

RESEARCH

Open Access



Semi-supervised learning for detecting human trafficking

Hamidreza Alvari^{1*} , Paulo Shakarian¹ and J. E. Kelly Snyder²

Abstract

Human trafficking is one of the most atrocious crimes and among the challenging problems facing law enforcement which demands attention of global magnitude. In this study, we leverage textual data from the website “Backpage”—used for classified advertisement—to discern potential patterns of human trafficking activities which manifest online and identify advertisements of high interest to law enforcement. Due to the lack of ground truth, we rely on a human analyst from law enforcement, for hand-labeling a small portion of the crawled data. We extend the existing Laplacian SVM and present $S^3VM - R$, by adding a regularization term to exploit exogenous information embedded in our feature space in favor of the task at hand. We train the proposed method using labeled and unlabeled data and evaluate it on a fraction of the unlabeled data, herein referred to as *unseen* data, with our expert’s further verification. Results from comparisons between our method and other semi-supervised and supervised approaches on the labeled data demonstrate that our learner is effective in identifying advertisements of high interest to law enforcement.

Keywords: Human trafficking, Backpage, Semi-supervised support vector machines, Laplacian support vector machines

Background

According to the United Nation [1], human trafficking is defined as the modern slavery or the trade of humans mostly for the purpose of sexual exploitation and forced labor, via different improper ways including force, fraud and deception. The United States’ Trafficking Victim Protection Act of 2000 (TVPA 2000) [2] was the first US legislation passed against human trafficking. Human trafficking has ever since received increased national and societal concern [3] but still demands persistent fight against from all over the globe. No country is immune and the problem is rapidly growing with little to no law enforcement addressing the issue. This problem is amongst the challenging ones facing law enforcement as it is difficult to identify victims and counter traffickers.

Before the advent of the Internet, human traffickers were under risks of being arrested by law enforcement while advertising their victims on streets [4]. However, move to the Internet has made it easier and less

dangerous for sex sellers [5] as they no longer needed to advertise on the streets. There are now a plethora of websites that host and provide sexual services under categories of escort, adult entertainment, massage services, etc., which help sex sellers and buyers maintain their anonymity. Although some services such as the Craigslist’s adult section and myredbook.com were shut down recently, there are still many websites such as the Backpage.com that provide such services and many new are frequently created. Traffickers even use dating and social networking websites such as the Twitter, Facebook, Instagram and Tinder to reach out to sex buyers and their followers. Although the Internet has presented new trafficking related challenges for law enforcement, it has also provided readily and publicly available rich source of information which could be gleaned from online sex advertisements for fighting this crime [6]. However, the problem is we lack the ground truth and obtaining the labels through hand-labeling is indeed tedious and expensive even for a small subset of data—this is the point where the semi-supervised setting comes in handy.

Despite considerable attention which has been devoted to studying supervised, unsupervised and

*Correspondence: halvari@asu.edu

¹ Arizona State University, Tempe, Arizona, USA

Full list of author information is available at the end of the article

semi-supervised learning settings via different applications [7–13], semi-supervised learning, i.e., learning from labeled and unlabeled examples, is still one of the most interesting yet challenging problems in the machine learning community [14]. The idea is simple though—we shall have an approach that makes a better use of unlabeled data to boost performance. This is pretty close to the most natural learning that occurs in the world. For the most part, we as humans are exposed only to a small number of labeled instances; yet we successfully generalize well by effective utilization of a large amount of unlabeled data. This motivates us to use unlabeled samples to improve recognition performance while developing classifiers.

In this article, expanding on our previous work [15], we use the data crawled from the adult entertainment section of the website Backpage.com and extend the existing Laplacian SVM framework [14] to detect escort advertisements of high interest to law enforcement. Here, we merely focus on the online advertisements although the Internet has triggered many other activities including attracting the victims, communicating with customers and rating the escort services. We thus highlight several contributions of the current research as follows.

1. Based on the literature, we created different groups of features that capture the characteristics of potential human trafficking activities. The less likely human trafficking related posts were then filtered out using these features. We also conducted a feature importance analysis to demonstrate how these features contribute to the proposed learner.
2. We extended the Laplacian SVM [14] and proposed the semi-supervised support vector machine learning algorithm, $S^3VM - R$. In particular, we incorporated additional information of our feature space as a regularization term into the standard optimization formulation with regard to the Laplacian SVM. We also used geometry of the underlying data as an intrinsic regularization term in Laplacian SVM.
3. We trained our model on both of the *labeled* and *unlabeled* data and sent back the identified human trafficking related advertisements to an expert from law enforcement for further verification. We then validated our approach on a small subset of the unlabeled data (i.e. *unseen* data) with further verification of the expert.
4. We performed comparisons between our approach and several semi-supervised and supervised baselines on both of the labeled and unseen data (so-called *blind* evaluation).
5. We demonstrated the effect of varying different hyperparameters used in our learner on its performance.

The rest of the paper is organized as follows. In section “[Related work](#)”, we review the prior studies on human trafficking. Section “[Data preparation](#)” covers our data preparation, feature engineering, unsupervised filtering and expert assisted labeling. We detail our semi-supervised learning approach in section “[Semi-supervised learning framework](#)” by deriving the required equations. Section “[Experimental study](#)” provides in-depth explanation of our experiments. Section “[Conclusion](#)” concludes the paper by providing future research directions.

Related work

Recently, several studies have examined the role of the Internet and related technology in facilitating human trafficking [16–18]. For example, the work of [16] studied how closely sex trafficking is intertwined with new technologies. According to [17], sexual exploitation of women and children is a global human right crisis that is being escalated by the use of new technologies. Researchers have studied relationships between new technologies and human trafficking and advantages of the Internet for sex traffickers. For instance, findings from a group of experts from the Council of Europe demonstrated that the Internet and sex industry are closely interlinked and volume and content of the material on the Internet promoting human trafficking are unprecedented [18].

