Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)

Automatic Opioid User Detection From Twitter: Transductive Ensemble Built On Different Meta-graph Based Similarities Over Heterogeneous Information Network

Yujie Fan, Yiming Zhang, Yanfang Ye *, Xin Li
Department of Computer Science and Electrical Engineering, West Virginia University, WV, USA
{yf0004,ymzhang}@mix.wvu.edu, {yanfang.ye,xin.li}@mail.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, in this paper, we propose a novel framework named *HinOPU* to automatically detect opioid users from Twitter, which will assist in sharpening our understanding toward the behavioral process of opioid addiction and treatment. In *HinOPU*, to model the users and the posted tweets as well as their rich relationships, we introduce structured heterogeneous information network (HIN) for representation. Afterwards, we use meta-graph based approach to characterize the semantic relatedness over users; we then formulate different similarities over users based on different meta-graphs on HIN. To reduce the cost of acquiring labeled samples for supervised learning, we propose a transductive classification method to build the base classifiers based on different similarities formulated by different meta-graphs. Then, to further improve the detection accuracy, we construct an ensemble to combine different predictions from different base classifiers for opioid user detection. Comprehensive experiments on real sample collections from Twitter are conducted to validate the effectiveness of *HinOPU* in opioid user detection by comparisons with other alternate methods.

1 Introduction

Opioids are a class of drugs including the illegal drug heroin and powerful pain relievers by legal prescription, such as morphine and oxycodone [NIDA, 2014]. Opioid addiction has become one of the largest and deadliest epidemics in the U.S. [NIDA, 2017], which has also turned into a serious national crisis that affects public health as well as social and economic welfare. There was a skyrocketing increase of opioid related death in the past decade: according to Centers for Disease Control and Prevention (CDC), in 2015, more than 32,000 Americans died from opioid drugs overdoses and over 13,000 people died from heroin overdose, both reflecting significant increase from 2002 [NIDA, 2017]. Opioid addiction is a chronic mental illness that requires long-term treatment and care [McLellan *et al.*, 2000]. Although Medication Assisted Treatment (MAT) using methadone or buprenorphine has been proven to provide best outcomes for opioid addiction recovery, stigma (i.e., bias) associated with MAT has limited its utilization [Saloner and Karthikeyan, 2015]. Therefore, there is an urgent need for novel tools and methodologies to gain new insights into the behavioral processes of opioid addiction and treatment.

In recent years, the role of social media in biomedical knowledge mining has become more and more important [Hawn, 2009; Alkhateeb *et al.*, 2011]. Due to the increasing use of the Internet, never-ending growth of data is generated from the social media which offers opportunities for the users to freely share opinions and experiences in online communities. For example, Twitter, as one of the most popular social media platforms, has averaged at 328 million monthly active users as of the second quarter of 2017 [Neha, 2017]. Many Twitter users are willing to share their experiences of using opioids (e.g., "*It started with anorexia as I was 11, with 13 I was on heroin. At every time I leave heroin I fall into anorexia*"), and perceptions toward MAT (e.g., "*I stopped with heroin bc I'm having the methadone, that is the most effective treatment.*"). Thus, the data from social media may contribute information beyond the knowledge of domain professionals (e.g., psychiatrists and epidemics researchers) and could assist in sharpening our understanding toward the behavioral process of opioid addiction and treatment.

To combat the opioid addiction epidemic and promote the practice of MAT, in this paper, we propose a novel framework named *HinOPU*, a transductive ensemble classification model built on different meta-graph based similarities over heterogeneous information network (HIN), to automatically detect opioid users from Twitter. To model the users and the posted tweets as well as their rich semantic relationships, in *HinOPU*, we first introduce a HIN [Sun *et al.*, 2011] for representation. To capture the complex relationship (e.g. two users are relevant if they have posted tweets that discussed the same topic, mentioned same people, and also reposted same tweets from others), we use meta-graphs [Zhao *et al.*, 2017] to incorporate higher-level semantics to build up relatedness over users. Different similarities over users can then

*Corresponding author.

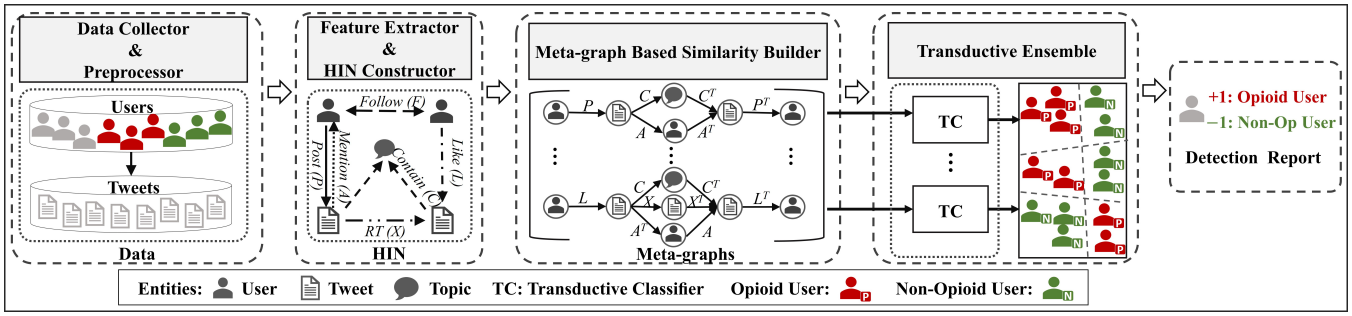

Figure 1: System architecture of *HinOPU*.

