

# Opioid Relapse Prediction with GAN

Zhou Yang, Long H. Nguyen, and Fang Jin

*Dept. of computer science*

*Texas Tech University*

*Lubbock, USA*

{zhou.yang, long.nguyen, fang.jin}@ttu.edu

**Abstract**—Opioid addiction is a severe public health threat in the U.S, causing massive deaths and many social problems. Accurate relapse prediction is of practical importance for recovering patients since relapse prediction promotes timely relapse preventions that help patients stay clean. In this paper, we introduce a Generative Adversarial Networks (GAN) model to predict the addiction relapses based on sentiment images and social influences. Experimental results on real social media data from Reddit.com demonstrate that the GAN model delivers a better performance than comparable alternative techniques. The sentiment images generated by the model show that relapse is closely connected with two emotions ‘joy’ and ‘negative’. This work is one of the first attempts to predict relapses using massive social media (Reddit.com) data and generative adversarial nets. The proposed method, combined with knowledge of social media mining, has the potential to revolutionize the practice of opioid addiction prevention and treatment.

**Index Terms**—Opioid Addicts Detection, Opioid Relapse Prediction, Generative Adversarial Nets

## I. INTRODUCTION

Opioid addiction is a severe public health threat in the U.S, causing massive deaths and many social problems [1]. According to the latest statistics of National Institute on Drug Abuse (NIDA, 2017), more than 115 Americans die after overdosing on opioids on a daily basis, and nearly 64,000 people died of drug overdoses in the US in 2016, the most lethal year of the drug overdose epidemic (NIDA, 2017). Moreover, millions of Americans have been influenced by opioid-related problems. It is estimated that 2.1 million people suffer from substance use disorders related to prescription opioid pain relievers in the United States alone [1], where there has been a significant increase in opioid-related deaths from 2000 to 2017, and the death toll is still rising [2]. The opioid crisis has social impacts beyond the increased death toll. The status of the opioid crisis in the US is shown in Figure 1; other consequences include a rise in the number of infants born dependent on opioids [3] as well as the spread of infectious diseases such as HIV and hepatitis C [4]. Opioid epidemic has now deteriorated into a full-scale national pandemic, leading

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

ASONAM '19, August 27-30, 2019, Vancouver, BC, Canada

© 2019 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6868-1/19/08. \$15.00

<https://doi.org/10.1145/3341161.3342951>

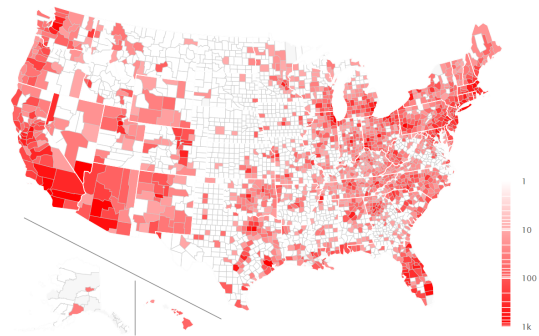


Fig. 1. County level drug overdose deaths in the U.S.A 2017 [2].

to national concern because of its negative impacts on health, social security and economics.

Effective ways to mitigate the opioid crisis are thus urgently needed. Most of the previous work in this area has tended to focus on discovering drug-related adverse events [5], for example by proposing platforms to identify prescription drug abuse, monitor medication abuse via social media, or detect opioid addicts. As yet, however, there have been few attempts at opioid addiction relapse prediction. The ability to predict relapses is critically important; as the data shown in Figure 2 indicate, since this could potentially provide substantial support for better addiction prevention and treatment.