One of the earliest works which leveraged data mining techniques for online human trafficking was [18], wherein the authors conducted data analysis on the adult section of the website Backpage.com. Their findings confirmed that female escort post frequency would increase in Dallas, Texas, leading up to the Super Bowl 2011 event. In a similar attempt, other studies [19, 20] have investigated impact of large public events such as the Super Bowl on sex trafficking by exploring advertisement volume, trends and movement of advertisements along with the scope and volume of demand associated with such events. The work of [19] for instance, concluded that large events such as the Super Bowl which attract significant amount of concentration of people in a relatively short period of time and in a confined urban area, could be a desirable location for sex traffickers to bring their victims for commercial sexual exploitation. Similarly, the data-driven approach of [20] showed that in some but not all events, one can see a correlation between occurrence of the event and statistically significant evidence of an influx of sex trafficking activity. Also, certain studies [21] have tried to build large distributed systems to store and

process available online human trafficking data in order to perform entity resolution and create ontological relations between entities.

Beyond these works, the work of [22] studied the problem of isolating sources of human trafficking from online advertisements with a pairwise entity resolution approach. Specifically, they used phone number as a strong feature and trained a classifier to predict if two ads are from the same source. This classifier was then used to perform entity resolution using a heuristically learned value for the score of classifier. Another work of [6] used Backpage.com data and extracted most likely human trafficking spatio-temporal patterns with the help of law enforcement. Note that unlike our method, this work did not employ any machine learning methodologies for automatically identifying human trafficking related advertisements. The work of [23] also deployed machine learning for the advertisement classification problem, by training a supervised learning classifier on labeled data (based on phone numbers of known traffickers) provided by a victim advocacy group. We note that while phone numbers can provide a very precise set of positive labeled data, there are clearly many posts with previously unseen phone numbers.

In contrast, we do not solely rely on phone numbers for labeling our data. Instead, our expert analyze each post's content to identify whether it is human trafficking related or not. To do so, we first filter out most likely advertisements using several feature groups and pass a small sample to the expert for hand-labeling. Then, we train our semi-supervised learner on both of the labeled and unlabeled data which in turn lets us evaluate our approach on new coming (unseen) data later. We note that our semi-supervised approach can also be used as a complementary method to procedures such as those described in [23] as we can significantly expand the training set for use with supervised learning.

Finally, note that our current research is different from our previous work [15] and we list the key nuances here:

- In this study we experiment with a much larger dataset. To obtain such dataset, we use the same raw data from [15], but this time with slight modifications of the thresholds that were used for filtering out less likely human trafficking related advertisements.
- As opposed to our previous research which deployed only one feature space, in this work, two feature spaces that have complementary roles to each other are used.
- In this paper we present a new framework based on the existing Laplacian SVM [14], by adding a regularization term to the standard optimization problem and solving the new optimization equation derived

from there. In contrast, [15] utilized the off-the-shelf graph based semi-supervised learner, LabelSpreading method [24], without any further manipulation of the original approach.

- Unlike [15] in which we did not compare our method with other approaches, this work compares our proposed framework against other semi-supervised and supervised learners. Also unlike our previous work in which only *one* group of human trafficking related advertisements were passed to *two* experts for validation, here in order to reduce the inconsistency, *two* control groups of advertisements—those of interest to law enforcement and those of not—are sent to only *one* expert for verification.

Data preparation

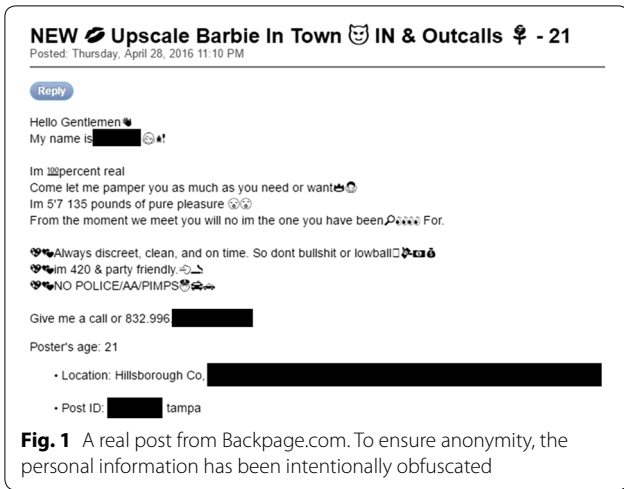
We collected about 20K publicly available listings from the US posted on Backpage.com in March, 2016. Each post includes a title, description, time stamp, poster's age, poster's ID, location, image, and sometimes video and audio. The description usually lists the attributes of the individual(s) and contact phone numbers. In this work, we only focus on the textual component of the data. This free-text data required significant cleaning due to a variety of issues common to textual analytics (i.e. misspellings, format of phone numbers, etc.). We also acknowledge that the information in data could be intentionally inaccurate, such as poster's name, age and even physical appearance (e.g. bra cup size, weight). Figure 1 shows an actual post from Backpage.com. To illustrate geographic diversity of the listings, we use the Tableau¹ software to visualize choropleth map of phone frequency with respect to the different states in Fig. 2, wherein darker colors mean higher frequencies.

Next, we will explain most important characteristics of potential human trafficking advertisements which are captured by our feature groups.

Feature engineering

Though many advertisements on Backpage.com are posted by posters selling their own services without coercion and intervention of traffickers, some do exhibit many common trafficking triggers. For example, in contrast to Fig. 1, Fig. 3 shows an advertisement that could be an evidence of human trafficking. This advertisement indicates several potential properties of human trafficking, including advertising for multiple escorts with the first individual coming from Asia and very young. In what follows, such common properties of human trafficking related advertisements are discussed in more detail.

¹ <https://www.tableau.com/>.



Inspired by the literature, we define and extract 6 groups of features from advertisements (see Table 1). These features could be amongst the strong indicators of human trafficking. Let us now briefly describe each group of features used in our work. Note each feature listed here is ultimately treated as a *binary* variable.