be formulated by different meta-graphs. As obtaining the labeled data (either opioid users or not) from Twitter is both time-consuming and cost-expensive, to reduce the cost of acquiring labeled samples for supervised learning, we propose a transductive classification method to build the base classifiers based on different similarities formulated by different meta-graphs; to further improve the detection accuracy, we then present an ensemble approach to combine different predictions from different base classifiers for opioid user detection.

2 System Architecture

The system architecture of *HinOPU* is shown in Figure.1, which consists of the following four major components:

- ***Data collector and preprocessor.*** We first develop web crawling tools to collect opioid-related tweets (i.e., the tweets include opioid names or their common street names) as well as users' profiles from Twitter. For privacy protection, UserID is used to anonymize each user. For the collected tweets, the preprocessor will remove all the links, punctuations, stopwords, and conduct lemmatization using Stanford CoreNLP [Manning *et al.*, 2014]; as street names can be used to imply opioids (e.g., *smack* is used to refer *heroin*) and may also have different meanings when given different contexts, we will then perform word sense disambiguation (WSD) [Zhong and Ng, 2010] to decide if a tweet containing the street name is an opioid-related tweet.
- ***Feature extractor and HIN constructor.*** A bag-of-words feature vector is extracted to represent each user's posted tweet(s); then the relationships among users, tweets and topics will be further analyzed, such as, i) *user-post-tweet*, ii) *user-like-tweet*, iii) *user-follow-user*, iv) *tweet-mention-user*, v) *tweet-RT-tweet*, and vi) *tweet-contain-topic*. Based on these extracted features, a structural HIN will be constructed. (See Section 3.1.)
- ***Meta-graph based similarity builder.*** In this module, different meta-graphs are first built from HIN to capture the relatedness between users. Then, we integrate similarity of users' posted tweets and relatedness depicted by each meta-graph to formulate a set of similarity measures over users. (See Section 3.2.)
- ***Transductive ensemble.*** Based on different similarities formulated by different meta-graphs in the previous module, to reduce the cost of acquiring labeled samples for supervised learning, transductive classification method is proposed to build the base classifiers; then an ensemble approach is presented to combine different results from the base classifiers to make final predictions. (See Section 3.3.)

3 Proposed Method

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

3.1 HIN Construction

To detect whether a user is an opioid user or not from Twitter, besides the users' posted tweets, the following kinds of relationships are also considered:

- ***R1***: To describe the relation of a user and his/her posted tweet, we build the ***user-post-tweet*** matrix $\mathbf{P}$ where each element $p_{i,j} \in \{0, 1\}$ denotes if user i posts tweet j.
- ***R2***: To denote the relation that a user appreciates a tweet, we generate the ***user-like-tweet*** matrix $\mathbf{L}$ where each element $l_{i,j} \in \{0, 1\}$ means if user i likes tweet j.
- ***R3***: Two users can follow each other in Twitter. To represent such kind of user-user relationship, we generate the ***user-follow-user*** matrix $\mathbf{F}$ where each element $f_{i,j} \in \{0, 1\}$ denotes whether user i and user j follow each other.
- ***R4***: If a tweet includes symbol @ followed by a user name, it means that the user is mentioned and talked publicly in the tweet. To describe such tweet-user relationship, we build the ***tweet-mention-user*** matrix $\mathbf{A}$ where each element $a_{i,j} \in \{0, 1\}$ indicates if tweet i mentions user j.
- ***R5***: A tweet can be a repost of another tweet. To represent such relationship between two tweets, we build the ***tweet-RT-tweet*** matrix $\mathbf{X}$ where element $x_{i,j} \in \{0, 1\}$ denotes whether tweet i or tweet j is a repost of the other.
- ***R6***: To represent the relation that a tweet contains a specific topic, we generate the ***tweet-contain-topic*** matrix $\mathbf{C}$ whose element $c_{i,j} \in \{0, 1\}$ indicates whether tweet i contains topic j. In our case, we use Latent Dirichlet Allocation [Blei *et al.*, 2003] for topic extraction from posted tweets.

In order to depict users, tweets, topics and the rich relationships among them (i.e., user-user, user-tweet, tweet-tweet, tweet-topic relations), 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 HIN to represent the users by using the features described above. We first present the concept related to HIN as below.

Definition 1 *A **heterogeneous information network (HIN)** [Sun* et al.*, 2011] is 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 $|\mathcal{R}| > 1$. The **network schema** [Sun* et al.*, 2011] for $\mathcal{G}$, 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 $\mathcal{R}$.*

In our case, we have three entity types (i.e., user, tweet and topic) and six relation types as aforementioned. Based on the definitions above, the network schema for HIN in our application is shown in Figure. 2.