The addiction relapse process involves three main phases, namely emotional, mental, and physical relapses [6], [7]. Relapse is a gradual process that could start days to weeks before the actual consumption of drugs. Typically, addiction relapse begins on an emotional level, which then goes on to become a physical relapse if no action is taken to support the recovering addict. These emotional states, once confirmed, can thus be regarded as an indicator for relapse prediction. An accurate relapse prediction is essential to address the opioid epidemic since a precise relapse prediction makes it possible to implement a well-timed relapse prevention plan. Moreover, if combined with medical care and mental support, relapse prediction can prevent relapse or extend the period that an addict stays clean. Even though relapse prediction could potentially be helpful for relapse prevention, however, it remains a very challenging task for two main reasons: (1) addicts tend to be very emotional as a result of the way drugs affect their brains, making their responses and behavior patterns inherently unpredictable [8]; and (2) traditional supervised methods for

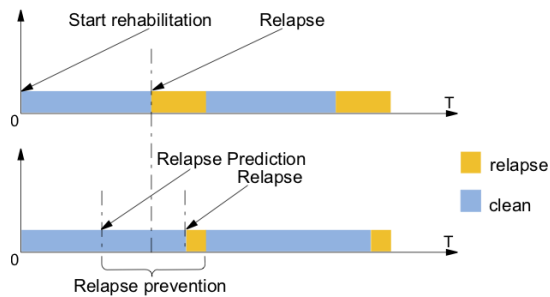


Fig. 2. Rehabilitation process. Relapse is an inevitable part of the rehabilitation process, but accurate relapse prediction combined with other supports can postpone or prevent relapse, and hence compress the relapsing period [11], [12].

relapse prediction demand a significant amount of training data if they are to achieve a reliable accuracy.

Social media is now playing an increasingly important role in addiction rehabilitation as it provides addicts with mental support during the recover process. According to a report from the National Institute on Drug Abuse (NIDA) [9], opioid addiction recovery calls for not only medical treatment (such as methadone, buprenorphine and naltrexone), but also support from families, medical professionals and communities. However, family members or friends do not always know or are unsure how to help addicts to alleviate the side effects of addiction. This situation spurs addicts to turn to social media for help, posting their problems and confusion, searching for opportunities for group therapy, and seeking answers in specific forums. Many people who have managed to get rid of opioid dependencies are willing to share their experience with others online. For instance, on Reddit, some discussion groups allow users to share their opinions or experience regarding opioid addiction and rehabilitation “*I was the same way, I only got clean cuz I didn’t wanna go back to jail tbh, and had to go cold turkey in rehab*”. Also, social media allows opioid addicts to express their joy and sadness, trust and disgust, anticipation and anger. These sentiments, along with addicts’ online behaviors, could disclose their treatment status and thus serve as indicators for relapse prediction. We therefore sought to, capture addicts’ subtle sentiments such as fear, joy, sadness or anger and then use these as target variables to predict whether an addict will relapse based on their posts on social media.

In this paper, we present the implementation of a Generative Adversarial Net (GAN) [10], and describe how we incorporated sentiment analysis and network influence analysis, to model the relapse prediction problem. If a softmax layer is applied to the output of the discriminator, the resulting GAN model can be used for relapse prediction.

We gathered appropriate social media data, especially data from groups such as “r/opiates”, “r/opiatesRecovery” and “r/Drugs” on Reddit, and built a classifier to extract data about the addicts. We then quantified the emotions represented in the comments made by the addicts and transformed these sentiment scores into sentiment images that could be fed into

the discriminator. Using the GAN framework, the generator was able to learn the emotion patterns using a convolution neural network while the discriminator inferred the relapse probabilities based on a neural network with a softmax layer. Our key contributions are:

- We designed an innovative Generative Adversarial Network to predict opioid relapse that includes a discriminator that is capable of learning the data distribution patterns of both relapsing and non-relapsing addicts.
- This work provides an entirely new approach to opioid relapse prediction. Classifiers are trained to identify two types of addicts: those who are struggling to get off drugs and those who enjoy the drugs and have no plans to rehabilitate. Using social media data, we seek to predict whether a recovering opioid addict will relapse or not.
- We propose the concept of sentiment images that adequately capture the dynamic emotion fluctuations of addicts, as well as considering the social influences between users and the rich relationships between them.
- We performed a comprehensive set of experiments using a real-world dataset that produced results that consistently outperformed those obtained using other approaches, which could potentially revolutionize the practice of opioid addiction prevention and treatment.

## II. RELATED WORK

We review three categories of related work as follows.