Advertisement language pattern

The first group consists of different language related features. For the first and second features, we identify posts which have third person language (more likely to

be written by someone other than the escort) and posts which contain first person plural pronouns such as ‘we’ and ‘our’ (more likely to be an organization) [6].

To ensure their anonymity, traffickers would deploy techniques to generate diverse information and hence make their posts look more complicated. They usually do this to avoid being identified by either human analysts or automated programs. Thus, to obtain the third feature we take an approach from complexity theory, namely *Kolmogorov complexity*, which is defined as length of shortest program to reproduce a string of characters on a universal machine such as the Turing Machine [25]. Since the Kolmogorov complexity is not computable, we approximate the complexity of an advertisement content by first removing stop words and then computing entropy of the content [25]. To illustrate this, let X denote the content and x_i be a given word in the content. We use the following equation [31] to calculate the entropy of the content and thus approximate the Kolmogorov complexity of X :

$$K(X) \approx - \sum_{i=1}^n P(x_i) \log_2 P(x_i) \tag{1}$$

We expect higher values of the entropy correspond to human trafficking. Finally, we discretize the result by using the threshold of 4 which was found empirically in our experiments.

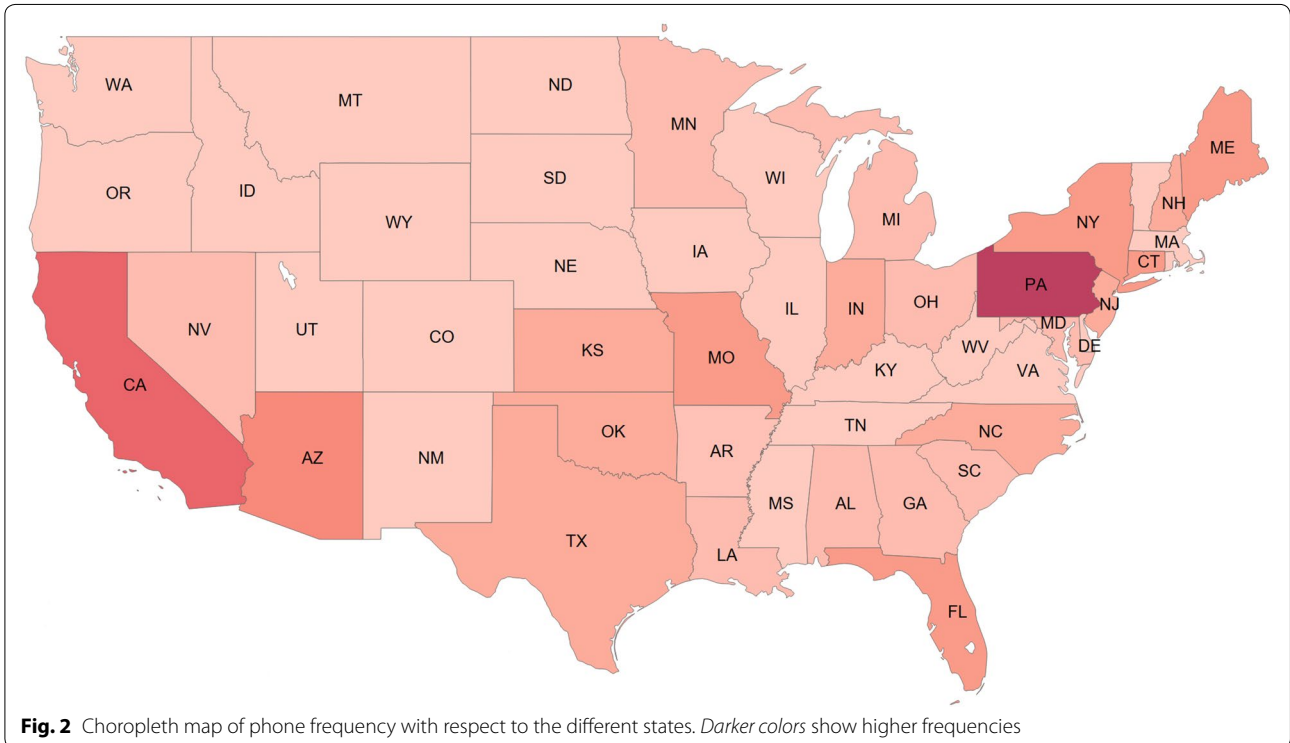


Fig. 2 Choropleth map of phone frequency with respect to the different states. Darker colors show higher frequencies

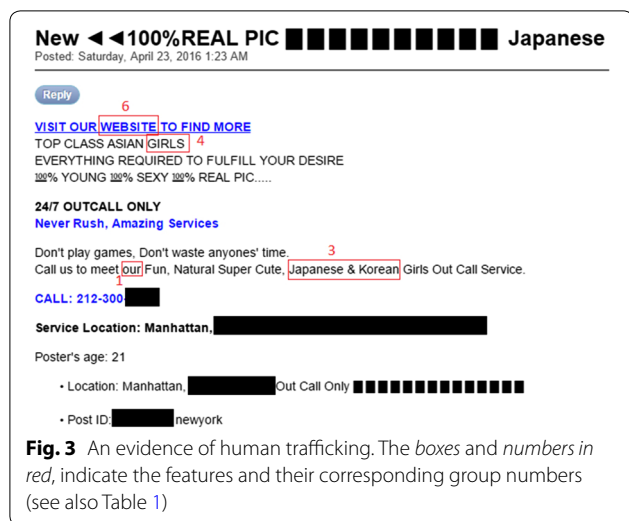


Table 1 Different features and their corresponding groups

No.	Feature group	References
1	Advertisement language pattern Third person language First person plural pronouns Kolmogorov complexity <i>n</i> -grams (1) <i>n</i> -grams (2) <i>n</i> -grams (3)	[6, 25, 26]
2	Words and phrases of interest	[27–29]
3	Countries of interest	[3]
4	Multiple victims advertised	[6]
5	Victim weight	[2, 30]
6	Reference to website or spa massage therapy Reference to a website Reference to a spa massage therapy	[6]

For the next features, we use word-level *n*-grams to find common language patterns of advertisements. This particular choice is because of the fact that character-level *n*-grams have already shown to be useful in detecting unwanted content for spam detection [26]. We set $n = 4$ and use the range of (4,4) to compute normalized *n*-grams (using TF-IDF) of each advertisement content. We ultimately create a matrix whose rows and columns correspond to the advertisements contents and their associated 4-grams, respectively. We rank all elements of this matrix in a descending order and pick the top 3 ones. Finally for each advertisement content, 3 elements with the column numbers associated with the top elements are chosen. This way, 3 more features will be added to our feature set. Overall, we have 6 features related to the language of the advertisement.