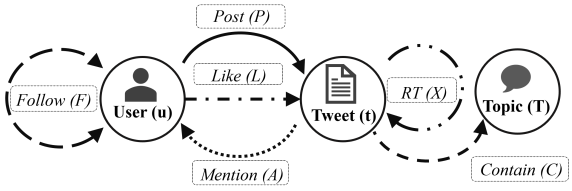

Figure 2: Network schema for HIN.

3.2 Meta-graph Based Similarity

The different types of entities and relations motivate us to use a machine-readable representation to enrich the semantics of relatedness among users. To handle this, the concept of meta-path has been proposed.

Definition 2 *A **meta-path** [Sun* et al.*, 2011] $\mathcal{P}$ 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 $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \ldots \xrightarrow{R_L} A_{L+1}$, 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 based on HIN schema shown in Figure.2 is: $user \xrightarrow{post} tweet \xrightarrow{contain} topic \xrightarrow{contain^{-1}} tweet \xrightarrow{post^{-1}} user$, which states that two users can be connected through their posted tweets containing same topics. Although meta-path has been shown to be useful for relatedness measure between users, it fails to capture a more complex relationship, e.g., two users have posted tweets that discussed the same topic and also mentioned same people. This calls for a better characterization to handle such complex relationship. Meta-graph is proposed to use a directed acyclic graph of entity and relation types to capture more complex relationship between two HIN entities. The concept of meta-graph is given as follow.

Definition 3 *A **meta-graph** [Zhao* et al.*, 2017] $\mathcal{S}$ is a directed acyclic graph with a single source node n_s and a single target node n_t, defined on a HIN $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with schema $\mathcal{T}_{\mathcal{G}} = (\mathcal{A}, \mathcal{R})$. Formally, a meta-graph is defined as $\mathcal{S} = (\mathcal{V}_S, \mathcal{E}_S, \mathcal{A}_S, \mathcal{R}_S, n_s, n_t)$, where $\mathcal{V}_S \subseteq \mathcal{V}$, $\mathcal{E}_S \subseteq \mathcal{E}$ constrained by $\mathcal{A}_S \subseteq \mathcal{A}$ and $\mathcal{R}_S \subseteq \mathcal{R}$, respectively.*

Based on the HIN schema displayed in Figure.2, we generate seven meaningful meta-graphs to characterize the relatedness over users (i.e., ***SID1–SID7*** shown in Figure.3). Different meta-graphs measure the relatedness between two users at different views. For example, ***SID1*** depicts that two users are related if they have posted tweets that discussed same topic and also mentioned same people; while ***SID7*** describes that two users are connected if there're tweets they like that discussed same topic, mentioned same people, and also reposted same tweet from others. Actually, a meta-path is a special case of a meta-graph. But meta-graph is capable to express more complex relationship in a convenient way. In Figure.3, the meta-paths *PID1–PID8* (*left*) are the special cases of the constructed meta-graphs *SID1–SID7* (*right*).

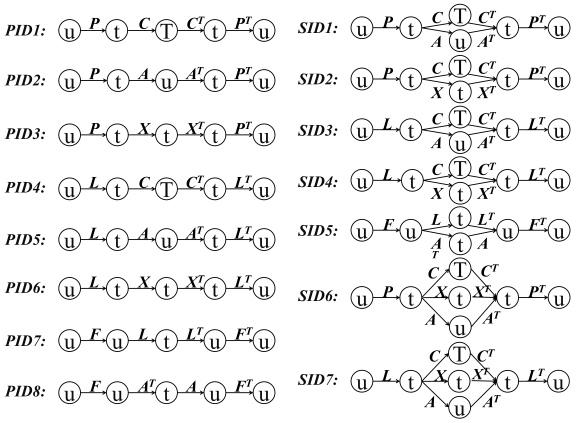

Figure 3: Meta-paths (*left*) and meta-graphs (*right*). (The symbols in this figure are the abbreviations shown in Figure. 2.)

To measure the relatedness over users using a particular meta-graph designed above, we use commuting matrix [Sun *et al.*, 2011; Zhao *et al.*, 2017] to compute the counting-based similarity matrix for a meta-graph. Take *SID1* as an example, the commuting matrix of users computed using *SID1* is $\mathbf{M}_{\mathcal{S}_1} = \mathbf{P}[(\mathbf{C}\mathbf{C}^T) \circ (\mathbf{A}\mathbf{A}^T)]\mathbf{P}^T$, where **P, C, A** are the adjacency matrices between two corresponding entity types, $\circ$ denotes the Hadamard product of two matrices. $\mathbf{M}_{\mathcal{S}_1}(i, j)$ denotes the number of tweet pairs posted by user i and user j which contain same topic and also mention same people.

