Social Media for Opioid Addiction Epidemiology: Automatic Detection of Opioid Addicts from Twitter and Case Studies

Yujie Fan, Yiming Zhang
Department of Computer Science and Electrical Engineering
West Virginia University, WV, USA
{yf0004,ymzhang}@mix.wvu.edu

Yanfang Ye *, Xin li
Department of Computer Science and Electrical Engineering
West Virginia University, WV, USA
{yanfang.ye,xin.li}@mail.wvu.edu

Wanhong Zheng
Department of Behavioral Medicine and Psychiatry
West Virginia University, WV, USA
wzheng@hsc.wvu.edu

ABSTRACT

Opioid (e.g., heroin and morphine) addiction has become one of the largest and deadliest epidemics in the United States. To combat such deadly epidemic, there is an urgent need for novel tools and methodologies to gain new insights into the behavioral processes of opioid abuse and addiction. The role of social media in biomedical knowledge mining has turned into increasingly significant in recent years. In this paper, we propose a novel framework named *AutoDOA* to automatically detect the opioid addicts from Twitter, which can potentially assist in sharpening our understanding toward the behavioral process of opioid abuse and addiction. In *AutoDOA*, to model the users and posted tweets as well as their rich relationships, a structured heterogeneous information network (HIN) is first constructed. Then meta-path based approach is used to formulate similarity measures over users and different similarities are aggregated using Laplacian scores. Based on HIN and the combined meta-path, to reduce the cost of acquiring labeled examples for supervised learning, a transductive classification model is built for automatic opioid addict detection. To the best of our knowledge, this is the first work to apply transductive classification in HIN into drug-addiction domain. Comprehensive experiments on real sample collections from Twitter are conducted to validate the effectiveness of our developed system *AutoDOA* in opioid addict detection by comparisons with other alternate methods. The results and case studies also demonstrate that knowledge from daily-life social media data mining could support a better practice of opioid addiction prevention and treatment.

CCS CONCEPTS

• **Information systems applications** → Data mining; • **Life and medical sciences** → Health care information systems;

KEYWORDS

Social Media; Opioid Addict Detection; Heterogeneous Information Network; Transductive Classification

CIKM'17 , November 6–10, 2017, Singapore, Singapore

ACM ISBN 978-1-4503-4918-5/17/11...$15.00
https://doi.org/10.1145/3132847.3132857

1 INTRODUCTION

Opioids are a group of drugs which include the illegal drug heroin and powerful pain relievers by legal prescription [16], such as morphine and oxycodone. Opioid addiction has become one of the largest and deadliest epidemics in the United States - "more Americans now die every year from drug overdoses than they do from motor vehicle crashes" [27]. In 2014, 4.3 million Americans age 12 and up reported current non-medical use of prescription pain relievers, while 435,000 people had heroin addiction [20]. There was a skyrocketing increase of opioid related death in the past decade: according to National Institute on Drug Abuse (NIDA), in 2014, 18,893 Americans died from opioid analgesics and 10,574 people died from heroin overdose, both reflecting significant increase from 2001 [17]. Opioid addiction has also turned into a serious global concern because of its negative health, social and economic impacts (e.g., family breakdown, domestic violence, child abuse, etc). Opioid addiction is a chronic mental illness that requires long-term treatment and care [14]. It is a psychiatric challenge because of high relapse and drop-out rates. To combat such deadly epidemic, there is an imminent need for novel tools and methodologies to gain new insights into the behavioral processes of opioid abuse and addiction.

The role of social media in biomedical knowledge mining, such as interactive healthcare [6] and drug pharmacology [1], has become more and more important in recent years. Due to the increasing use of the Internet, never-ending growth of data are generated from the social media which offers opportunities for the users to share opinions and experiences freely in online communities. For example, Twitter, as one of the most popular social media platforms, has more than 140 million active users posting over 500 million 140 character tweets every day [23]. A larger number of Twitter users are willing and able to share their experiences of opioid addiction (e.g., "*I have a crippling heroin addiction and it's destroying my life*"), sentiment tendencies (e.g., "*I am snorting so much heroin and i don't even feel happy. just depressed*"), and perceptions of Medication-Assisted Treatment (MAT) (e.g., "*Crack and Heroin are the Problem! I'm a former addict and am on a Subutex Program and its working for me. But still smoke weed but no crime*"), etc. Therefore, the data from social media may contribute information beyond the knowledge of domain professionals (e.g., psychiatrists and epidemics researchers) and could potentially assist in sharpening our understanding toward the behavioral process of opioid abuse and addiction.

To achieve the goal above, in this paper, we propose a novel framework named *AutoDOA* to automatically detect the opioid addicts from Twitter. As the moral says "*man is known by the company he keeps*", to detect whether a user is opioid addict or not, we not only analyze the posted tweets but also his/her social network. For example, a user posted a tweet "*I'll bring some heroin*", which might not be sufficient to determine if he/she is an addict. However, with the information that one of his/her *tweeps* (i.e., Twitter people that follow each other) @ him/her in the tweet "*Let's shoot heroin tonight hahahah. where's the needles at?*", we can conclude that the user is highly possible an opioid addict. Besides the social relations of ***user-follow-user*** (i.e., people follow each other) and ***user-@-tweet*** (i.e., user being @ in the tweet means one person talks publicly to another in this tweet), ***user-like-tweet*** (i.e., user shows appreciation for a tweet) and ***tweet-RT-tweet*** (i.e., a repost of another user's tweet) are also used in our application for opioid addict detection. To model the users and posted tweets as well as their rich semantic relationships, a structured heterogeneous information network (HIN) [22], which is capable to be composed of different types of entities and relations, is first introduced. Then we use meta-path [25] to incorporate higher-level semantics to build up the relatedness of users. In this way, a similarity between two users can not only capture whether they are posting similar tweets but also capture whether they have strong social relations, such as follow/talk to each other, like/RT each other's tweets. Since there can be multiple meta-paths to define different similarities, we incorporate all useful meta-paths with their weights computed by Laplacian scores [7]. To reduce the cost of acquiring labeled samples for supervised learning, we construct a transductive classification model [31] to detect the opioid addicts based on HIN and the combined meta-path. In short, our developed system named *AutoDOA* has the following major traits:

- ***Novel feature representation and similarity measure:*** Instead of using the posted tweets to describe users, we further utilize rich relationships among users and tweets (i.e., user-user, user-tweet, tweet-tweet relations). Based on both content- and relation-based features, the users will be represented by a structural heterogeneous information network (HIN), and a meta-path based approach will be used to link the users. In this way, the similarity measure between two users is an aggregation of different similarities defined by different meta-paths. We then use Laplacian scores to determine the importance of different meta-paths. The proposed solution is comprehensive yet elegant than traditional approaches in feature engineering.
- ***Transductive classification in HIN:*** Inductive classification has been applied in HIN to predict the unlabeled entities. However, it usually requires large number of labeled data to achieve better accuracy. In other words, when training data decreases, its detection accuracy may greatly compromise. In our application, obtaining the labeled data (either opioid addicts or non-opioid addicts) from Twitter is both time-consuming and cost-expensive. To overcome this challenge, we present a transductive classification model in HIN to reduce the cost of acquiring labeled samples for opioid addict detection.
- ***A practical developed system to support a better practice of opioid addiction prevention and treatment:*** We first use our own developed web crawling tools to collect over **4,447,507 tweets** including keywords of opioids (e.g., heroin, morphine, buprenorphine, etc.) from nearly **4,051,423 users** through March 2007 to January 2017, under certain user privacy agreements. To obtain the prelabeled data for training, after data collection, based on the 19,722 tweets including keywords of heroin and morphine corresponding to 2,312 users, **five groups of annotators (i.e., 18 persons)** with knowledge from **domain professional (i.e., psychiatrist)** spent **three months** to label whether they are opioid addicts or not by cross-validations. Based on these data collections and annotations from Twitter, we develop a practical system *AutoDOA* for automatic opioid addict detection and provide comprehensive experimental studies to validate the effectiveness and efficiency of our proposed method in comparisons with other alternate approaches. The results and case studies also demonstrate that knowledge from daily-life social media data mining could support a better practice of opioid addiction prevention and treatment.

The remainder of this paper is organized as follows. Section 2 introduces our system architecture. Section 3 presents our proposed method in detail. In Section 4, based on the real sample collections and annotations from Twitter, we systematically evaluate the performance of our method in comparisons with other alternate approaches. Section 5 presents the case studies based on the detected opioid addicts using our developed system. Section 6 discusses the related work. Finally, Section 7 concludes.

2 SYSTEM ARCHITECTURE

Figure 1 shows the system architecture of our proposed framework *AutoDOA* for automatic opioid addict detection from Twitter, which consists of the following three major components:

- ***Data Collector and Preprocessor.*** Our own developed web crawling tools are used to collect the tweets including keywords of opioids (e.g. heroin, morphine) as well as users' profiles from Twitter. Note that, to protect the users' privacy, we use UserID to represent each individual user whose information is anonymous. For the collected tweets, the preprocessor will further remove all the links, punctuation and stopwords, and conduct lemmatization using Stanford CoreNLP [13].
- ***Feature Extractor.*** A bag-of-words [30] feature vector will be extracted to represent each user's posted tweet(s). Besides, the relationships among users and tweets (i.e., user-user, user-tweet, tweet-tweet relations) will be further analyzed, which include i) *user-follow-user* (i.e., two users follow each other), ii) *user-@-tweet* (i.e., one user is @ in the tweet by the other, which means one talks publicly to the other in this tweet), iii) *user-like-tweet* (i.e., one user shows appreciation for a tweet posted by the other), and iv) *tweet-RT-tweet* (i.e., a repost of the other user's tweet).
- ***Transductive Classifier in HIN.*** In this module, a structural heterogeneous information network (HIN) is first constructed based on the features extracted from the previous components to represent users. Then different meta-paths associated with their corresponding commuting matrices are generated from HIN to measure the similarities between users. Laplacian scores are further computed to weight the importance of different meta-paths. Given the weighted meta-paths, the different commuting matrices is combined to formulate a more powerful similarity measure over users. To reduce the cost of acquiring labeled samples for supervised learning, a transductive classification model in HIN is built to automatically detect the opioid addicts. (See Section 3 for details.)

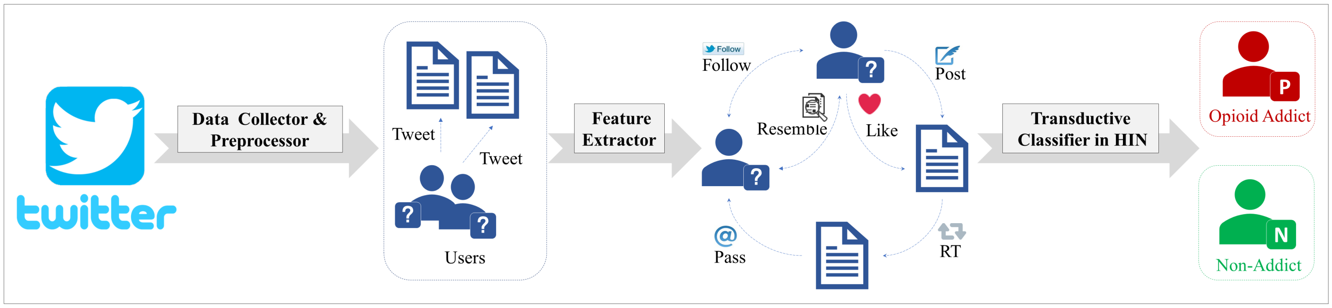

Figure 1: System architecture of *AutoDOA*.

3 PROPOSED METHOD

In this section, we introduce the detailed approaches of how we represent Twitter users, and how to solve the problem of opioid addict detection based on this representation.

3.1 HIN Construction

As mentioned above, to detect the opioid addicts from Twitter, we not only use the tweets a user post but also his/her social network in online communities. To measure the similarity of two users, we consider not only the similarity of their posted tweets but also the closeness of their relationships (e.g., whether they follow or talk to each other, whether they like or RT each other's tweets), which includes the followings.

- ***R1***: To describe the relation of a user and his/her posted tweet, we generate the ***user-post-tweet*** matrix **P** where each element $p_{i,j} \in \{0, 1\}$ denotes if user i posts tweet j.
- ***R2***: To represent how similar two users are based on their posted tweets, we generate the ***user-resemble-user*** matrix **T** where each element $t_{i,j} \in \{0, 1\}$ denotes the semantic similarity of the tweets user i and user j post.
- ***R3***: If two users follow each other, it could imply that they might be friends or have similar interests. To represent such kind of user-user relationship, we generate the ***user-follow-user*** matrix **F** where each element $f_{i,j} \in \{0, 1\}$ denotes whether user i and user j follow each other.
- ***R4***: Like the physical world, user can talk publicly to another in Twitter – if a tweet includes the symbol of @ followed by a user name, it means that this message is passed to the user who is @ in the tweet. To describe this type of user-tweet relationship, we build the ***user-@-tweet*** matrix **A** where each element $a_{i,j} \in \{0, 1\}$ indicates whether tweet i is passed to user j.
- ***R5***: To denote the relation that a user appreciates a tweet, we generate the ***user-like-tweet*** matrix **L** where each element $l_{i,j} \in \{0, 1\}$ means if user i likes tweet j.
- ***R6***: A tweet can be a repost of another tweet. To represent such relationship between two tweets, we build the ***tweet-RT-tweet*** matrix **X** where element $x_{i,j} \in \{0, 1\}$ denotes whether tweet i or tweet j is a repost of the other.