Words and phrases of interest

Despite the fact that advertisements on Backpage.com do not directly mention sex with children, customers who prefer children know to look for words and phrases such as “sweet, candy, fresh, new in town, new to the game” [27–29]. We thus investigate within the posts to see if they contain such words as they could be highly related with human trafficking in general.

Countries of interest

We identify if the individual being escorted is coming from other countries such as those in Southeast Asia (especially from China, Vietnam, Korea and Thailand, as we observed in our data) [3].

Multiple victims advertised

Some advertisements advertise for multiple women at the same time. We consider the presence of more than one victim as a potential evidence of organized human trafficking [6].

Victim weight

We take into account the weight of the individual being escorted as a feature (if it is available). This information is particularly useful assuming that for the most part, lower body weights (under 115 lbs) correlate with smaller and underage girls [2, 30] and thereby human trafficking.

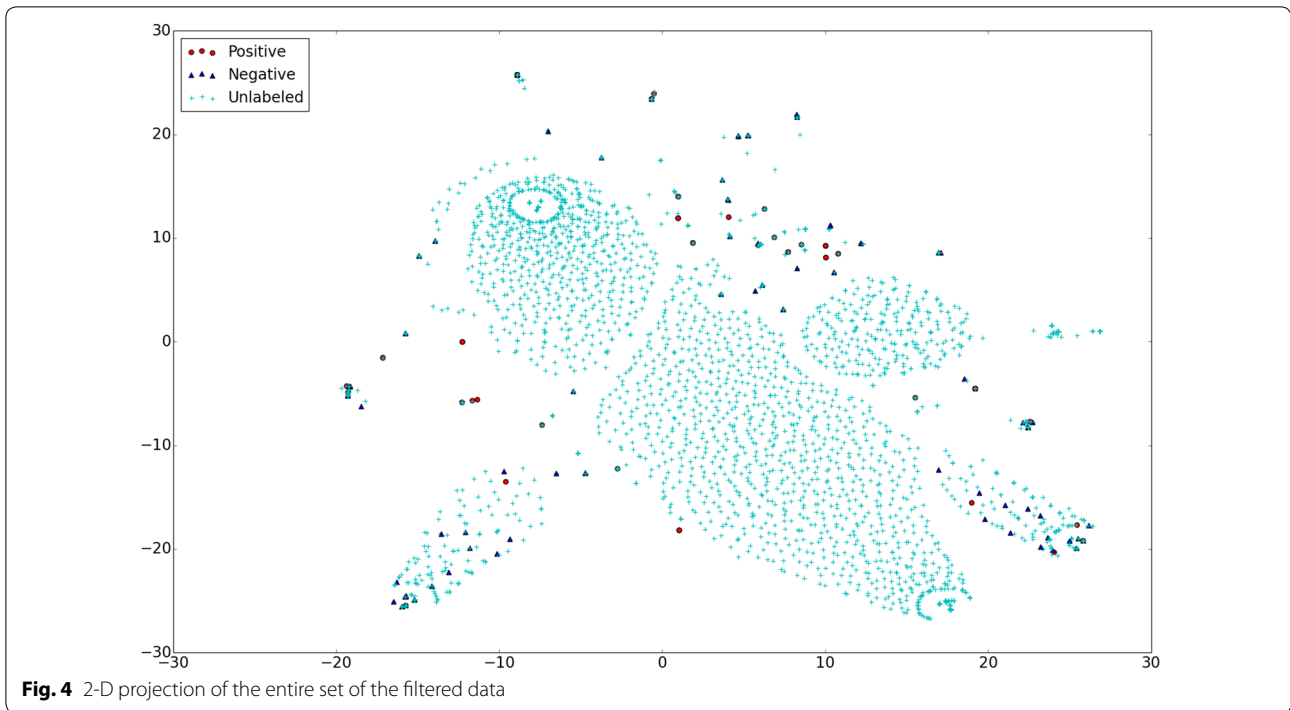
Reference to website or spa massage therapy

The presence of a link in the advertisement either referencing to an outside website (especially infamous ones) or spa massage therapy could be an indicator of more elaborate organization [6]. In particular, in case of spa therapy, we observed many advertisements interrelated with advertising for young Asian girls and their erotic massage abilities. Therefore, the last group of features has two binary features for presence of any website and spa.

Finally, in order to extract all of the above features, we first clean the original data and conduct preprocessing. By applying these features, we draw a random sample of 3543 instances out of our dataset for further analysis to see if they are evidences of human trafficking—this is described in the next section.

Unsupervised filtering

Having detailed our feature set, we now construct a feature vector for each instance by creating a vector of 12 binary features that correspond to the important characteristics of human trafficking. Hereafter, we refer to this feature space, as our *first* feature space and denote it with \mathcal{F}_1 . As mentioned earlier, we draw 3543 instances from our raw data by filtering out those that do not possess any of the binary features. We will refer to this as our *filtered*



dataset. For the sake of visualization, a 2-D projection (using the t-SNE transformation [32]) of the filtered dataset is depicted in Fig. 4. The purpose of this figure is to demonstrate how hard it is for basic clustering techniques such as the K-means, to correctly assign labels to unlabeled instances using only few existing labeled ones.