After characterizing the relatedness between users, we utilize both content- and relation-based information to measure the similarity over users: we integrate similarity of users' posted tweets and relatedness depicted by meta-graph to form a similarity measure matrix over users. The similarity matrix over users is denoted as $\mathbf{M}'$, whose element is a combination of content-based similarity and meta-graph based relatedness. We define the similarity matrix $\mathbf{M}'_{\mathcal{S}_k}$ as: $\mathbf{M}'_{\mathcal{S}_k}(i, j) = [1 + log(\mathbf{M}_{\mathcal{S}_k}(i, j) + 1)] \times tSim(i, j)$, where $\mathbf{M}_{\mathcal{S}_k}(i, j)$ is the relatedness between user *i* and *j* under meta-graph $\mathcal{S}_k$, $tSim(i, j)$ is the cosine similarity between two bag-of-words feature vectors generated from user *i* and *j*'s posted tweets.

3.3 Transductive Ensemble Classification in HIN

Base Classifier

Compared with inductive classification methods [Lu and Getoor, 2003] which only use objects with known labels for training, transductive classification models [Zhu *et al.*, 2003; Ye *et al.*, 2011] can utilize the *relatedness* between objects to *propagate* labels and thus reduce the cost of acquiring labeled data. In our case, since it is time-consuming and cost-expensive to obtain the labeled data (either opioid users or not) from Twitter, we propose to use transductive classification in HIN to build the base classifier. The concepts of transductive classification in HIN are introduced as follows.

Definition 4 *Given a HIN $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with m types of entities $\mathcal{V} = \bigcup_{i=1}^{m} \mathcal{V}_i$, where $\mathcal{V}_i \in \mathcal{A}_i$ $(i = 1, ..., m)$. Suppose $\mathcal{V}'$ is a subset of $\mathcal{V}$ which is with class labels of $C = \{C_1, ..., C_c\}$, where c is the number of classes. The task of* ***transductive classification in HIN*** *[Ji* et al.*, 2010] is to predict the labels for all the unlabeled entities $\mathcal{V} - \mathcal{V}'$.*

Though we have three types of entities (i.e., user, tweet, and topic), we are only interested in classifying one certain entity (i.e., user), denoted as A_u.To solve this problem, topology shrinking sub-network (TSSN) was proposed in [Li *et al.*, 2016] to perform the transductive classification in HIN. We here further extend the definition of TSSN in our application.

Definition 5 *Given a HIN $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the* ***topology shrinking sub-network (TSSN)*** *[Li* et al.*, 2016] of a certain entity type A_u derived from a meta-graph $\mathcal{S}$ is a graph whose nodes consist of only entities of type A_u and edges connect entities that are related by $\mathcal{S}$. Formally, the TSSN is the graph $G_{A_u,\mathcal{S}} = (\mathcal{V}_u, E_{A_u}, R_{A_u})$, where $E_{A_u} = \{e_{i,j} | s_{v_i \rightsquigarrow v_j} \in \mathcal{S}, v_i, v_j \in \mathcal{V}_u\}$ and $R_{A_u} = \{r_{i,j} | r_{i,j}$ is the weight of $e_{i,j}\}$.*

In our case, R_{A_u} is the similarity matrix $\mathbf{M}'_{\mathcal{S}}$ formulated by the meta-graph $\mathcal{S}$ in HIN described in Section 3.2.

In transductive classification model, there are two assumptions of consistency: (1) *smoothness constraint*: entities with tight relationship tend to have a high possibility being in the same class; and (2) *fitting constraint*: the classification results should consist with the prelabeled information. Following these two assumptions, Gaussian Fields Harmonic Function (GFHF) [Zhu *et al.*, 2003] method was proposed for classification in homogeneous information network. In our application, given the prelabeled information and a TSSN $G_{A_u,\mathcal{S}} = (\mathcal{V}_u, E_{A_u}, R_{A_u})$, we propose an approach (named *TCHin*) to further extend the GFHF in HIN to assign the labels (either opioid users or non-opioid users) to the unlabeled entities (i.e., users), which works as follow.

Suppose we have h labeled users and q unlabeled users. Let $\mathbf{Y}$ be a $(h + q) * c$ matrix (c is the number of classes), whose element $y_{i,j}$ denotes if user i belongs to class j. Let $\mathbf{Y}_h$ be the top *h* rows of $\mathbf{Y}$ and $\mathbf{Y}_q$ the remaining *q* rows. To initialize $\mathbf{Y}_q$, a random class is assigned to each element in $\mathbf{Y}_q$. Based on the TSSN, $\mathbf{D}$ is defined as a diagonal matrix whose (i, i)-element is equal to the sum of the *i*-th row of the weight matrix R_{A_u} of $G_{A_u,\mathcal{S}}$. Then, a probabilistic transition matrix $\mathbf{B}$ is calculated via $\mathbf{B} = \mathbf{D}^{-1} R_{A_u}$. We then have the following algorithm: (1) propagate labels $\mathbf{Y} = \mathbf{BY}$; (2) clamp $\mathbf{Y}_h$ to the labeled users; (3) repeat until $\mathbf{Y}$ converges. Note that $\mathbf{Y}_h$ won't change since it is clamped in step (2). To get $\mathbf{Y}_q$, we partition $\mathbf{B}$ into four sub-matrices:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{hh} & \mathbf{B}_{hq} \\ \mathbf{B}_{qh} & \mathbf{B}_{qq} \end{bmatrix},$$