A summary of the description of the above relations and their elements in the relation matrices is shown in Table 1.

Table 1: The description of each matrix and its element.

Matrix	Element	Description
P	$p_{i,j}$	if tweet j is posted by user i, then $p_{i,j} = 1$; otherwise, $p_{i,j} = 0$
T	$t_{i,j}$	similarity of the tweets user i and user j post
F	$f_{i,j}$	if user i and user j follow each other, then $f_{i,j} = 1$; otherwise, $f_{i,j} = 0$
A	$a_{i,j}$	if tweet i is passed to user j, then $a_{i,j} = 1$; otherwise, $a_{i,j} = 0$
L	$l_{i,j}$	if user i likes tweet j, then $l_{i,j} = 1$; otherwise, $l_{i,j} = 0$
X	$x_{i,j}$	if tweet i or tweet j is a repost of the other, then $x_{i,j} = 1$; otherwise, $x_{i,j} = 0$

In order to depict users, tweets and the rich relationships between them, it is important to model them in a proper way so that different kinds of relations can be better and easier handled. We introduce how to use heterogeneous information network, which is capable to be composed of different types of entities and relations, to represent the users by using the features described above (including their posted tweets and the social relationships among them). We first present some concepts related to heterogeneous information network as follows.

Definition 3.1. ***Heterogeneous information network (HIN)*** [22]. An HIN is defined as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with an entity type mapping ϕ: $\mathcal{V} \rightarrow \mathcal{A}$ and a relation type mapping ψ: $\mathcal{E} \rightarrow \mathcal{R}$, where $\mathcal{V}$ denotes the entity set and $\mathcal{E}$ is the relation set, $\mathcal{A}$ denotes the entity type set and $\mathcal{R}$ is the relation type set, and the number of entity types $|\mathcal{A}| > 1$ or the number of relation types $|\mathcal{R}| > 1$.

Definition 3.2. ***Network schema*** [25]. The network schema for an HIN $\mathcal{G}$, denoted as $\mathcal{T}_{\mathcal{G}} = (\mathcal{A}, \mathcal{R})$, is a graph with nodes as entity types from $\mathcal{A}$ and edges as relation types from $\mathcal{R}$.

HIN not only provides the network structure of the data associations, but also provides a high-level abstraction of the categorical association. For the detection of opioid addicts, we have two entity types (i.e., user and tweet) and six types of relations between them as described above. Based on the definitions above, the network schema for HIN in our application is shown in Figure 2.

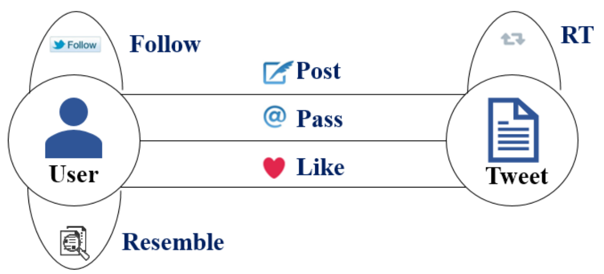

Figure 2: Network schema for HIN.

3.2 Meta-Path Based Similarity

The different types of entities and different relations between them motivate us to use a machine-readable representation to enrich the semantics of similarities among users. Meta-path [25] is used in the concept of HIN to formulate the semantics of higher-order relationships among entities. Here we follow this concept and extend it for the detection of opioid addicts in our application.

Definition 3.3. ***Meta-path*** [25] . A meta-path $\mathcal{P}$ is a path defined on the graph of network schema $\mathcal{T}_{\mathcal{G}} = (\mathcal{A}, \mathcal{R})$, and is denoted in the form of $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} A_{L+1}$, which defines a composite relation $R = R_1 \cdot R_2 \cdot \ldots \cdot R_L$ between types A_1 and A_{L+1}, where $\cdot$ denotes relation composition operator, and L is the length of $\mathcal{P}$.

An example of a meta-path for users is $user \xrightarrow{like} tweet \xrightarrow{like^{-1}} user$, which means two user both like a tweet; another example is $user \xrightarrow{post} tweet \xrightarrow{RT} tweet \xrightarrow{post^{-1}} user$, which denotes a user reposts the other user's tweet. To compute entity similarities using a particular meta-path, we use the following commuting matrix [25] to give a general form.

Definition 3.4. ***Commuting matrix*** [25]. Given a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and its network schema $\mathcal{T}_{\mathcal{G}}$, a commuting matrix $\mathbf{M}_{\mathcal{P}}$ for a meta-path $\mathcal{P} = (A_1 A_2 \ldots A_{L+1})$ is defined as $\mathbf{M}_{\mathcal{P}} = \mathbf{G}_{A_1 A_2} \mathbf{G}_{A_2 A_3} \cdots \mathbf{G}_{A_l A_{L+1}}$, where $\mathbf{G}_{A_i A_j}$ is the adjacency matrix between types A_i and A_j. $\mathbf{M}_{\mathcal{P}}(i, j)$ represents the number of path instances between entity $x_i \in A_1$ and entity $y_i \in A_{L+1}$ under meta-path $\mathcal{P}$.

For the former example, the adjacently matrix between users and tweets is $\mathbf{G}_{user,tweet}$. Then the commuting matrix of users computed using the meta-path $user \xrightarrow{like} tweet \xrightarrow{like^{-1}} user$ is $\mathbf{G}_{user,tweet} \mathbf{G}^T_{user,tweet}$, which is $\mathbf{L}\mathbf{L}^T$ whose element denotes the number of tweets liked by this pair of users. Given a network schema with different types of entities and relations, we can enumerate a lot of meta-paths. In our application, based on the collected data, resting on the six different kinds of relationships described in Table 1, we construct nine meaningful meta-paths as listed in Table 2 for similarity measures over users.