Now, we shall define our *second* feature space, namely \mathcal{F}_2 , which will be used to compute geometry of the underlying data. Note that our proposed framework will utilize both of the feature spaces in the form of regularization terms, to detect advertisements of high interest to law enforcement. After conducting standard preprocessing techniques on the filtered dataset, we build \mathcal{F}_2 by transforming the filtered data into a 3543×3543 matrix of TF-IDF similarity features. Each entry in this matrix simply shows the similarity between a pair of advertisements in our filtered dataset.

Note that since we lack the ground truth, we would rely on a human analyst (expert) for labeling the listings as either ‘of interest’ or ‘of not interest’ to law enforcement. In the next section, we select a smaller yet finer grain subset of this data to be sent to the expert. This alleviates the burden of the tedious work of hand-labeling.

Expert assisted labeling

We first obtain a sample of 200 listings from the filtered dataset. This set of listings was labeled by our expert from law enforcement who is specialized in this type of crime. From this subset, the law enforcement professional

identified 70 instances to be of interest to law enforcement and the rest to be not human trafficking related. However, we are still left with a large amount of the unlabeled examples (3343 instances) in our dataset. The ratio of the labeled to unlabeled instances in our dataset is very small (about 0.06). The statistics of our dataset is summarized in Table 2.

Semi-supervised learning framework

Here, we first introduce some preliminary notations necessary for the rest of the discussion and then outline our proposed semi-supervised approach, $S^3VM - R$, for detecting online human trafficking. Note as said earlier, our framework is an extension to the existing Laplacian SVM [14]. In particular, we incorporated another regularization term into the standard Laplacian SVM to leverage the additional information of our first feature space and then solved the associated optimization problem. Consequently, similar notation is adopted throughout the

Table 2 Description of the dataset

Name	Value	
Raw	20,822	
Filtered	3543	
Unlabeled	3343	
Labeled	Positive	Negative
	70	130

following section. Furthermore, we shall once again note that our current research does not utilize any off-the-shelf graph based semi-supervised learner in contrast to our previous research [15].

Technical preliminaries

We assume a set of l labeled pairs $\{(x_i, y_i)\}_{i=1}^l$ and an unlabeled set of u instances $\{x_{l+i}\}_{i=1}^u$, where $x_i \in \mathbb{R}^n$ and $y_i \in \{+1, -1\}$. Recall for the standard soft-margin support vector machine, the following optimization problem is solved:

$$\min_{f_\theta \in \mathcal{H}_k} \gamma \|f_\theta\|_k^2 + C_l \sum_{i=1}^l H_1(y_i f_\theta(x_i)) \tag{2}$$

In the above equation, $f_\theta(\cdot)$ is a decision function of the form $f_\theta(\cdot) = w \cdot \Phi(\cdot) + b$ where $\theta = (w, b)$ are the parameters of the model, and $\Phi(\cdot)$ is the feature map which is usually implemented using the kernel trick [33]. Also, the function $H_1(\cdot) = \max(0, 1 - \cdot)$ is the Hinge Loss function.

The classical Representer theorem [34] suggests that solution to the optimization problem exists in a Hilbert space \mathcal{H}_k and is of the following form:

$$f_\theta^*(x) = \sum_{i=1}^l \alpha_i^* \mathbf{K}(x, x_i) \tag{3}$$

where \mathbf{K} is the $l \times l$ Gram matrix over labeled samples. Equivalently, the above problem can be written as:

$$\min_{w, b, \epsilon} \frac{1}{2} \|w\|_2^2 + C_l \sum_{i=1}^l \epsilon_i \tag{4}$$

$$\begin{aligned} \text{s.t. } & y_i(w \cdot \Phi(x_i) + b) \geq 1 - \epsilon_i, \quad i = 1, \dots, l \\ & \epsilon_i \geq 0, \quad i = 1, \dots, l \end{aligned} \tag{5}$$

Next, we will use the above optimization equation as our basis to derive the formulations for our proposed semi-supervised learner.

The proposed method

The basic assumption behind semi-supervised learning methods is to leverage unlabeled instances in order to restructure hypotheses during the learning process. In this paper, exogenous information extracted from both of our feature spaces is further exploited to make a better use of the unlabeled examples. To do so, we first introduce matrix \mathbf{F} in \mathcal{F}_1 and over both of the labeled and unlabeled samples with \mathbf{F}_{ij} defined as follows:

$$\mathbf{F}_{ij} = \frac{1}{n_f} (\Phi(x_i) \cdot \Phi(x_j)) \tag{6}$$

where n_f is the number of features in \mathcal{F}_1 (here, $n_f = 12$). We force the instances x_i and x_j in our dataset to have same label if they both possess same features. To account for this, a regularization term is added to the standard equation and the following optimization is solved:

$$\min_{f_\theta \in \mathcal{H}_k} \frac{1}{2} \sum_{i=1}^l \mathbf{F}_{ij} \|f_\theta(x_i) - f_\theta(x_j)\|_2^2 = \mathbf{f}_\theta^T \mathcal{L}^T \mathbf{f}_\theta \tag{7}$$

where $\mathbf{f} = [f(x_1), \dots, f(x_{l+u})]^T$ and \mathcal{L} is the Laplacian matrix based on \mathbf{F} given by $\mathcal{L} = \mathbf{D} - \mathbf{F}$, and $\mathbf{D}_{ii} = \sum_{j=1}^{l+u} \mathbf{F}_{ij}$. The intuition here is that any two instances which are composed of same features are more likely to have same labels than others. Next, by solving a similar optimization problem, we are able to capture data geometry in \mathcal{F}_2 as $\mathbf{f}_\theta^T \mathcal{L}'^T \mathbf{f}_\theta$ (also referred to as the intrinsic smoothness penalty term [14]). Here, \mathcal{L}' is the Laplacian of matrix \mathbf{A} associated with the data adjacency graph \mathbf{G} in \mathcal{F}_2 .

