the Deep Self-Taught Learning Models

Our task of validation concerns three key issues: **(1)** Which parameter configurations are optimal for the baseline models, i.e., SVM, RF, and NB? **(2)** Which boosting model is the best in terms of accuracy, recall, and $F1$-value, given the 6,794 annotated tweets and the 3 million unlabeled tweets? and **(3)** Which vectorization setting is more effective? To address these concerns, our series of experiments are as follows:

Figure 3 illustrates the accuracy, recall, and $F1$-value of each algorithm with different parameter configurations, i.e., term frequency tf, tf-idf, and *word2vec*, on the (annotated) seed dataset. The term *"custom"* is used to indicate the word2vec embedding trained in our own drug abuse-related tweets, compared with the pre-trained Google News word2vec embedding, denoted as "google." It is clear that the SVM model using the custom-trained word2vec embedding achieves the best and the most balanced performance in terms of all three measures, i.e., accuracy, recall, and $F1$-value, at approximately 67%. Other configurations usually have a lower recall, which suggests that the decisions they make bias towards the major class, i.e., tweets that are not drug abuse-related or negative tweets. Note that Naive Bayes had a complete failure, i.e., no correct positive tweet prediction was made, when using tf-idf. Therefore, these results were not shown in Figure 3. From the angle of classifiers, SVM model achieves the best overall performance. Random Forest has slightly less average accuracy than the SVM model, but worse recall and $F1$-value. Furthermore, from the view of vectorization approach, it is clear that word2vec embedding outperforms term frequency and tf-idf in most of the cases.

As shown in the previous experiment, SVM model using the custom-trained word2vec embedding achieves the best performance, we decided to apply the same model structure to compare with our deep self-taught learning approaches. In this experiment, 1,000 labeled tweets were randomly sampled and held out

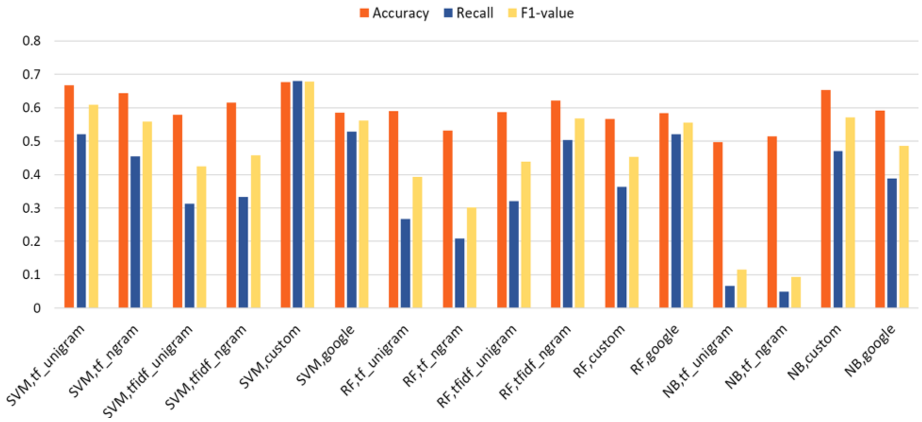


Fig. 3. Accuracy, recall, and $F1$-value of each baseline models on the seed dataset.

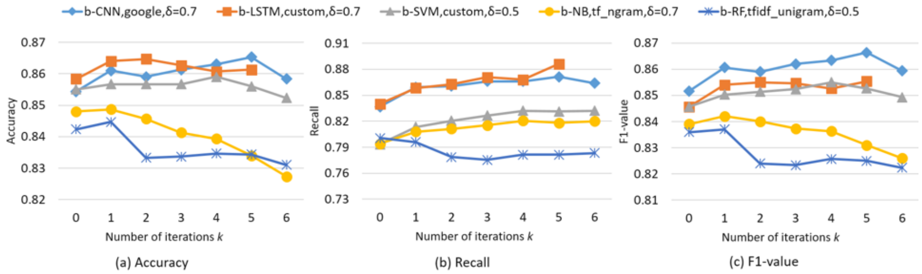
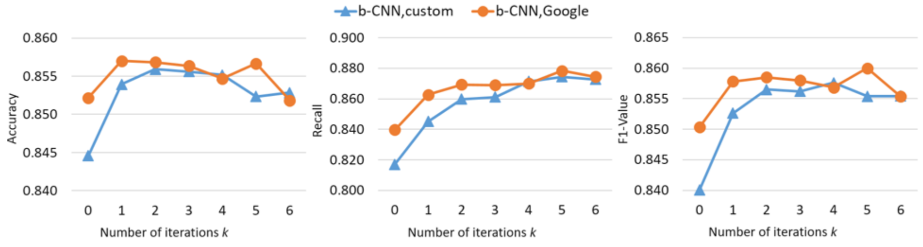


Fig. 4. Accuracy, recall, and $F1$-value of the five self-taught learning models, including b-CNN, b-LSTM, b-SVM, b-NB, and b-RF.