### a) *Biomedical Knowledge Mining on Social Network:*

As social media are profoundly and widely intertwined with people’s everyday lives, they contain more and more biomedical information about users. As a result, social media represent an increasingly important data source for discovering biomedical knowledge in applications. Bian et al. proposed the use of a Support Vector Machine to identify drug users and adverse events based on Twitter messages [5], while Fan et al. introduced an AutoDOA to automatically detect opioid addicts by building a Heterogeneous Information Network (HIN) [13].

b) *Emotional States and Addiction Relapse:* Negative emotional states, such as stress, depression, anxiety, and other emotional states are known to be related to relapse. Shiffman evaluated reports of relapse in ex-smokers, finding that most of the subjects (71%) had negative emotions preceding their relapse, with the most common mood state being anxiety, followed by anger, frustration and depression [14].

### c) *Adversarial Neural Networks for Classification:*

Generative Adversarial Nets (GANs) stand out for their performance over a wide range of applications. Since their introduction by Goodfellow et al [15], they have been used to generate images or videos, protect medical information privacy, and classify images, among others. These earlier studies on GAN laid the foundation for applying GAN to some interesting fields, including classification and prediction.

## III. PROBLEM FORMULATION

In this section, we introduce the formal definition of the problem, and present the solution. Our mainly focus here is

to build a practical and reasonable model using social media data to predict who is likely to relapse **in the next week**. The reason why we are limiting our goal to predict relapse in the next week is based on several facts. First, relapse predictions for a shorter or a longer periods makes no sense, as 99% of the recovering addicts will continue to abstain for a day since their last relapse or since beginning of the detoxification process, while more than 85% will relapse within a year [16]. Moreover, the majority of current relapse prediction research is measured in terms of weeks [17], [18], hence our focus on predicting relapse in the week ahead, utilizing it as the default setting in our experiments.

*Definition 3.1 (Drug Relapse):* Drug relapse is defined as the resumption of drug use after a period of abstinence.

A relapse is generally associated with younger age, heavy use before treatment, a history of injecting, and not following up with aftercare [19]. However, it is also frequently associated with extreme emotions and stressors [20]. Specifically, people in rehabilitation suffer from extreme emotional swings. Because of this, this paper predicts the probability of relapsing based on emotional behaviors.

*Definition 3.2 (Social Community):* Each social community of an addict consists of several social groups, which is called subreddit, denoted as  $S_j = \{s_1^j, s_2^j, s_3^j, \dots\}$ . A subreddit consists of a collection of posts, which is denoted by  $s_i^j = \{p_{i1}^j, p_{i2}^j, p_{i3}^j, \dots\}$ .

We can conclude that the influence an addict receives is  $P = \sum_{j=1}^J \sum_{i=1}^I \sum_{k=1}^K p_{ik}^j$ . For an addict  $a_j$ , we quantify it with  $\langle P_j, F_j, Y_j \rangle$ , where  $P_j$  is the social influence received,  $Y_j$  is the relapsing indicator, and  $F_j$  represents the features of the addicts. In short, we combine  $\langle P_j, F_j \rangle$  and rewrite it as  $X_j = \langle P_j, F_j \rangle$ .

*Definition 3.3 (Task):* Given a set of addicts  $A = \{a_1, a_2, a_3, \dots\}$ , and their corresponding feature variable  $X_j = \langle P_j, F_j \rangle$  derived from their social community  $S_j = \{s_1^j, s_2^j, s_3^j, \dots\}$ , a task is to learn a function  $f : X_j \rightarrow Y_j$ , such that  $Y_j \in [0, 1]$ , where  $X_j$  and  $Y_j$  are feature variable and target variable respectively.

*Definition 3.4 (Relapsing Prediction Task):* Given a set of addicts  $A = \{a_1, a_2, a_3, \dots\}$ , and their state features  $X_j = \{P_j, F_j\} \in R^m$ , find a mapping  $\Psi : R^m \rightarrow Y$ , where  $Y \in [0, 1]$ , such that the task error is minimized. Formally, we have

$$\arg \min_{\Psi} \sum_{j=1}^m \|f(X_j) - Y_j\|_2 \quad (1)$$

In the following part, we will discuss how to apply Generative Adversarial Networks to the relapsing prediction problem. Given real data with distribution  $P_r(x)$  over a feature space  $X$ , and expect to learn a distribution  $P_g(x)$  that is as close as possible to  $P_r(x)$ . The closeness between two probability distribution  $P_1$  and  $P_2$  is defined by the divergence function  $D(P_1, P_2)$ . So the optimality is defined by

$$P_g^* = \arg \inf_{P_g} D(P_g, P_r) \quad (2)$$

where  $P_g$  is the data distribution generated by generator,  $P_r$  is the data distribution of addicts and  $\inf_{P_g} D(P_g, P_r)$  is the infimum of  $D(P_g, P_r)$ .