We construct \mathbf{G} with $(l + u)$ nodes in \mathcal{F}_2 , and by adding an edge between each pair of nodes $\langle i, j \rangle$, if the edge weight W_{ij} exceeds a given threshold. For computing the edge weights, we use the heat kernel [35] as a function of the Euclidean distance between two samples in \mathcal{F}_2 , hence we set $W_{ij} = \exp^{-\|x_i - x_j\|^2 / 4t}$.

Following the notations used in [14] and by including our regularization term as well as the intrinsic smoothness penalty term, we would extend the standard equation by solving the following optimization:

$$\min_{f_\theta \in \mathcal{H}_k} \gamma \|f_\theta\|_k^2 + C_l \sum_{i=1}^l H_1(y_i f_\theta(x_i)) + C_r \mathbf{f}_\theta^T \mathcal{L} \mathbf{f}_\theta + C_s \mathbf{f}_\theta^T \mathcal{L}' \mathbf{f}_\theta \tag{8}$$

Note one typical value for the smoothness penalty coefficient C_s is $\frac{\gamma_l}{(l+u)^2}$, where $\frac{1}{(l+u)^2}$ is a natural scale factor for empirical estimate of the Laplace operator and γ_l is a regularization term [14]. Again, solution in \mathcal{H}_k would be in the following form:

$$f_\theta^*(x) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(x, x_i) \tag{9}$$

Here \mathbf{K} is the $(l + u) \times (l + u)$ Gram matrix over all samples. The Eq. 8 could be then written as follows:

$$\begin{aligned} \min_{\alpha, b, \epsilon} & \frac{1}{2} \alpha^T \mathbf{K} \alpha + C_l \sum_{i=1}^l \epsilon_i + \frac{C_r}{2} \alpha^T \mathbf{K} \mathcal{L} \mathbf{K} \alpha \\ & + \frac{\gamma_l}{2(l+u)^2} \alpha^T \mathbf{K} \mathcal{L}' \mathbf{K} \alpha \end{aligned} \tag{10}$$

$$\begin{aligned}
 \text{s.t. } & y_i \left(\sum_{j=1}^{l+u} \alpha_j \mathbf{K}(x_i, x_j) + b \right) \geq 1 - \epsilon_i, \quad i = 1, \dots, l \\
 & \epsilon_i \geq 0, \quad i = 1, \dots, l
 \end{aligned} \tag{11}$$

With introduction of the Lagrangian multipliers β and γ , we write the Lagrangian function of the above equation as follows:

$$\begin{aligned}
 L(\alpha, \epsilon, b, \beta, \gamma) = & \frac{1}{2} \alpha^T \mathbf{K} \left(I + C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right) \alpha + C_l \sum_{i=1}^l \epsilon_i \\
 & - \sum_{i=1}^l \beta_i \left(y_i \left(\sum_{j=1}^{l+u} \alpha_j \mathbf{K}(x_i, x_j) + b \right) - 1 + \epsilon_i \right) - \sum_{i=1}^l \gamma_i \epsilon_i
 \end{aligned} \tag{12}$$

Obtaining the dual representation, requires taking the following steps:

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \beta_i \gamma_i = 0 \tag{13}$$

$$\frac{\partial L}{\partial \epsilon_i} = 0 \rightarrow C_l - \beta_i - \gamma_i = 0 \rightarrow 0 \leq \beta_i \leq C_l \tag{14}$$

With the above equations, we formulate the reduced Lagrangian as a function of only α and β as follows:

$$\begin{aligned}
 L^R(\alpha, \beta) = & \frac{1}{2} \alpha^T \mathbf{K} \left(I + C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right) \alpha \\
 & - \sum_{i=1}^l \beta_i \left(y_i \left(\sum_{j=1}^{l+u} \alpha_j \mathbf{K}(x_i, x_j) + b \right) - 1 + \epsilon_i \right)
 \end{aligned} \tag{15}$$

This equation is further simplified as follows:

$$\begin{aligned}
 L^R(\alpha, \beta) = & \frac{1}{2} \alpha^T \mathbf{K} \left(I + C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right) \alpha \\
 & - \alpha^T \mathbf{K} \mathbf{J}^T \mathbf{Y} \beta + \sum_{i=1}^l \beta_i
 \end{aligned} \tag{16}$$

In the above equation, $\mathbf{J} = [\mathbf{I} \ \mathbf{0}]$ is a $l \times (l+u)$ matrix, \mathbf{I} is the $l \times l$ identity matrix and \mathbf{Y} is a diagonal matrix consisting of the labels of the labeled examples.

In the followings, we first take the derivative of L^R with respect to α and then set $\frac{\partial L^R(\alpha, \beta)}{\partial \alpha} = 0$:

$$\mathbf{K} \left(I + C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right) \alpha - \mathbf{K} \mathbf{J}^T \mathbf{Y} \beta = 0 \tag{17}$$

Accordingly, we obtain α^* by solving the following equation:

$$\alpha^* = \left(I + C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right)^{-1} \mathbf{J}^T \mathbf{Y} \beta^* \tag{18}$$

Next, we obtain the dual problem in the form of a quadratic programming problem by substituting α back in the reduced Lagrangian function:

$$\beta^* = \operatorname{argmax}_{\beta \in \mathbb{R}^l} - \frac{1}{2} \beta^T \mathbf{Q} \beta + \sum_{i=1}^l \beta_i \tag{19}$$

$$\text{s.t. } \sum_{i=1}^l \beta_i \gamma_i = 0 \tag{20}$$

$$0 \leq \beta_i \leq C_l$$

where $\beta = [\beta_1, \dots, \beta_l]^T \in \mathbb{R}^l$ are the Lagrangian multipliers and \mathbf{Q} is obtained as follows:

$$\mathbf{Q} = \mathbf{Y} \mathbf{K} \left(I + \left(C_r \mathcal{L} + \frac{\gamma I}{(l+u)^2} \mathcal{L}' \right) \mathbf{K} \right)^{-1} \mathbf{J}^T \mathbf{Y} \tag{21}$$

We summarize the proposed semi-supervised framework in Algorithm 1. Our optimization problem is very similar to the standard optimization problem solved for SVMs, hence we use a standard optimizer for SVMs to solve our problem.