from the 6,794 labeled tweets as testing set. The remaining 5,794 labeled tweets were used as the initial training dataset. At each epoch, 10,000 machine-labeled tweets were randomly sampled from 3 million unlabeled tweets and merged into the training set. Figure 4 shows the experimental results of the five self-taught learning models, including self-taught CNN (**b-CNN**), self-taught LSTM (**b-LSTM**), self-taught SVM (**b-SVM**), self-taught NB (**b-NB**), and self-taught RF (**b-RF**). All configurations of classifiers and vectorization methods are tested. For the sake of clarity, we only illustrate the best-performing setting for each model in Figure 4. It is clear that our proposed deep self-taught learning approaches (i.e., b-LSTM and b-CNN) outperform traditional models, i.e., b-SVM, b-NB, and b-RF, in terms of accuracy, recall, and $F1$-value, in all cases. Our models achieve 86.53%, 88.6%, and 86.63% in terms of accuracy, recall, and $F1$-value correspondingly.

The impact of two different word2vec representations on b-CNN, i.e., the custom word2vec embedding we trained from our corpus, and pre-trained Google News word2vec embedding, is shown in Figure 5. The Google News word2vec achieves 0.1%, 0.4%, and 0.3% improvements in terms of accuracy, recall, and $F1$-value (86.63%, 89%, 86.83%, respectively) compared with the custom trained



(a) Accuracy (b) Recall (c) F1-value

Fig. 5. Performance comparison between custom word2vec embedding and Google News word2vec embedding.

word2vec embedding. In addition, it is clear that Google News word2vec embedding outperforms the custom trained word2vec in most of the cases. This is because the Google News word2vec embedding was trained on a large-scale corpus, which is significantly richer in contextual information, compared with our short, noisy, and sparse Twitter datasets.

5 Discussion

According to our experimental results, our deep self-taught learning models achieved promising performance in drug abuse-related risk behavior detection in Twitter. However, many assumptions call for further experiments. First, how to optimize the classification performance by exploring the correlations among parameters and experimental configurations. For instance, for SVM and RF models, unigram feature works better than n-gram feature on term frequency; however, for tf-idf, it is the opposite situation. Second, the pre-trained Google News word2vec embedding performs better than the custom-trained word2vec embedding may also be situational. These findings indicate the necessity of leveraging size and quality of the training data for training word embedding, given that the available data may better fit the classification task but be short in quantity. Nevertheless, among the measures, recall receives a more significant boost than accuracy and $F1$-value. We may argue that the proposed deep self-taught learning algorithm helped correcting the bias in the classifiers caused by the imbalanced nature of the training dataset. However, more experiments need to be conducted to verify this interesting point.

6 Conclusion

In this paper, we proposed a large-scale drug abuse-related tweet collection mechanism based on supervised machine learning and data crowd-sourcing techniques. Challenges came from the noisy and sparse characteristics of Twitter data, as well as the limited availability of annotated data. To address this problem, we propose a deep self-taught learning algorithm to improve drug abuse-related tweet detection models by leveraging a large number of unlabeled tweets. An extensive experiment and data analysis were carried out on 3 million drug abuse-related tweets with geo-location information, to validate the effectiveness and reliability of our system. Experimental results shown that our models outperform traditional models. In fact, our models correspondingly achieve 86.63%, 89%, and 86.83% in terms of accuracy, recall, and $F1$-value. This is a promising result.