### A. Generative Adversarial Networks

The GANs framework consists of two neural networks that are trained jointly. The first neural work is implemented as a generator, and it is initialized with some random noise  $x_0$ . The discriminator is trained using traditional supervised learning techniques. The input of the discriminator then consists of two types, real data (labeled) and fake data created by the generator. The output is scalar, simply indicating if the data is real or fake. In theory, these two opponents playing in the game are represented by two functions, each of which must be differentiable both concerning its inputs  $x$  and its parameters  $\theta$ . The discriminator is represented by a function  $D$  that takes  $x$  (a combination of real and fake data) as input and uses  $\theta^{(D)}$  as parameters. Similarly, the generator is represented by a function  $G$  that takes  $z$  as input and use  $\theta^{(G)}$  as parameters. The objective function of the generator is to minimize the cost  $J^{(G)}(\theta^{(D)}, \theta^{(G)})$  while controlling only  $\theta^{(G)}$ . Similarly, the objective function of the discriminator is to minimize the cost  $J^{(D)}(\theta^{(D)}, \theta^{(G)})$  while controlling only  $\theta^{(D)}$ . The solution to this game is local differential Nash equilibrium, which is described by a tuple  $(\theta^{(D)}, \theta^{(G)})$  [21].

---

#### Algorithm 1: Generative Adversarial Nets [15]

---

```

Initialize  $\theta_d$  for D and  $\theta_G$  for G;
for each training iteration do
  for K steps do
    Sample m examples  $x^1, x^2, \dots, x^m$  from data
    distribution  $P_{data}(x)$ ;
    Sample m noise samples  $z^1, z^2, \dots, z^m$  from the
    prior  $P_{prior}(z)$ ;
    Obtaining generated data  $\tilde{x}^1, \tilde{x}^2, \dots, \tilde{x}^m$  by using
     $\tilde{x}^i = G(z^i)$ ;
    Update discriminator parameters  $\theta_d$  to maximize
     $\tilde{V} = \frac{1}{m} \sum_{i=1}^m \log D(x^i) + \frac{1}{m} \sum_{i=1}^m \log(1 - D(\tilde{x}^i))$ 
    and  $\theta_d \leftarrow \theta_d + \eta \nabla \tilde{V}(\theta_d)$ ;
  end
  Sample another m noise samples  $z^1, z^2, \dots, z^m$  from
  the prior  $P_{prior}(z)$ ;
  Update generator parameters  $\theta_g$  to minimize
   $\tilde{V} = \frac{1}{m} \sum_{i=1}^m \log D(x^i) + \frac{1}{m} \sum_{i=1}^m \log(1 - D(z^i))$  and
   $\theta_d \leftarrow \theta_d - \eta \nabla \tilde{V}(\theta_d)$ ;
end

```

---

### B. Cost Functions

Since the discriminators of GANs are similar to a supervised binary classifier, they have the same cost function. Suppose the discriminator is approximated by a function  $D : R^n \rightarrow [0, 1]$ ,

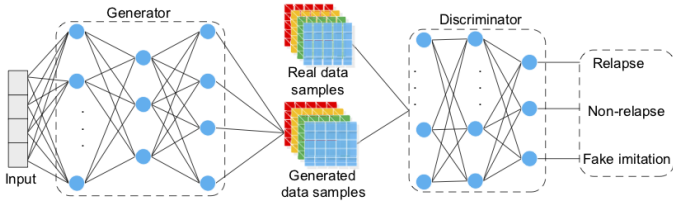


Fig. 3. Generative Adversarial Net Framework.

and let  $(x_i, y_i) \in (R^n, 0, 1)$  be the input-label pairing for training data points, the cross-entropy of the binary discriminator is defined as

$$H(G) = - \sum_{i=1}^N (y_i \log D(x_i) + (1 - y_i) \log(1 - D(x_i))) \quad (3)$$