Algorithm 1 The Proposed Semi-Supervised Framework

Input: $\{(x_i, y_i)\}_{i=1}^l, \{x_{l+i}\}_{i=1}^u, \mathcal{F}_1, \mathcal{F}_2, C_l, C_r, C_s$.

Output: Estimated function $f_\theta : \mathbb{R}^n \rightarrow \mathbb{R}$

- 1: Construct matrix \mathbf{F} based on the features in \mathcal{F}_1
 - 2: Compute the corresponding Laplacian matrix \mathcal{L} .
 - 3: Construct \mathbf{A} according to the features in \mathcal{F}_2 .
 - 4: Compute the graph Laplacian matrix \mathcal{L}' .
 - 5: Construct gram matrix over all examples using $\mathbf{K}_{ij} = k(x_i, x_j)$ where k is the kernel function.
 - 6: Compute α^* and β^* using Eq. 18 and Eq. 19 and standard QP solvers.
 - 7: Compute function $f_\theta^*(x) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(x, x_i)$
-

Experimental study

In this section, we provide a comprehensive analysis of the proposed framework by designing a series of experiments on the filtered dataset. First, we explain several approaches used in this study. Next, various results are discussed: (1) comparisons on the *labeled* data were made between our method and other approaches, (2) experiments were performed on a fraction of the unlabeled data (i.e., unseen data), and the results were further verified by our expert to see what fraction is of interest to law enforcement, (3) *blind* evaluation was conducted to examine other approaches on the unseen data, and finally, (4) experiments were designed to analyze effect of varying different hyperparameters on our method as well as impact of different groups of features in \mathcal{F}_1 on our approach.

Approaches

We present results for the following methods:

- *Semi-supervised $S^3VM - R$* , Laplacian support vector machines [14], graph inference based label spreading approach [24] with radial basis function (RBF) and K-nearest neighbors (KNN) kernels, and co-training learner [36] with two support vector machines classifiers (SVM).
- *Supervised SVM*, KNN, Gaussian naïve Bayes, logistic regression, adaboost and random forest.

For the sake of fair comparison, all algorithms were implemented and run in Python. More specifically, the Python package CVXOPT² was used to implement $S^3VM - R$ and Laplacian support vector machines, and all other approaches were implemented with the help of the Scikit-learn³ package in Python. Note for those methods that require special tuning of parameters, we performed grid search to choose the best set of parameters. Before going any further, we first define main parameters used in each method and then demonstrate their best values picked by our grid search. The discussion on the effect of varying the hyperparameters on our learner is provided in the section “Hyperparameter sensitivity”.

- $S^3VM - R$ we set the penalty parameter as $C_l = 0.6$ and the regularization parameters $C_r = 0.2$ and $C_s = 0.2$. Linear kernel was used in our approach.
- *Laplacian SVM* we used linear kernel and set the parameters $C_l = 0.6$ and $C_s = 0.6$.

- *LabelSpreading (RBF)* RBF Kernel was used and γ was set to the default value of 20.
- *LabelSpreading (KNN)* KNN kernel was used and the number of neighbors was set to 5.
- *Co-training (SVM)* we followed the algorithm introduced in [36] and used two SVM as our classifiers. For both SVMs we set the tolerance for stopping criteria to 0.001 and the penalty parameter $C = 1$.
- *SVM* tolerance for stopping criteria was set to the default value of 0.001. Penalty parameter C was set to 1 and linear kernel was used.
- *KNN* number of neighbors was set to 5.
- *Gaussian NB* there were no specific parameter to tune.
- *Logistic regression* we used the ‘l2’ penalty. We also set the parameter $C = 1$ (the inverse of regularization strength) and tolerance for stopping criteria to 0.01.
- *Adaboost* number of estimators was set to 200 and we also set the learning rate to 0.01.
- *Random forest* we used 200 estimators and the ‘entropy’ criterion was used.

Classification results

Here, we first evaluate the entire set of approaches on a small portion of the data for which we already know the labels, i.e., the *labeled* examples. We note that expert-generated judgmental labeling might be error-prone, though it is served as a surrogate to the ground truth problem.

We used tenfold cross-validation on the labeled data in the following way. We first divided the set of the labeled samples into 10 different sets of approximately equal size. Each time we held one set out for validation (by removing their labels and adding them to the unlabeled samples) and used the remaining along with the unlabeled samples for the training—this was performed for all approaches for the sake of fair comparison. Finally, we reported the average of 10 different runs, using different combinations of the feature spaces and various evaluation metrics, including the area under curve (AUC), accuracy, precision, recall and F1-score. In Table 3, we reported the average AUC and accuracy for each method and each feature space. On the other hand, for precision, recall and F1-score, we reported separate results for each feature space, in Tables 4, 5 and 6, respectively. Note, each of these tables includes separate scores for the positive and negative classes. In general, we observed the followings:

- Overall, our approach achieved highest performance on \mathcal{F}_1 (Tables 3, 4) and $\{\mathcal{F}_1, \mathcal{F}_2\}$ (Table 6), in terms of all metrics. However it did not perform well using solely \mathcal{F}_2 (Table 5), i.e. when $C_r = 0$. This clearly demonstrates the importance of using C_r over C_s .

² <http://cvxopt.org/>.

³ <http://scikit-learn.org/stable/>.