Different meta-paths measure the similarities between two users at different views. For example, (1) the meta-path of *PID-1* calculates the similarity between two users based on their posted tweets; (2) *PID-3* estimates the similarity of two users according to the common tweeps they have, like common friends two persons have in physical world; and (3) *PID-7* weighs the relatedness of two users in the light of the tweets they both talk publicly to another user. In sum, the meta-path of *PID-1* measures the similarity of two users using content-based feature, while the meta-paths of *PID-2* to *PID-9* utilize relation-based features for similarity measures. Instead of using a single meta-path for similarity measure over two users, since HIN can naturally provide us different similarities with different semantic meanings, we propose to combine different meta-paths and weight each of them for user classification (i.e., whether he/she is an opioid addict) in our application. Suppose there are K meta-paths $\mathcal{P}_k$ with their corresponding commuting matrices $\mathbf{M}_{\mathcal{P}_k}$, $k = 1, 2, \ldots, K$. Following [28, 29], after the normalization of each commuting matrix, we combine different meta-paths to form a new similarity measure:

$$\mathbf{M}'(i,j) = \frac{2 \times \sum_{k=1}^{K} w_k \mathbf{M}_{\mathcal{P}_k}(i,j)}{\sum_{k=1}^{K} w_k \mathbf{M}_{\mathcal{P}_k}(i,i) + \sum_{k=1}^{K} w_k \mathbf{M}_{\mathcal{P}_k}(j,j)}, \tag{1}$$

where $\boldsymbol{w} = [w_1, w_2, \ldots, w_K]$ is the weighted vector of the meta-paths. In our application, we use Laplacian score [7] to learn the weight of each meta-path, since it can be computed to reflect the locality preserving power of each feature.

3.3 Transductive Classification in HIN

Compared with inductive classification methods [11, 26] which only use objects with known labels for training, transductive classification models [31, 32] can also utilize the *relatedness* between objects to *propagate* labels and thus reduce the cost of acquiring labeled data for training. In recent years, transductive classification algorithms have been devised in HIN [10] for the applications such as classifying the bibliographic data into research communities [9]. In our case, since it is time-consuming and cost-expensive to obtain the labeled data (either opioid addicts or non-opioid addicts) from Twitter, we propose to use transductive classification in HIN for opioid addict detection. We first introduce the concept of transductive classification in HIN as follow.

Definition 3.5. ***Transductive classification in HIN*** [9]. Given an HIN $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with m types of entities $\mathcal{V} = \bigcup_{i=1}^{m} \mathcal{V}_i$, where $\mathcal{V}_i \in \mathcal{A}_i$ $(i = 1, \ldots, m)$. Suppose $\mathcal{V}'$ is a subset of $\mathcal{V}$ which is with class labels of $C = \{C_1, \ldots, C_c\}$, where c is the number of classes. The classification task is to predict the labels for all the unlabeled entities $\mathcal{V} - \mathcal{V}'$.

In transductive classification model, there are two assumptions of consistency: *Assumption* (1) – entities with tight relationship tend to have a high possibility being in the same class; and *Assumption* (2) – the classification results should consist with the prelabeled information. Following these two assumptions, learning with local and global consistency algorithm (LLGC) in homogeneous information network was proposed in [31] for classification. In our application, we further extend the LLGC framework to classify the entities in heterogeneous information network for opioid addict detection, whose cost function can be denoted as follow:

$$Q(\mathbf{F}) = \frac{1}{2}\Big(\sum_{i,j=0}^{n} \mathbf{M}'_{(i,j)} \parallel \frac{\mathbf{F}_i}{\sqrt{D_{ii}}} - \frac{\mathbf{F}_j}{\sqrt{D_{jj}}} \parallel^2 + \mu \sum_{i=1}^{n} \parallel \mathbf{F}_i - \mathbf{Y}_i \parallel^2\Big), \tag{2}$$

Table 2: The description of each meta-path.

PID	Meta-path	Matrix $\mathbf{M}$	Description of each element m_{ij} in $\mathbf{M}$
1	$u \xrightarrow{T} u$	$\mathbf{T}$	the similarity of the tweets user i and user j post
2	$u \xrightarrow{F} u$	$\mathbf{F}$	whether user i and user j are tweeps (i.e., two users follow each other)
3	$u \xrightarrow{F} u \xrightarrow{F^T} u$	$\mathbf{FF}^T$	the number of common tweeps user i and user j have
4	$u \xrightarrow{L} t \xrightarrow{L^T} u$	$\mathbf{LL}^T$	the number of tweets liked by user i and user j
5	$u \xrightarrow{P} t \xrightarrow{X} t \xrightarrow{X^T} t \xrightarrow{P^T} u$	$\mathbf{PXX}^T\mathbf{P}^T$	the number of tweets reposted by user i and user j
6	$u \xrightarrow{L} t \xrightarrow{X} t \xrightarrow{X^T} t \xrightarrow{L^T} u$	$\mathbf{LXX}^T\mathbf{L}^T$	the number of reposted tweets liked by user i and user j
7	$u \xrightarrow{P} t \xrightarrow{A} u \xrightarrow{A^T} t \xrightarrow{P^T} u$	$\mathbf{PAA}^T\mathbf{P}^T$	the number of tweets user i and user j talk publicly to another user
8	$u \xrightarrow{L} t \xrightarrow{A} u \xrightarrow{A^T} t \xrightarrow{L^T} u$	$\mathbf{LAA}^T\mathbf{L}^T$	the number of tweets talked publicly to another user and liked by both user i and user j
9	$u \xrightarrow{F} u \xrightarrow{L} t \xrightarrow{L^T} u \xrightarrow{F^T} u$	$\mathbf{FLL}^T\mathbf{F}^T$	the number of tweets liked by the tweeps of user i and user j

where n is the number of entities (i.e., Twitter users) in HIN, $\mathbf{M'}$ is the similarity matrix combing different meta-paths, $\mathbf{F}$ is a $n * c$ (c is the number of classes) matrix whose element $F(i, j)$ represents the possibility of user i belonging to class j, $\mathbf{Y}$ is also a $n * c$ matrix containing the prelabeled information, $\mathbf{D}$ is a diagonal matrix whose (i, i)-element is equal to the sum of the i-th row of $\mathbf{M'}$, and $\mu > 0$ is the regularization parameter. The first term of the right-hand side in Eq.(2) is called *smoothness* constraint satisfying *Assumption* (1), which means a good classifying function should not change too much between nearby points; the second term is *fitting* constraint following *Assumption* (2), which indicates a good classifying function should not change too much from the initial label assignment. The parameter μ captures the trade-off between these two competing constraints. Note that the fitting constraint contains both labeled and unlabeled data. In our application, to initialize the matrix $\mathbf{Y}$ [8], a classier is trained (i.e., SVM) resting on the content-based features (i.e., tweets posted by labeled users) to assign an initial label for each unlabeled entity (i.e., user) in HIN.

Based on Eq.(2), the classifying function can be defined as

$$\mathbf{F}^* = \arg\min Q(\mathbf{F}). \tag{3}$$

To obtain the optimal $\mathbf{F}$, we differentiate $Q(\mathbf{F})$ with respect to $\mathbf{F}$ and then have

$$\frac{\partial Q}{\partial \mathbf{F}} = \mathbf{F}^* - \mathbf{S}\mathbf{F}^* + \mu(\mathbf{F}^* - \mathbf{Y}), \tag{4}$$

where $\mathbf{S} = \mathbf{D}^{-1/2}\mathbf{M'}\mathbf{D}^{-1/2}$. Eq.(4) can be further transformed into [31]

$$\mathbf{F}^* = \beta(\mathbf{I} - \alpha\mathbf{S})^{-1}\mathbf{Y}, \tag{5}$$

where $\alpha = \dfrac{1}{1+\mu}$, $\beta = \dfrac{\mu}{1+\mu}$.