where $\mathbf{B}_{hh}$ denotes the probabilistic transition matrix for labeled users to labeled users, $\mathbf{B}_{hq}$ is the one for labeled users to unlabeled users - and vise versa for $\mathbf{B}_{qh}$, and $\mathbf{B}_{qq}$ is the one for unlabeled users to unlabeled users. Based on the above algorithm, we have the following iterative propagation rule:

$$\mathbf{Y}^{t+1} = \begin{bmatrix} \mathbf{Y}_h^{t+1} \\ \mathbf{Y}_q^{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B}_{qh} & \mathbf{B}_{qq} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_h^{t} \\ \mathbf{Y}_q^{t} \end{bmatrix}.$$

This equation will converge to [Zhu *et al.*, 2003]:

$$\mathbf{Y}' = \begin{bmatrix} \mathbf{Y}'_h \\ \mathbf{Y}'_q \end{bmatrix} = \lim_{t \to \infty} \mathbf{Y}^t = \begin{bmatrix} \mathbf{Y}_h \\ (\mathbf{I} - \mathbf{B}_{qq})^{-1} \mathbf{B}_{qh} \mathbf{Y}_h \end{bmatrix}.$$

Then we have $\mathbf{Y}'_q = (\mathbf{I} - \mathbf{B}_{qq})^{-1} \mathbf{B}_{qh} \mathbf{Y}_h$, where the element $y'_{i,j}$ in $\mathbf{Y}'_q$ is the possibility of user *i* being in class *j*. At this point, we can get the class label for each unlabeled user *i*: $Class(i) = \text{argmax}_{j \in \{1,...,c\}} \{y'_{i,j}\}$.

Ensemble

Ensemble methods are a popular way to overcome instability and increase performance in many machine learning tasks [Ye *et al.*, 2009b]. An ensemble of classifiers is a set of classifiers whose individual decisions are combined in some way (e.g., by weighted or unweighted voting) to classify new samples, which is shown to be much more accurate than the individual classifiers that make them up [Ye *et al.*, 2009a]. Typically, an ensemble can be decomposed into two components [Ye *et al.*, 2017]: (1) create base classifiers with necessary accuracy and diversity; (2) aggregate all the outputs of the base classifiers into a numeric value as the final output of the ensemble.

As described in Section 3.2, different meta-graphs formulate the similarities between users at different views (i.e., with different semantic meanings). To detect opioid users from Twitter, we first use the above proposed transductive classification method in HIN (*TCHin*) to build the base classifiers based on different similarities formulated by different meta-graphs (i.e., ***SID1–SID7*** shown in Figure.3); we then use weighted voting (i.e., the weight of k-th classifier w_k is determined by its learning accuracy) to aggregate the classification results from different base classifier to make final prediction. The proposed transduction ensemble classification algorithm (named *TEHin*) is shown in Algorithm 1.

4 Experimental Results And Analysis

In this section, we fully evaluate the performance of our developed system *HinOPU* for opioid user detection.

4.1 Data Collection and Annotation

To obtain the data from Twitter, we develop a set of tools to crawl the tweets including keywords of opioids (e.g., *heroin*, *morphine*) and their commonly used street names (e.g., *smack*, *morpho*), as well as users' profiles in a period of time. By the date, ***4,537,615 tweets*** from ***4,015,276 users***

Algorithm 1 Transductive Ensemble Classification (*TEHin*)

Input:
h labeled users, q unlabeled users
meta-graphs $\{\mathcal{S}_k\}_{k=1}^{7}$
similarities $\mathbf{M}'_{\mathcal{S}_k}$ formulated by meta-graphs $\{\mathcal{S}_k\}_{k=1}^{7}$
Output:
Labels for q unlabeled users

1: **for** k=1 to 7 **do**
2: Build TSSN $G^{(k)}_{A_u,\mathcal{S}}$ based on $\mathcal{S}_k$ and $\mathbf{M}'_{\mathcal{S}_k}$;
3: Initialize $\mathbf{Y}$;
4: Build the base classifier using *TCHin* to get $\mathbf{Y}'^{(k)}_q$;
5: **end for**
6: Combine the results using weighted voting: $Class(i) = \text{argmax}_{j\in\{1,\ldots,c\}} \sum_{k=1}^{7} w_k \mathbf{Y}'^{(k)}_q(i,j)$.

through Nov. 2007 to Nov. 2017 have been collected and preprocessed in *HinOPU*. As heroin addiction occupies the majority of today's opioid addiction and morphine is one of the most popular legal prescriptions [NIDA, 2014], we first study the tweets containing these two keywords and their commonly used street names, as well as the related users. To obtain the ground truth, seven groups of annotators (i.e., ***35 persons***) with knowledge from domain professional (i.e., psychiatrist) ***spent 180 days to label*** whether the users are opioid users or not and whether the tweets (including opioid names and their street names) are opioid-related tweets or not, based on the posted tweets, users' profiles, and all other relation-based information by cross-validations. The mutual agreement is above 95%, and only the ones with agreements are retained. The annotated datasets include:

- $\mathbf{DB}_1$: This labeled dataset is based on the *heroin*-related tweets (i.e., tweets include keywords of heroin or its common street names) and their corresponding users. It consists of 11,560 users (5,660 are labeled as opioid users and 5,900 are non-opioid users) related to 98,610 tweets (53,695 are posted by opioid users and 44,915 are posted by non-opioid users).
- $\mathbf{DB}_2$: This labeled dataset is based on the *morphine*-related tweets (i.e., tweets include keywords of morphine or its common street names) and their corresponding users. It contains 11,870 users (5,900 are opioid users and 5,970 are not) related to 74,490 tweets (43,890 are posted by opioid users and 30,600 are posted by non-opioid users).

4.2 Baseline Methods

We validate the performance of our proposed method in *HinOPU* for opioid users detection by comparisons with different groups of baseline methods.

First, we evaluate different types of features for opioid user detection from Twitter. Our proposed method is general for HINs. Thus, a natural baseline is to see whether the knowledge we add in should be represented as HIN instead of other features. Here we compare four types of features: two traditional features (BoW and BoW+Relations) and two HIN-based features (meta-path and meta-graph).

- **BoW (*f-1*)**: Each user's collected tweet(s) is represented as a bag-of-words feature vector.
- **BoW+Relations (*f-2*)**: This augments bag-of-words with relations of ***R1–R6*** described in Section 3.1 as flat features.
- **Meta-path (*f-3*)**: We combine the generated eight meta-paths *PID1–PID8* (*left* of Figure.3) to formulate the relatedness over users, where Laplacian scores [He *et al.*, 2006] are used to learn the weights; then content-based information is incorporated using the method proposed in Section 3.2 to construct a meta-path based similarity matrix.
- **Meta-graph (*f-4*)**: Similar to the above setting, we utilize the seven meta-graphs *SID1–SID7* (*right* of Figure.3) to form a meta-graph based similarity matrix.

Second, based on the HIN-based features (meta-path and meta-graph), we evaluate our proposed learning algorithms (i.e., *TCHin* and *TEHin*) with other learning mechanisms, including two inductive methods (i.e., Naive Bayes (NB) and Support Vector Machine (SVM)) and a transductive method (i.e., Learning with Local and Global Consistency (LLGC) algorithm proposed in [Zhou *et al.*, 2003]). For SVM, we use LibSVM and the penalty is empirically set to be 1,000 while other parameters are set by default.

4.3 Comparisons

Based on $\mathbf{DB}_1$ and $\mathbf{DB}_2$, we show the comparison results for all the above methods for opioid user detection. We randomly select a portion of the labeled data (ranging from 90% to 50%) from $\mathbf{DB}_1$ and $\mathbf{DB}_2$ to simulate the experiments. We use F1 measure for evaluations and the experimental results are illustrated in Table 1.

From the results, we can see that feature engineering (***f-2***: BoW+Relations) helps the performance of machine learning, since the rich semantics encoded in different types of relations can bring more information. However, the use of this information is simply flat features, i.e., concatenation of different features altogether, which is less informative than the HIN-based features (***f-3***: meta-path and ***f-4***: meta-graph). Furthermore, the meta-graph feature does perform better than meta-path. The reason behind this is that meta-graphs are more expressive to characterize a complex relatedness over users than meta-paths.

Based on the HIN-based features (***f-3***: meta-path and ***f-4***: meta-graph), we can see that transductive methods (i.e., *LLGC*, *TCHin* and *TEHin*) outperform inductive methods (i.e., NB and SVM) in both datasets especially when labeled data is insufficient. To put this into perspective, for these two inductive methods, when labeled data decreases from 90% to 50%, their performances greatly compromise (i.e., F1 drops more than 10%); while the performances of transductive methods decrease more elegantly as labeled samples decrease. This is because transductive methods not only use the information of labeled data for prediction but also utilize the relatedness among labeled and unlabeled samples to propagate labels. We also observe that (1) our proposed transductive classification algorithm *TCHin* outperforms LLGC over HIN in opioid user detection under the same experimental settings; (2) the ensemble method *TEHin* achieves better detection performance compared with *TCHin*, which demonstrates

Method		NB				SVM				LLGC		*TCHin*		***TEHin***
Settings		***f-1***	***f-2***	***f-3***	***f-4***	***f-1***	***f-2***	***f-3***	***f-4***	***f-3***	***f-4***	***f-3***	***f-4***	***Alg.1***
$\mathbf{DB}_1$	90%	0.743	0.756	0.776	0.798	0.809	0.818	0.839	0.860	0.840	0.859	0.854	0.872	**0.886**
	70%	0.696	0.705	0.739	0.760	0.761	0.771	0.791	0.814	0.819	0.830	0.832	0.850	**0.863**
	50%	0.638	0.648	0.674	0.697	0.706	0.716	0.735	0.759	0.798	0.815	0.811	0.832	**0.841**
$\mathbf{DB}_2$	90%	0.754	0.766	0.785	0.804	0.805	0.811	0.841	0.861	0.843	0.862	0.858	0.875	**0.888**
	70%	0.708	0.717	0.738	0.755	0.766	0.779	0.800	0.811	0.819	0.831	0.830	0.849	**0.870**
	50%	0.652	0.660	0.681	0.704	0.703	0.711	0.739	0.760	0.789	0.803	0.802	0.821	**0.841**

Table 1: Comparisons of different learning methods in opioid user detection when labeled data decreases from 90% to 50%.