References

1. Aphinyanaphongs, Y., Lulejian, A., Penfold-Brown, D., Bonneau, R., Krebs, P.: Text classification for automatic detection of e-cigarette use and use for smoking cessation from twitter: A feasibility pilot. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing **21**, 480–91 (2016)
2. Bengio, Y.: Learning deep architectures for ai. Found. Trends Mach. Learn. **2**(1), 1–127 (2009)
3. Bettge, A., Roscher, R., Wenzel, S.: Deep self-taught learning for remote sensing image classification. CoRR **abs/1710.07096** (2017)
4. Bosley, J.C., Zhao, N.W., Hill, S., Shofer, F.S., Asch, D.A., Becker, L.B., Merchant, R.M.: Decoding twitter: Surveillance and trends for cardiac arrest and resuscitation communication. Resuscitation **84**(2), 206 – 212 (2013)
5. Chary, M., Genes, N., McKenzie, A., Manini, A.F.: Leveraging social networks for toxicovigilance. Journal of Medical Toxicology **9**(2), 184–191 (Jun 2013)
6. CL, H., B, C., S, B., C, G.C.: An exploration of social circles and prescription drug abuse through twitter. J Med Internet Res **15**(9), e189 (Sept 2013)
7. CL, H., SH, B., C, G.C., JH, W., MD, B., B, H.: Tweaking and tweeting: Exploring twitter for nonmedical use of a psychostimulant drug (adderall) among college students. J Med Internet Res **15**(4), e62 (Apr 2013)
8. Coloma, P.M., Becker, B., Sturkenboom, M.C.J.M., van Mulligen, E.M., Kors, J.A.: Evaluating social media networks in medicines safety surveillance: Two case studies. Drug Safety **38**(10), 921–930 (Oct 2015)
9. Dong, X., Meng, D., Ma, F., Yang, Y.: A dual-network progressive approach to weakly supervised object detection. In: Proceedings of the 2017 ACM on Multimedia Conference. pp. 279–287. MM '17 (2017)
10. on Drug Abuse, N.I.: Overdose death rates, september 15, 2017. National Institute on Drug Abuse (Jan 20, 2018), https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates
11. on Drug Abuse, N.I.: Twitter by the numbers: stats, demographics and fun facts, 2018. Omnicore (March 7, 2018), https://www.omnicoreagency.com/twitter-statistics/
12. Ex-DEA Agent: Opioid crisis fueled by drug industry and congress. CBS 60 Minutes (Oct 17, 2017)
13. Gan, J., Li, L., Zhai, Y., Liu, Y.: Deep self-taught learning for facial beauty prediction. Neurocomputing **144**, 295 – 303 (2014)
14. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997)
15. Hossain, N., Hu, T., Feizi, R., White, A.M., Luo, J., Kautz, H.A.: Precise localization of homes and activities: Detecting drinking-while-tweeting patterns in communities. In: ICWSM (2016)
16. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
17. Marino, T.: Withdraws in latest setback for trump's opioid fight. New York Times (Oct 17, 2017)

18. McNaughton, E.C., Black, R.A., Zulueta, M.G., Budman, S.H., Butler, S.F.: Measuring online endorsement of prescription opioids abuse: an integrative methodology. Pharmacoepidemiology and Drug Safety **21**(10), 1081–1092
19. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. CoRR **abs/1301.3781** (2013)
20. Monitoring the Future: A continuing study of american youth http://www.monitoringthefuture.org
21. Myslín, M., Zhu, S.H., Chapman, W., Conway, M.: Using twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res **15**(8), e174 (Aug 2013)
22. National Institute on Drug Abuse: Gun violence archive, past summary ledgers. (n.d.). Gun Violence Archive (Jan 20, 2018), http://www.gunviolencearchive.org/past-tolls
23. National Poisoning Data System: National Poisoning Data System (Jan 16, 2017), http://www.aapcc.org/data-system/
24. Phan, N., Chun, S.A., Bhole, M., Geller, J.: Enabling real-time drug abuse detection in tweets. In: 2017 IEEE 33rd International Conference on Data Engineering (ICDE). pp. 1510–1514 (2017)
25. Raina, R., Battle, A., Lee, H., Packer, B., Ng, A.Y.: Self-taught learning: Transfer learning from unlabeled data. In: Proceedings of the 24th International Conference on Machine Learning. pp. 759–766. ICML '07 (2007)
26. SAMHSA: Key substance use and mental health indicators in the united states, 2015. SAMHSA (n.d.) (Jan 20, 2018), https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2015/NSDUH-FFR1-2015/NSDUH-FFR1-2015.htm
27. SAMHSA: Key substance use and mental health indicators in the united states, 2016. SAMHSA (n.d.) (Jan 20, 2018), https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2016/NSDUH-FFR1-2016.htm
28. Sarker, A., O'Connor, K., Ginn, R., Scotch, M., Smith, K., Malone, D., Gonzalez, G.: Social media mining for toxicovigilance: Automatic monitoring of prescription medication abuse from twitter. Drug Safety **39**(3), 231–240 (Mar 2016)
29. Shutler, L.e.a.: Prescription opioids in the twittersphere: A contextual analysis of tweets about prescription drugs. Annals of Emergency Medicine **62**(4), S122
30. Signorini, A., Segre, A.M., Polgreen, P.M.: 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), 1–10 (05 2011)
31. Substance Abuse and Mental Health Services Administration Center for Behavioral Health Statistics and Quality (formerly the Office of Applied Studies): The dawn report: highlights of the 2009 drug abuse warning network (dawn) findings on drug-related emergency department visits (Dec 28, 2010)
32. The National Center on Addiction and Substance Abuse: Commonly used illegal drugs (Jan 16, 2017), http://www.centeronaddiction.org/addiction/commonly-used-illegal-drugs
33. US FDA: Medwatch: the fda safety information and adverse event reporting program (Jan 16, 2017), http://www.fda.gov/Safety/MedWatch/
34. Weston, J., Ratle, F., Collobert, R.: Deep learning via semi-supervised embedding. In: Proceedings of the 25th International Conference on Machine Learning. pp. 1168–1175. ICML '08 (2008)
35. Yuan, Y., Liang, X., Wang, X., Yeung, D., Gupta, A.: Temporal dynamic graph LSTM for action-driven video object detection. CoRR **abs/1708.00666** (2017)