### C. Relapsing Prediction

The relapse prediction model is a combination of GAN and a standard neural network classifier. For a standard classifier that classifies a data point  $x$  into one of  $K$  possible classes, it takes in  $x$  as input and outputs a  $K$ -dimensional vector of logits  $\{l_1, \dots, l_K\}$ , that can be turned into class probabilities by applying the softmax:  $p_{model}(y = i|x) = \frac{\exp(l_i)}{\sum_{k=1}^K \exp(l_k)}$ . This

model is then trained by minimizing the cross-entropy between the observed labels and the model predictive distribution  $p_{model}(y|x)$ . Similarly, a GAN classifier is built with standard neural classifier by simply adding samples from the GAN generator to the data set. The samples are then labeled with a new ‘generated’ class  $y = K + 1$ , and correspondingly increasing the dimension of the output of the classifier from  $K$  to  $K + 1$ . In this case,  $p_{model}(y = K + 1|x)$  indicates the probability that  $x$  is fake, corresponding to  $1 - D(x)$  in the original GAN framework. This process is given in Figure 3.

Once the generative adversarial net model has been trained, we obtain a model that can be used either to predict relapse or to generate sample data. On the one hand, the discriminator has now learned the data distribution of the relapsing population and can thus be used to predict relapses. On the other hand, since the generator has learned what real relapsing data looks like, it can produce realistic fake data by imitating the real data. This process is shown in Figure 3.

## IV. MODEL DEVELOPMENT, APPLICATION AND RESULTS

### A. Experimental Design

In this section, we utilize an experimental study using real data from Reddit.com to comprehensively evaluate the performance of our proposed new GANs model. A series of complex intermediate experiments were involved including, for example, natural language processing and image processing, and these are described in turn below. Finally, the proposed method is evaluated by comparing its performance with those of alternative methods.

The GANs model described here consists of a generator and a discriminator that are trained together. In this case,

the generator generates data about relapsing addicts, while the discriminator evaluates whether the input data it receives is from a real relapsing addict or not. The input data of the discriminator therefore consists of a combination of real data from addicts and fake data generated by the generator. Meanwhile, the generator continues to create new data that imitates real data from relapsing addicts in an attempt to fool the discriminator. The data it generates is fed into the discriminator, along with a stream of real data from actual relapsing addicts. The goal of the discriminator is to identify the data produced by the generator as fake. The discriminator receives three categories of data: data about addicts who have relapsed, data about addicts who are struggling to stay clean and data generated by the generator. The output of the discriminator is thus a vector of probabilities with three elements that indicate the probability that the input belongs to each category. The model that emerges at the end of this process can be used both to predict relapse probabilities and generate sample data.

### B. Data collection

A web crawling tool<sup>1</sup> was developed using PRAW (Python Reddit API Wrapper) to collect data from Reddit.com, a social media platform open to specific groups and sessions. The dataset for this paper came primarily from "r/Opiates", "r/OpiatesRecovery" and "r/Drugs" subreddit (see Definition 3.2) on Reddit.com in April 2018. Within each subreddit, there are a series of discussion posts (see Definition 3.2). We collected the data using keyword-based method, with the keywords being a list of opioid-related drug names and its abbreviations, slangs or jargons that are related to opioids, such as ‘oxy (oxytocin)’, ‘Xanax’, ‘dope’, etc. The dataset consists of two parts. The first consists of the subreddits, which include comments and interactions within specific posts. After collecting the subreddit data, the user IDs of the users who either create a post or make at least one comment are extracted to produce a list user IDs from the corresponding subreddit. The second part of the dataset consists of personal data, which was collected based on the user list extracted from the first part. Based on each user ID, comments from the previous 37 days was collected using keyword-based methods; and comments/posts from the first 30 days were then used to build sentiment images, which are explained in more detail in Section IV-F, and posts/comments from the final 7 days were used to label individual users as relapsed or not lapsed based on self-reporting. This involved manually reading the posts/comments from the last 7 days; if a user posts/comments that he/she has relapsed, then they are labeled as a relapsed addict. This labeling process was done by several health science researchers, and labeling agreement among the researchers was required to reduce labeling errors that might otherwise have resulted from different labeling metrics. For the addiction classifier, we followed DSM-5 criteria [22].