TEHin can overcome the instability and improve the detection performance compared with individual classifier.

4.4 Stability and Scalability

In this set of experiments, we systematically evaluate the stability and scalability of the developed system *HinOPU*. Figure 4 shows the overall receiver operating characteristic (ROC) curves of *HinOPU* based on 10-fold cross validations: *HinOPU* achieves an impressive 0.923 average TP rate at the 0.134 average FP rate in $\mathbf{DB}_1$ and 0.931 average TP rate at the 0.131 average FP rate in $\mathbf{DB}_2$. We also further evaluate the running time of *HinOPU* with different sizes of the datasets. Figure 5 shows its scalability, which illustrates that the running time is quadratic to the number of samples. When dealing with more data, approximation or parallel algorithms can be developed. From the results and analysis above, for practical use, our approach is feasible for real application in automatic opioid user detection.

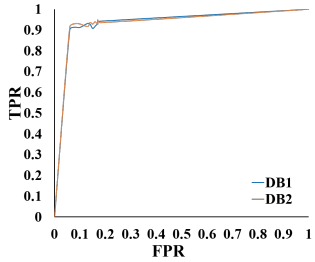

Figure 4: Stability evaluation.

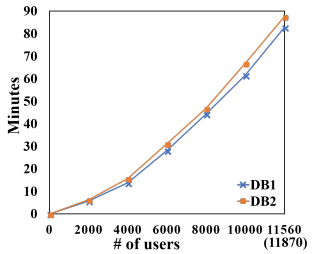

Figure 5: Scalability evaluation.

5 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 [Nikfarjam *et al.*, 2015]; SVM classifiers based on the content of twitter messages were built to find drug users and the potential adverse events [Bian *et al.*, 2012]. However, most of these studies merely used content-based features (e.g., posted tweets or messages) for their applications. Actually, the relations among users are also important for targeted user detection. In this paper, we propose to utilize not only the users' posted tweets but also the relationships among users and tweets (i.e., user-user, user-tweet, tweet-tweet, and tweet-topic relations) for opioid user detection from Twitter; we then present HIN for feature representations.

HIN has been intensively studied in recent years. It has been applied to various applications, such as scientific publication network analysis [Sun *et al.*, 2011], social network analysis for Twitter users [Fan *et al.*, 2017], and malware detection [Hou *et al.*, 2017]. Several studies have already investigated the use of HIN information for relevance computation, however, most of them only used meta-path [Sun *et al.*, 2011] to measure the similarity. Such simple path structure fails to capture a more complex relationship between two entities. To address this problem, meta-graph [Zhao *et al.*, 2017] were proposed to measure the proximity between two entities. In this paper, for opioid user detection, we take into consideration of different meta-graphs which characterize the relatedness over Twitter users at different views (i.e., with different semantic meanings). As acquiring the labeled information (either opioid users or not) from Twitter is both time-consuming and cost-expensive, we propose to use transductive classification [Ji *et al.*, 2010; Li *et al.*, 2016] to solve this learning problem in HIN. To combine different results generated from different base transductive classifiers guided by different meta-graphs, an ensemble approach is further presented to make final predictions.

6 Conclusion

In this paper, we propose a framework called *HinOPU* to automatically detect opioid users from Twitter. In *HinOPU*, we first construct a HIN to leverage the information of users, tweets and topics as well as the rich relationships among them, which gives the user a higher-level semantic representation. Then, meta-graph based approach is used to characterize the semantic relatedness over users, based on which different similarities over users are formulated. To reduce the cost of acquiring labeled samples, a transductive classification method (*TCHin*) is proposed to build the base classifiers based on different similarities guided by different meta-graphs; then an ensemble approach (*TEHin*) is presented to combine different predictions from the base classifiers for opioid user detection. The promising experimental results based on the real data collections from Twitter demonstrate that *HinOPU* outperforms other alternate methods.

Acknowledgments

This work is partially supported by the U.S. NSF Grant (CNS-1618629) and WV HEPC Grant (HEPC.dsr.18.5).

References

[Alkhateeb *et al.*, 2011] Fadi M Alkhateeb, Kevin A Clauson, and David A Latif. Pharmacist use of social media. *International Journal of Pharmacy Practice*, 19(2):140–142, 2011.

[Bian *et al.*, 2012] Jiang Bian, Umit Topaloglu, and Fan Yu. Towards large-scale twitter mining for drug-related adverse events. In *SHB*, pages 25–32. ACM, 2012.