4 EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we show three sets of experimental studies using real sample collections from Twitter to fully evaluate the performance of our developed system *AutoDOA* for automatic detection of opioid addicts: (1) in the first set of experiments, we evaluate the effectiveness of HIN and meta-path based similarity; (2) in the second set of experiments, we evaluate the proposed transductive classification method in HIN by comparisons with other inductive classification approaches; and (3) in the third set of experiments, we systematically evaluate our developed system *AutoDOA* in real application for opioid addict detection. The measures for evaluation of different methods are shown in Table 3.

Table 3: Performance indices of opioid addict detection.

Indices	Description
TP	# correctly classified as opioid addicts
TN	# correctly classified as non-opioid addicts
FP	# mistakenly classified as opioid addicts
FN	# mistakenly classified as non-opioid addicts
Precision	$TP/(TP + FP)$
Recall	$TP/(TP + FN)$
ACC	$(TP + TN)/(TP + TN + FP + FN)$
F1	$2 * Precision * Recall/(Precision + Recall)$

4.1 Data Collection and Annotation

To obtain the real data collections from Twitter, we develop web crawling tools to collect the tweets including keywords of opioids (e.g., heroin, morphine, buprenorphine, etc.) as well as users' information (e.g., user profiles) in a period of time. By the date of this paper, we have used our own developed web crawlers to collect over **4,447,507 tweets** including keywords of opioids from nearly **4,051,423 users** through March 2007 to January 2017. All the collected data is anonymous. Note that a user may post multiple tweets including the keywords of opioids. In this case, for each user, we convert his/her posted tweet(s) into a bag-of-words feature vector. After data collection and preprocessing, the six relationships as introduced in Section 3.1 are further extracted.

As addicting to heroin and morphine occupies the majority of today's opioid addiction [16], in this paper, we first study the tweets containing these two keywords and their related users. To obtain the prelabeled data for training, after data collection, based on the

19,722 tweets with keywords of heroin and morphine corresponding to 2,312 users, **five groups of annotators (i.e., 18 persons)** with knowledge from **domain professional (i.e., psychiatrist)** spent **three months** to label whether they are opioid addicts or not by cross-validations. The mutual agreement is above 95%, and only the ones with agreements are retained. The annotated data includes two sets: (1) the first sample set is annotated for preliminary study which consists of 2,466 tweets related to 312 users (152 of which are labeled as opioid addicts, while 160 are non-opioid addicts); (2) the second dataset with larger size is further labeled including 17,256 tweets related to 2,000 users (980 are opioid addicts and 1,020 are non-opioid addicts). The summarization of annotated data is shown in Table 4 (opioid addict is denoted as "+" and non-opioid addict is denoted as "-").

Table 4: Description of annotated datasets.

DataSet	+ users #	+ tweets #	- users #	- tweets #
DB 1	152	1,757	160	709
DB 2	980	8,982	1,020	8,274

4.2 Evaluation of HIN and Meta-path Based Similarity

In this set of experiments, based on the first sample set *DB1* described in Section 4.1, resting on the 2,466 tweets including keywords of heroin and morphine posted by 312 users and the six different kinds of relationships generated among them (***R1-R6***) (as described in Section 3.1), we construct nine meta-paths based on HIN (shown in Table 2) and compare their performances of opioid addict detection by using Support Vector Machine (SVM). We randomly select 90% of the data for training, while the remaining 10% is used for testing, and repeat this procedure 10 times. For each meta-path, we use its related commuting matrix as the kernel fed to SVM. We also evaluate the combined similarity [28, 29] of all the constructed meta-paths using Laplacian scores as their weights [7]. The experimental results are shown in Table 5.

From Table 5, we can see that different meta-paths indeed show different detection performance which can be ranked as: $PID1 \rightarrow PID3 \rightarrow PID2 \rightarrow PID4 \rightarrow PID9 \rightarrow PID6 \rightarrow PID7 \rightarrow PID5 \rightarrow PID8$. Figure 3 also shows the positive correlations between different meta-paths and their related Laplacian scores. Though the content-based meta-path $PID1$ works better than the others, the performances of some relation-based meta-paths such as $PID3$ (i.e., common tweeps two users have) and $PID2$ (i.e., whether the two users are tweeps) are also as competitive as $PID1$, which demonstrates the moral "*a man is known by the company he keeps*". When we combine these meta-paths together, the detection performance is further improved, since it utilizes not only content-based features (users' posted tweets) but also diverse relation-based features which include rich semantic information in opioid addict detection.

We also further evaluate the parameter sensitivity of combined meta-path using SVM with different values of the penalty parameter C. From Figure 4, we can see that in a wide range of numbers, the performance of combined meta-path is stable and not very sensitive to the penalty parameter.

Table 5: Evaluation of HIN and meta-path based similarity.

PID	Meta path	LS	ACC	F1
1	$\mathbf{T}$	0.9855	0.8021	0.7993
2	$\mathbf{F}$	0.8931	0.7779	0.7768
3	$\mathbf{FF}^T$	0.9747	0.7836	0.7788
4	$\mathbf{LL}^T$	0.8764	0.7453	0.7354
5	$\mathbf{PXX}^T\mathbf{P}^T$	0.3681	0.7236	0.7213
6	$\mathbf{LXX}^T\mathbf{L}^T$	0.6179	0.7379	0.7296
7	$\mathbf{PAA}^T\mathbf{P}^T$	0.5664	0.7361	0.7279
8	$\mathbf{LAA}^T\mathbf{L}^T$	0.1239	0.7128	0.7069
9	$\mathbf{FLL}^T\mathbf{F}^T$	0.8597	0.7451	0.7317
10	***Combined Path***	-	**0.8336**	**0.8215**

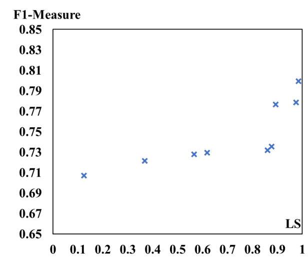

Figure 3: Laplacian score and F1 correlation.

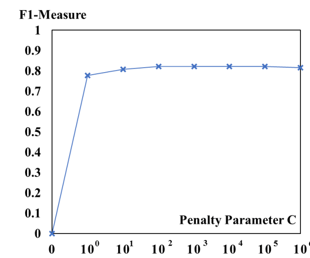

Figure 4: Parameter sensitivity evaluation.