<sup>1</sup>[https://github.com/yangzhou95/opioid\\_on\\_reddit/tree/master/code/reddit-collector](https://github.com/yangzhou95/opioid_on_reddit/tree/master/code/reddit-collector)





to form the sentiment image, a 10\*10 matrix with each row representing the average sentiment score of three comments. The sentiment images of the addicts are then fed into the discriminator as real data from relapsing addicts.

Social influences or interactions are also considered when building the sentiment images. Just like a color image that has 3 RGB (Red, Green and Blue) channels, the sentiment image we utilized here has two channels. The first of these channels is a 10\*10 matrix, which is built by stacking up ten sentiment vectors. The second channel is constructed by taking the social network influences into consideration. Specifically, if an addict starts a discussion by asking a question or describing his or her current status, other users may make comments and then the discussion creator may reply. By exchanging opinions in this way, people who exchange opinions can influence each other. For the purposes of this research, we defined social network influence as follows:

$$influence = \begin{cases} \frac{1}{n} \sum_{i=1}^n S_i & \text{if } n \text{ comments is replied} \\ 0 & \text{no reply} \end{cases} \quad (8)$$

If no one replies, the discussion creator receives a 0 influence score. If dozens of users reply and the creator replies, then the creator receives an average of all the comments. Figure 6 shows the social interaction for users from ‘r/Opiates’, ‘OpiatesRecovery’ and ‘Drugs’. By analyzing the group interaction, we find that in most of the posts, the total number of comments are smaller than 10. Those participating in ‘r/Drugs’ discuss illegal drug usage more, while people in ‘r/OpiatesRecovery’ and ‘Opiates’ tend to discuss how to get rid of drugs.

The 2-channel sentiment images were then fed into the generative adversarial network to train the model.

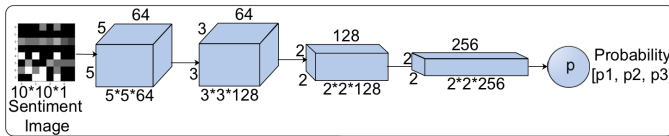


Fig. 8. Input for the Discriminator and tensor flows

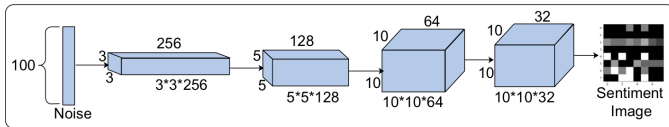


Fig. 9. Input for the Generator and tensor flows

### G. Relapse Prediction

The relapse prediction model was modified from the discriminator of a standard generative adversarial net by applying a softmax layer to the output of a discriminator [24]. We trained the model for 7,000 epochs with 1,000 labeled sentiment images (consisting of 330 addicts who have managed to stay clean for a while, and 670 addicts who have relapsed).

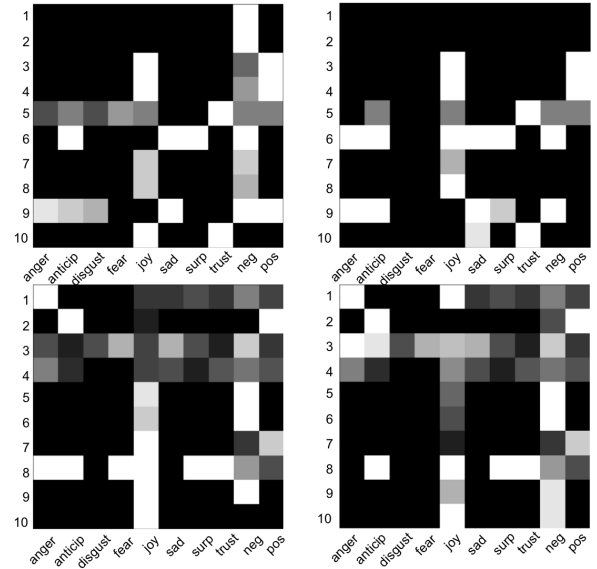


Fig. 10. These sentiment images are from four addicts produced by GAN when model converged. It can be observed that the emotion ‘joy’ and emotion ‘negative’ are closely connected to relapse.