[Blei *et al.*, 2003] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *JMLR*, 3(Jan):993–1022, 2003.

[Fan *et al.*, 2017] Yujie Fan, Yiming Zhang, Yanfang Ye, Wanhong Zheng, et al. Social media for opioid addiction epidemiology: Automatic detection of opioid addicts from twitter and case studies. In *CIKM*, pages 1259–1267. ACM, 2017.

[Hawn, 2009] Carleen Hawn. Take two aspirin and tweet me in the morning: how twitter, facebook, and other social media are reshaping health care. *Health Affairs*, 28(2):361–368, 2009.

[He *et al.*, 2006] Xiaofei He, Deng Cai, and Partha Niyogi. Laplacian score for feature selection. In *NIPS*, pages 507–514, 2006.

[Hou *et al.*, 2017] Shifu Hou, Yanfang Ye, Yangqiu Song, and Melih Abdulhayoglu. Hindroid: An intelligent android malware detection system based on structured heterogeneous information network. In *KDD*, pages 1507–1515. ACM, 2017.

[Ji *et al.*, 2010] Ming Ji, Yizhou Sun, Marina Danilevsky, Jiawei Han, and Jing Gao. Graph regularized transductive classification on heterogeneous information networks. In *PKDD*, pages 570–586. Springer, 2010.

[Li *et al.*, 2016] Xiang Li, Ben Kao, Yudian Zheng, and Zhipeng Huang. On transductive classification in heterogeneous information networks. In *CIKM*, pages 811–820. ACM, 2016.

[Lu and Getoor, 2003] Qing Lu and Lise Getoor. Link-based classification. In *ICML*, volume 3, pages 496–503, 2003.

[Manning *et al.*, 2014] Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. The stanford corenlp natural language processing toolkit. In *ACL (Demo)*, pages 55–60, 2014.

[McLellan *et al.*, 2000] A Thomas McLellan, David C Lewis, Charles P O'brien, and Herbert D Kleber. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. *The Journal of the American Medical Association*, 284(13):1689–1695, 2000.

[Neha, 2017] Kuhar Neha. Effectiveness of social media as marketing strategy. *IJRITCC*, 5:942–944, 2017.

[NIDA, 2014] NIDA. *America's Addiction to Opioids: Heroin and Prescription Drug Abuse*, 2014. https://goo.gl/HYZAnh.

[NIDA, 2017] NIDA. *Overdose Death Rates*, 2017. https://goo.gl/aVXGu2.

[Nikfarjam *et al.*, 2015] Azadeh Nikfarjam, Abeed Sarker, Karen O'Connor, Rachel Ginn, and Graciela Gonzalez. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the American Medical Informatics Association*, 22(3):671–681, 2015.

[Saloner and Karthikeyan, 2015] Brendan Saloner and Shankar Karthikeyan. Changes in substance abuse treatment use among individuals with opioid use disorders in the united states, 2004-2013. *The Journal of the American Medical Association*, 314(14):1515–1517, 2015.

[Sun *et al.*, 2011] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S Yu, and Tianyi Wu. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *VLDB*, 4(11):992–1003, 2011.

[Ye *et al.*, 2009a] Yanfang Ye, Lifei Chen, Dingding Wang, Tao Li, Qingshan Jiang, and Min Zhao. Sbmds: an interpretable string based malware detection system using svm ensemble with bagging. *JCV*, 5(4):283–293, 2009.

[Ye *et al.*, 2009b] Yanfang Ye, Tao Li, Qingshan Jiang, Zhixue Han, and Li Wan. Intelligent file scoring system for malware detection from the gray list. In *KDD*, pages 1385–1394. ACM, 2009.

[Ye *et al.*, 2011] Yanfang Ye, Tao Li, Shenghuo Zhu, Weiwei Zhuang, Egemen Tas, Umesh Gupta, and Melih Abdulhayoglu. Combining file content and file relations for cloud based malware detection. In *SIGKDD*, pages 222–230, 2011.

[Ye *et al.*, 2017] Yanfang Ye, Tao Li, Donald Adjeroh, and S Sitharama Iyengar. A survey on malware detection using data mining techniques. *ACM Computing Surveys*, 50(3):41, 2017.

[Zhao *et al.*, 2017] Huan Zhao, Quanming Yao, Jianda Li, Yangqiu Song, and Dik Lun Lee. Meta-graph based recommendation fusion over heterogeneous information networks. In *KDD*, pages 635–644. ACM, 2017.

[Zhong and Ng, 2010] Zhi Zhong and Hwee Tou Ng. It makes sense: A wide-coverage word sense disambiguation system for free text. In *ACL (Demo)*, pages 78–83. Association for Computational Linguistics, 2010.

[Zhou *et al.*, 2003] Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf. Learning with local and global consistency. In *NIPS*, volume 16, pages 321–328, 2003.

[Zhu *et al.*, 2003] Xiaojin Zhu, Zoubin Ghahramani, and John D Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *ICML*, pages 912–919, 2003.