4.3 Evaluation of Transductive Classification in HIN

Since it's time-consuming and cost-expensive to obtain the labeled data (i.e., opioid addicts and non-opioid addicts): as described in Section 4.1, five groups of annotators (i.e., 18 persons) with knowledge from domain professional (i.e., psychiatrist) spent three months to annotate 2,312 users based on their posted tweets, profiles, and relation-based information. To reduce the cost of acquiring labeled samples for supervised learning, we look for an effective classification model whose detection performance won't greatly compromise when the size of training data decreases. In this section, based on the first sample set *DB1* described in Section 4.1, we randomly select a portion of the labeled data (range from 90% to 50%) to simulate the experiments. To evaluate the performances for automatic detection of opioid addicts when training size decreases, we compare our developed system *AutoDOA*, which integrates transductive classification model in HIN, with three typical inductive classification methods, i.e., Naive Bayes (NB), Decision Tree (DT) and Support Vector Machine (SVM), based on the extracted features discussed in Section 3.1. In other words, we put all HIN-related entities and relations as features for different methods to learn. The experimental results are illustrated in Table 6 and Figure 5, from which we can see that *AutoDOA* using transductive classification model in HIN and combined meta-path similarity outperforms NB, DT and SVM in automatic opioid addict detection. To put this into perspective, for the three inductive classification methods, when training data decreases (from 90% to 50%), their detection performances greatly

Table 6: Comparison of transductive and inductive classifications in HIN when training samples decrease.

		With different sizes of training samples				
		90%	80%	70%	60%	50%
NB	ACC	0.7245	0.6963	0.6687	0.6448	0.6115
	F1	0.7103	0.6884	0.6636	0.6319	0.6068
DT	ACC	0.7569	0.7245	0.7060	0.6674	0.6387
	F1	0.7206	0.7013	0.6740	0.6478	0.6118
SVM	ACC	0.8336	0.8167	0.7752	0.7426	0.7021
	F1	0.8215	0.8002	0.7536	0.7241	0.6956
AutoDOA	ACC	**0.8578**	**0.8454**	**0.8297**	**0.8087**	**0.8016**
	F1	**0.8448**	**0.8369**	**0.8242**	**0.8015**	**0.7967**

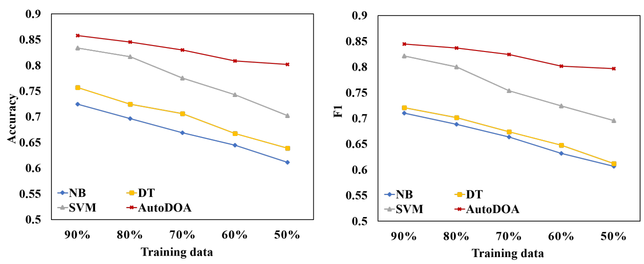

Figure 5: Comparisons of different classification methods when training samples decrease (ACCs and F1s)

compromise (i.e., both ACC and F1 drop more than 10%); while the detection performances of *AutoDOA* don't change too much as training samples decrease. This is because *AutoDOA* not only uses the information from training data for prediction but also utilizes the relatedness among training samples and testing objects to propagate labels.

4.4 Evaluation of Proposed Method in Real Application

Based on the real data collection from Twitter and larger dataset annotated by human experts (i.e. *DB2* described in Section 4.1 including 2,000 labeled users, 980 of which are opioid addicts and 1,020 are non-opioid addicts), we systematically evaluate the performance of our developed system *AutoDOA*, including the detection effectiveness and scalability.

Figure 6 shows the overall receiver operating characteristic (ROC) curves of *AutoDOA* based on the ten-fold cross validations. From Figure 6, we can see that our proposed framework *AutoDOA* achieves an impressive 0.9148 average TP rate at the 0.142 average FP rate for opioid addict detection from Twitter. Note that many of the existing microblog classifications [15, 19], especially on Twitter, achieve the TP rates at the range of [0.647, 0.877].

We also further evaluate the training time of *AutoDOA* with different sizes of the training data sets. Figure 7 shows the scalability of our proposed method. It is illustrated that the running time is quadratic to the number of training samples. When dealing with more data, approximation or parallel algorithms should be developed. However, as shown in Figure 8, for such automatic opioid addict

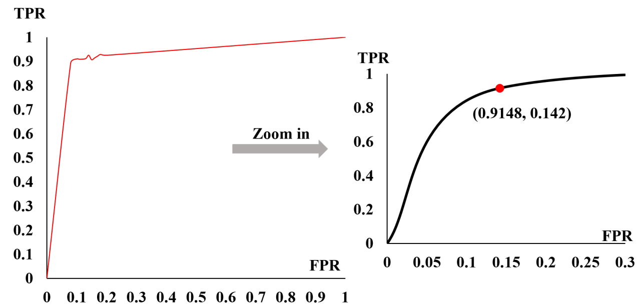

Figure 6: *Left*: ROC curve of AutoDOA, *Right*: Zoom-in.

detection problem, the need of more labels is not as important as the need of more expressive representations of data. The reasons behind this are: (1) *AutoDOA* using HIN representation and meta-path similarity measure well describes the rich semantic relationships among users and tweets (i.e., user-user, user-tweet, and tweet-tweet relations); (2) *AutoDOA* applying transductive classification in HIN utilizes the relatedness among training samples and testing objects to propagate labels and thus reduces the cost for acquiring labeled samples. Therefore, for practical use, our approach is feasible for real application in automatic opioid addict detection to support a better practice of opioid addiction prevention and treatment.

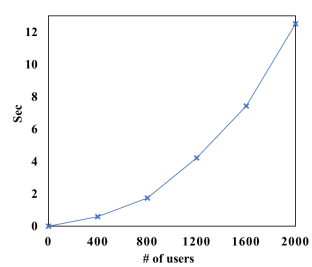

Figure 7: Scalability evaluation of *AutoDOA*.

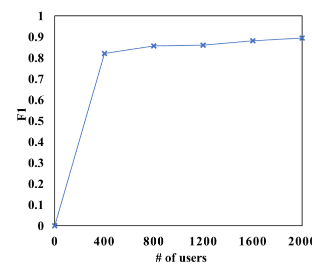

Figure 8: Comparisons when training data sizes vary.

5 CASE STUDIES

In this section, after the automatic detection of opioid addicts from Twitter using our developed system *AutoDOA*, to better understand opioid addiction epidemic and public perceptions toward Medication-Assisted Treatment (MAT), we further analyze the data and conduct some case studies based on the detected opioid addicts.