Since no labeled dataset is available online, we labeled the data of 1,000 users and transformed the data into images with two layers (sentiment layer and social influence layer, Channel=2). For each mini-batch, we randomly selected 128 images. For optimization we used Adam [25] with a learning rate=0.0001. Our implementation of the GANs model is based on Tensorflow [26]. The data flows inside the model can be visualized as shown in Figure 8 and Figure 9. The dimensions of the tensors are calculated as the tensor dimension calculation in convolution neural networks [27]. Figure 8 shows the input of a discriminator as a mixture of sentiment images and data from the generator, with the output being the probability that an image belongs. In Figure 9, the input for the generator is a vector of random noise sampled from a uniform distribution, after which the data goes through several convolution layers. Finally, the generator outputs an image.

The prediction results are presented in Table II and the new model’s performance is evaluated by comparing the results with those obtained using alternative methods. The generative adversarial net achieved an accuracy of 0.8390 and an F-score of 0.8547, clearly outperforming other alternative methods such as Logistic Regression, SVM and KNN. Interestingly, the sentiment images generated by the generator show that relapse is closely connected with the emotion ‘joy’ and the emotion ‘negative’. For instance, in Figure 10, which is a set of images generated by the generator, the emotions ‘joy’ and ‘negative’ appear consistently in the sentiment image with non-zero values (white spaces in the figure), indicating the corresponding comments express complex feelings. The rows in the sentiment image with multiple emotions (white spaces) reveal that the user has made some lengthy comments that contain a mixture of different emotions, while the rows with fewer white spaces, signify short comments such as ‘so sad’

and ‘that’s bad’.

TABLE II  
PREDICTION PERFORMANCE AND EVALUATION

Relapse prediction with different sizes of training data						
		90%	80%	70%	60%	50%
LR	ACC	0.715	0.692	0.652	0.632	0.602
	F1	0.7775	0.7582	0.7268	0.7065	0.6801
SVM	ACC	0.821	0.798	0.768	0.738	0.718
	F1	0.850	0.841	0.816	0.791	0.775
KNN	ACC	0.712	0.798	0.659	0.629	0.599
	F1	0.774	0.841	0.728	0.702	0.6753
GAN	ACC	<b>0.839</b>	<b>0.819</b>	<b>0.784</b>	<b>0.749</b>	<b>0.714</b>
	F1	<b>0.855</b>	<b>0.849</b>	<b>0.831</b>	<b>0.801</b>	<b>0.772</b>

## V. CONCLUSION

In this paper, we present a new application where generative adversarial nets are utilized to predict whether a recovering opioid addict will relapse or not within the next week. This paper is one of the first attempts to predict opioid addict relapses using massive social media data from Reddit.com. By combining social media data with generative adversarial networks, our model makes predictions based on addicts’ previous records and current sentiment status. Our experimental results demonstrate that the new GANs model consistently outperforms traditional prediction techniques and could thus be incorporated into a strategy for supporting practitioners working on opioid relapse prevention and addiction treatment. Unfortunately, the limitations of this study are clear: the size of manually labeled training data used in the experimental is not large enough to draw strong and consolidated conclusions. We are aware of the existence of this limitation but leave it as it was because processing and manually labeling a large amount of training data is very costly. This limitation can be fixed by investing more time to collect, process and manually label more training data. We leave the limitation for future research.

## VI. ACKNOWLEDGEMENT

This work was supported by the U.S. National Science Foundation under Grant CNS-1737634.