- **Case Study 1: epidemic surveillance of opioid abuse and addiction in the U.S.** To better understand the distribution and opioid addiction epidemic, a series of spatio-temporal statistics such as geo-location distribution analysis associated with different timelines are performed based on our detected opioid addicts. By making use of the profile data of Tweeter users which indicates their related geo-locations, Figure 9 shows the distribution of our detected opioid addicts (i.e., 1,132 heroin addicts) in different states of the U.S. from February 2016 to February 2017 (the darker color the more severe epidemic the state has). Similar to the statistics of heroin-related overdose from CDC [4]: Ohio, New York, Illinois, West Virginia and Maryland have larger numbers of heroin addicts than the others in the U.S. Though in some rural areas Twitter is not the primary platform for social communication and may have its biases, this case study still clearly reflects the actual status of opioid addiction epidemic in the U.S., which demonstrates that using social media for epidemic surveillance of opioid abuse and addiction is practical and feasible.

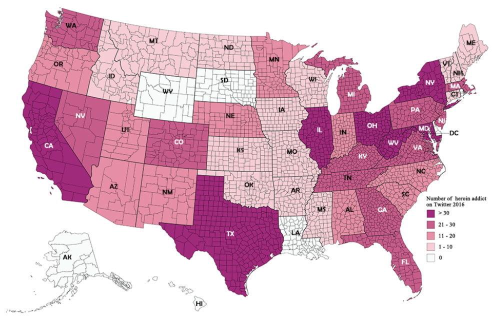

Figure 9: Distribution of heroin addicts on Twitter.

- **Case Study 2: deeper understanding of public perceptions toward MAT.** To further assess the public perceptions and stigmas of MAT, we randomly sample 30% of our detected opioid addicts (i.e. 356 heroin addicts) to further analyze and categorize their posted tweets. Table 7 shows different categories of the posted tweets as well as two cases of public perceptions toward MAT. The study reveals that (1) some users show appreciations of MAT (e.g., "*Methadone is an effective treatment which needs to be made more available in the U.S.*"), while (2) some of them still have significant stigmas toward MAT who mistakenly think of MAT as "*one drug replaced by another*". It also shows that there is a remarkable treatment gap suggesting the majority of people who need behavioral health treatment but have not received it due to various reasons (e.g., public stigma, financial burden).
- **Case Study 3: identification of the influential users to advertise the best practice of MAT.** To promote the perception of MAT, we believe it is best to first locate the group of users with apparent stigmas toward MAT and then use social network analysis to identify the most likely authoritative users that could influence the group of interests. In this study, the assumption is further validated. For the users who post their perceptions of MAT, we further analyze their social networks (e.g., their tweeps and people who like/repost/reply their tweets) and find that they actively interact with their virtual friends on Twitter, which indicates that they could be the influential users who have the power of authoritative sources in the linked environment and thus can help promote the perception of MAT. Figure 10 shows two examples of potential influential users who can help advertise the best practice of MAT.

The above case studies based on the automatically detected opioid addcits using our developed system *AutoDOA* demonstrate that knowledge from daily-life social media data mining could support a better practice of opioid addiction prevention and treatment.

Table 7: Deep understanding of detected heroin addicts.

Categories of the posted tweets	# tweets	Percentage
Need heroin	130	36.52%
Shoot heroin	103	28.92%
Love heroin	82	23.03%
Bought heroin	13	3.65%
Reasons of heroin addiction	11	3.09%
Attitude toward heroin addiction	8	2.25%
Perceptions toward MATs	5	1.40%
Consequences of heroin addiction	3	0.84%
Seek for help	1	0.28%
Examples of perceptions toward MAT		
1. **Appreciation for MAT:** "*heroin; I think this model of treatment (methadone) needs to be made available in the US, as it's the most effective treatment for opioid.*" 2. **Stigma toward MAT:** "*heroin mat is utterly fraudulent but expensive. You can treat all you want, the addicts will go right back to it.*"		

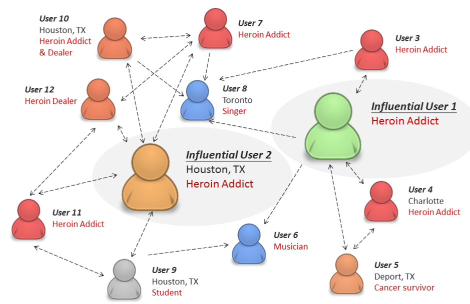

Figure 10: Identification of the influential users

6 RELATED WORK

In recent years, the role of social media in biomedical knowledge mining, such as interactive healthcare and drug pharmacology, has become increasingly important. For example, based on users' posted tweets, a machine learning-based concept extraction system ADRMine was introduced for adverse drug reactions (ADRs) analysis [18]; Support Vector Machine (SVM) classifiers based on the content of twitter messages were built to find drug users as well as the potential adverse events [2]. Unfortunately, the application of social media data analytics into drug-addiction domain has been scarce in the literature with few exceptions: Cameron et al. [3] developed a novel semantic web platform called PREDOSE (Prescription Drug Abuse Online Surveillance and Epidemiology) to facilitate the epidemiologic study of prescription and related drug abuse practices using social media; Sarker et al. [21] designed an automatic supervised classification technique to distinguish posts containing signals of medication abuse. However, most of these studies merely used content-based features (e.g., posted tweets or messages) for their applications. Actually, the relations among users and the generated contents are also very important for targeted user detection. Different from the existing works in drug-addition domain, in this paper, we propose to not only utilize users' posted tweets but also the relationships among users and tweets (i.e., user-user, user-tweet, tweet-tweet relations) for opioid addict detection from Twitter. Based on the extracted features, the users are represented by a structured heterogeneous information network (HIN), and a meta-path based approach is used to link the users.

Heterogeneous information network (HIN) has been intensively studied in recent years. Typically, HIN is used to model different types of entities and relations [22]. It has been applied to various applications, such as scientific publication network analysis [24, 25] and document analysis based on knowledge graph [28, 29]. Different from traditional graph similarities, such as shortest path, the similarity defined on HIN, i.e., PathSim [25], is more likely a natural extension to dot product. Different from the simple dot product, the similarity defined over HIN considers the semantics of the network meta-data. In our application, to measure the similarities over users, we develop a similarity based on multiple meta-paths using an unsupervised meta-path weighting mechanism [28, 29]. To solve the classification problem in HIN, compared with inductive methods [11, 26], transductive classification [5, 9, 10, 12] was proposed to reduce the cost of acquiring labeled samples in supervised learning. However, it has yet applied in biomedical knowledge mining. In this paper, we explore how to construct an effective transductive classification model in HIN for opioid addict detection from Twitter.

7 CONCLUSION

In this paper, we propose a framework called *AutoDOA* to automatically detect the opioid addicts from Twitter. In *AutoDOA*, we first construct a heterogeneous information network (HIN) to leverage the information of users and tweets as well as the rich relationships among them (i.e., user-user, user-tweet, tweet-tweet relations), which gives the user a higher-level semantic representation. Then, Laplacian scores are computed to weight different generated meta-paths and a combined meta-path is used for similarity measure over users. To reduce the cost of acquiring labeled samples, a transductive classification model in HIN is then built for opioid addict detection. The promising experimental results on the real data collections from Twitter demonstrate that our framework outperforms other alternate methods. The case studies also show that knowledge from daily-life social media data mining could support a better practice of opioid addiction prevention and treatment.

ACKNOWLEDGMENTS

This work is supported by WVU NT-NS grant (2016-2017). The work of Y. Fan, Y. Zhang, and Y. Ye is partially supported by the U.S. National Science Foundation under grant CNS-1618629.

REFERENCES

[1] Fadi M Alkhateeb, Kevin A Clauson, and David A Latif. 2011. Pharmacist use of social media. *International Journal of Pharmacy Practice* 19, 2 (2011), 140–142.
[2] Jiang Bian, Umit Topaloglu, and Fan Yu. 2012. Towards large-scale twitter mining for drug-related adverse events. In *SHB*. ACM, 25–32.
[3] Delroy Cameron, Gary A Smith, Raminta Daniulaityte, Amit P Sheth, Drashti Dave, Lu Chen, Gaurish Anand, Robert Carlson, Kera Z Watkins, and Russel Falck. 2013. PREDOSE: a semantic web platform for drug abuse epidemiology using social media. *JBI* 46, 6 (2013), 985–997.
[4] CDC. 2015. *Heroin Overdose Data*. https://www.cdc.gov/drugoverdose/data/heroin.html.
[5] Mukul Gupta, Pradeep Kumar, and Bharat Bhasker. 2017. HeteClass: A Meta-path based framework for transductive classification of objects in heterogeneous information networks. *Expert Systems with Applications* 68 (2017), 106–122.
[6] Carleen Hawn. 2009. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. In *Health affairs*. 361–368.
[7] Xiaofei He, Deng Cai, and Partha Niyogi. 2006. Laplacian score for feature selection. In *NIPS*. 507–514.
[8] Yann Jacob, Ludovic Denoyer, and Patrick Gallinari. 2011. Classification and annotation in social corpora using multiple relations. In *CIKM*. ACM, 1215–1220.
[9] Ming Ji, Yizhou Sun, Marina Danilevsky, Jiawei Han, and Jing Gao. 2010. Graph regularized transductive classification on heterogeneous information networks. In *PKDD*. Springer, 570–586.
[10] Xiang Li, Ben Kao, Yudian Zheng, and Zhipeng Huang. 2016. On Transductive Classification in Heterogeneous Information Networks. In *CIKM*. ACM, 811–820.
[11] Qing Lu and Lise Getoor. 2003. Link-based classification. In *ICML*, Vol. 3. 496–503.
[12] Chen Luo, Renchu Guan, Zhe Wang, and Chenghua Lin. 2014. Hetpathmine: A novel transductive classification algorithm on heterogeneous information networks. In *ECIR*. Springer, 210–221.
[13] Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit.. In *ACL (System Demonstrations)*. 55–60.
[14] A Thomas McLellan, David C Lewis, Charles P O'brien, and Herbert D Kleber. 2000. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. *Jama* 284, 13 (2000), 1689–1695.
[15] Kimberly McManus, Emily K Mallory, Rachel L Goldfeder, Winston A Haynes, and Jonathan D Tatum. 2015. Mining Twitter data to improve detection of schizophrenia. *AMIA* 2015 (2015), 122.
[16] NIDA. 2014. *Addiction to Opioids: Heroin and Prescription Drug Abuse*. https://www.drugabuse.gov/about-nida/legislative-activities/testimony-to-congress/2016/americas-addiction-to-opioids-heroin-prescription-drug-abuse.
[17] NIDA. 2015. *Overdose Death Rates*. http://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates.
[18] Azadeh Nikfarjam, Abeed Sarker, Karen O'Connor, Rachel Ginn, and Graciela Gonzalez. 2015. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *JAMIA* (2015), ocu041.
[19] Marco Pennacchiotti and Ana-Maria Popescu. 2011. Democrats, republicans and starbucks afficionados: user classification in twitter. In *KDD*. ACM, 430–438.
[20] SAMHSA. 2015. *Behavioral Health Trends in the United States: Results from the 2014 National Survey on Drug Use and Health*. https://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf.
[21] Abeed Sarker, Karen O'Connor, Rachel Ginn, Matthew Scotch, Karen Smith, Dan Malone, and Graciela Gonzalez. 2016. Social media mining for toxicovigilance: automatic monitoring of prescription medication abuse from Twitter. *DS* 39, 3 (2016), 231–240.
[22] Chuan Shi, Yitong Li, Jiawei Zhang, Yizhou Sun, and S Yu Philip. 2017. A survey of heterogeneous information network analysis. *TKDE* 29, 1 (2017), 17–37.
[23] Internet Live Stats. 2013. *Twitter Usage Statistics*. http://www.internetlivestats.com/twitter-statistics/.
[24] Yizhou Sun, Rick Barber, Manish Gupta, Charu C Aggarwal, and Jiawei Han. 2011. Co-author relationship prediction in heterogeneous bibliographic networks. In *ASONAM*. IEEE, 121–128.
[25] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S Yu, and Tianyi Wu. 2011. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *VLDB* 4, 11 (2011), 992–1003.
[26] Ben Taskar, Pieter Abbeel, and Daphne Koller. 2002. Discriminative probabilistic models for relational data. In *Proceedings of the 8th Conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 485–492.
[27] U.S.DEA. 2015. *2015 National Drug Threat Assessment Summary*.
[28] Chenguang Wang, Yangqiu Song, Haoran Li, Ming Zhang, and Jiawei Han. 2015. Knowsim: A document similarity measure on structured heterogeneous information networks. In *ICDM*. IEEE, 1015–1020.
[29] Chenguang Wang, Yangqiu Song, Haoran Li, Ming Zhang, and Jiawei Han. 2016. Text Classification with Heterogeneous Information Network Kernels.. In *AAAI*. 2130–2136.
[30] Jun Yang, Yu-Gang Jiang, Alexander G Hauptmann, and Chong-Wah Ngo. 2007. Evaluating bag-of-visual-words representations in scene classification. In *MIR*. ACM, 197–206.
[31] Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf. 2003. Learning with local and global consistency.. In *NIPS*, Vol. 16. 321–328.
[32] Xiaojin Zhu, Zoubin Ghahramani, and John D Lafferty. 2003. Semi-supervised learning using gaussian fields and harmonic functions. In *ICML*. 912–919.