## REFERENCES

- [1] M. W. Holly Hedegaard and A. M. Minino, “Drug overdose deaths in the united states,” <https://www.cdc.gov/nchs/products/databriefs/db294.htm>, December-2017, online, accessed 25-March-2018.
- [2] NIDA, “Overdose death rates,” <https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates>, September-2017, online, accessed 25-March-2018.
- [3] S. W. e. Patrick, “Increasing incidence and geographic distribution of neonatal abstinence syndrome: United states 2009 to 2012,” *Journal of Perinatology*, vol. 35, no. 8, p. 650, 2015.
- [4] C. Conrad, H. M. Bradley, D. Broz, S. Buddha, E. L. Chapman, R. R. Galang, D. Hillman, J. Hon, K. W. Hoover, M. R. Patel *et al.*, “Community outbreak of hiv infection linked to injection drug use of oxymorphone-indiana, 2015.” *MMWR. Morbidity and mortality weekly report*, vol. 64, no. 16, pp. 443–444, 2015.
- [5] J. e. Bian, “Towards large-scale twitter mining for drug-related adverse events,” in *Proc. of the 2012 international workshop on Smart health and wellbeing*. ACM, 2012, pp. 25–32.
- [6] Theraleighhouse, “3 stages of drug addiction relapse: Emotional, mental & physical,” <https://www.theraleighhouse.com/addiction-blog/resources-myself/3-stages-addiction-relapse-emotional-mental-physical>, 7-September-2017, online, accessed 25-March-2018.
- [7] K. D. e. Brownell, “Understanding and preventing relapse,” 1986.
- [8] G. F. Koob, “Dynamics of neuronal circuits in addiction: reward, antireward, and emotional memory,” *Pharmacopsychiatry*, vol. 42, no. Suppl 1, p. S32, 2009.
- [9] Theraleighhouse, “Seeking drug abuse treatment: Know what to ask,” <https://www.drugabuse.gov/publications/seeking-drug-abuse-treatment-know-what-to-ask/introduction>, 7-Sept-2017, online, accessed 25-March-2018.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [11] G. A. Marlatt and D. M. Donovan, *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. Guilford press, 2005.
- [12] J. R. e. Geddes, “Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review,” *The Lancet*, vol. 361, no. 9358, pp. 653–661, 2003.
- [13] Y. e. Fan, “Social media for opioid addiction epidemiology: Automatic detection of opioid addicts from twitter and case studies,” in *Proc. of the 2017 ACM CIKM*. ACM, 2017, pp. 1259–1267.
- [14] S. Shiffman, “Relapse following smoking cessation: a situational analysis.” *Journal of consulting and clinical psychology*, vol. 50, no. 1, p. 71, 1982.
- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems 27*, 2014, pp. 2672–2680.
- [16] R. Sinha, “New findings on biological factors predicting addiction relapse vulnerability,” *Current psychiatry reports*, vol. 13, no. 5, p. 398, 2011.
- [17] L. O. Bauer, “Predicting relapse to alcohol and drug abuse via quantitative electroencephalography,” *Neuropsychopharmacology*, vol. 25, no. 3, pp. 332–340, 2001.
- [18] R. L. e. Ashare, “The first seven days of a quit attempt predicts relapse: Validation of a measure for screening medications for nicotine dependence.” *Journal of addiction medicine*, vol. 7, no. 4, p. 249, 2013.
- [19] B. P. Smyth, J. Barry, E. Keenan, and K. Ducray, “Lapse and relapse following inpatient treatment of opiate dependence.” *Irish medical journal*, 2010.
- [20] J. D. Ford and E. Russo, “Trauma-focused, present-centered, emotional self-regulation approach to integrated treatment for posttraumatic stress and addiction: trauma adaptive recovery group education and therapy (target).” *American journal of psychotherapy*, vol. 60, no. 4, 2006.
- [21] L. J. e. Ratliff, “Characterization and computation of local nash equilibria in continuous games,” in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*. IEEE, 2013, pp. 917–924.
- [22] M. J. e. Tarrahi, “Latent class analysis of dsm-5 criteria for opioid use disorders: results from the iranian national survey on mental health,” *European addiction research*, vol. 21, no. 3, pp. 144–152, 2015.
- [23] S. M. Mohammad and P. D. Turney, “Crowdsourcing a word-emotion association lexicon,” vol. 29, no. 3, pp. 436–465, 2013.
- [24] T. e. Salimans, “Improved techniques for training gans,” in *Advances in Neural Information Processing Systems*, 2016, pp. 2234–2242.
- [25] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [26] A. e. Martín, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>
- [27] F.-F. Li, “Convolutional neural networks (cnns / convnets),” <http://cs231n.github.io/convolutional-networks/>, 2018.