Table 3 AUC and accuracy results with tenfold cross-validation on the labeled data

Learner	AUC			Accuracy		
	\mathcal{F}_1	\mathcal{F}_2	$\{\mathcal{F}_1, \mathcal{F}_2\}$	\mathcal{F}_1	\mathcal{F}_2	$\{\mathcal{F}_1, \mathcal{F}_2\}$
<i>S³VM – R</i>	<i>0.91</i>	0.9	<i>0.96</i>	<i>0.91</i>	0.9	<i>0.97</i>
Laplacian SVM	0.9	0.9	0.9	0.91	0.9	0.92
LabelSpreading (RBF)	0.78	0.87	0.84	0.8	0.85	0.86
LabelSpreading (KNN)	0.68	0.80	0.74	0.71	0.8	0.8
Co-training (SVM)	0.82	<i>0.94</i>	0.92	0.85	<i>0.94</i>	0.93
SVM	0.82	0.9	0.91	0.85	0.92	0.93
KNN	0.76	0.91	0.81	0.79	0.92	0.84
Gaussian NB	0.78	0.91	0.9	0.82	0.9	0.9
Logistic regression	0.82	0.89	0.88	0.85	0.92	0.92
AdaBoost	0.82	0.85	0.85	0.85	0.88	0.88
Random forest	0.81	0.89	0.89	0.83	0.91	0.92

The best performance is in italic

Table 4 Precision, recall and F1-score for the positive and negative classes using \mathcal{F}_1

Learner	Precision		Recall		F1-score	
	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>
<i>S³VM – R</i>	<i>0.91</i>	<i>0.92</i>	<i>0.91</i>	<i>0.93</i>	<i>0.91</i>	<i>0.92</i>
Laplacian SVM	0.86	0.89	0.88	0.9	0.87	0.88
LabelSpreading (RBF)	0.76	0.78	0.77	0.73	0.8	0.81
LabelSpreading (KNN)	0.65	0.7	0.71	0.68	0.69	0.73
Co-training (SVM)	0.81	0.84	0.71	0.92	0.73	0.87
SVM	0.86	0.83	0.68	0.96	0.74	0.88
KNN	0.72	0.8	0.63	0.88	0.65	0.83
Gaussian NB	0.79	0.81	0.72	0.85	0.73	0.81
Logistic regression	0.81	0.85	0.71	0.93	0.74	0.88
AdaBoost	0.86	0.83	0.68	0.95	0.74	0.88
Random forest	0.77	0.85	0.73	0.89	0.73	0.86

Experiments were run using tenfold cross-validation on the labeled data. The best performance is in italic

Table 5 Precision, recall and F1-score for the positive and negative classes using \mathcal{F}_2

Learner	Precision		Recall		F1-score	
	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>
<i>S³VM – R</i>	0.91	0.9	0.9	0.9	0.91	0.92
Laplacian SVM	0.91	0.9	0.9	0.91	0.89	0.92
LabelSpreading (RBF)	0.8	0.86	0.82	0.83	0.81	0.85
LabelSpreading (KNN)	0.7	0.75	0.73	0.78	0.79	0.77
Co-training (SVM)	<i>0.96</i>	0.91	0.91	<i>0.97</i>	0.93	<i>0.93</i>
SVM	0.93	0.91	0.84	<i>0.97</i>	0.87	<i>0.93</i>
KNN	0.87	0.92	0.88	0.94	0.87	<i>0.93</i>
Gaussian NB	0.78	<i>0.96</i>	<i>0.94</i>	0.87	0.84	0.91
Logistic regression	0.98	0.89	0.81	0.98	0.88	<i>0.93</i>
AdaBoost	0.88	0.88	0.75	0.95	0.78	0.91
Random forest	0.93	0.89	0.81	<i>0.97</i>	0.85	<i>0.93</i>

Experiments were run using tenfold cross-validation on the labeled data. The best performance is in italic

Table 6 Precision, recall and F1-score for the positive and negative classes using $\{\mathcal{F}_1, \mathcal{F}_2\}$

Learner	Precision		Recall		F1-score	
	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>	<i>class_p</i>	<i>class_n</i>
$S^3VM - R$	0.97	0.97	0.95	0.98	0.94	0.95
Laplacian SVM	0.96	0.94	0.91	0.96	0.91	0.93
LabelSpreading (RBF)	0.83	0.86	0.82	0.84	0.81	0.86
LabelSpreading (KNN)	0.71	0.74	0.75	0.78	0.8	0.78
Co-training (SVM)	0.92	0.9	0.9	0.94	0.91	0.92
SVM	0.96	0.92	0.84	0.97	0.89	0.94
KNN	0.84	0.83	0.67	0.95	0.73	0.88
Gaussian NB	0.77	0.96	0.94	0.87	0.84	0.91
Logistic regression	0.95	0.9	0.79	0.97	0.85	0.93
AdaBoost	0.88	0.88	0.75	0.95	0.78	0.91
Random forest	0.93	0.9	0.82	0.97	0.86	0.93

Experiments were run using tenfold cross-validation on the labeled data. The best performance is in italic

- When the feature space used is \mathcal{F}_2 , Co-training (SVM) is the best method. Next best methods are supervised learners KNN and Gaussian NB. Three remarks can be made here. First, our approach could not always defeat supervised learners as it is seen from Tables 3 and 5. This is not surprising and in fact lies at the inherent difference between semi-supervised and supervised methods—unlabeled examples could make the trained model susceptible to error propagation and thus wrong estimation. Second, as it is seen in Tables 4, 5 and 6, achieving very high recall on the negative examples and low score on the positive ones shall not be treated as a potent property, otherwise a trivial classifier which always assigns negative labels to all samples would be the best learner. Third, using C_r always improves the performance over C_s . One point that needs to be clarified is, our ultimate goal is not to achieve high performance on the labeled data, but rather to detect the suspicious (unlabeled) advertisements which could be human trafficking related—this will be explained in more details in “Blind evaluation”.
- Compared to the other semi-supervised approaches, our approach either achieved higher or comparable AUC scores. The reason we performed exactly the same as the Laplacian SVM, is because by setting $C_r = 0$, the two approaches are inherently the same.
- For the Laplacian SVM to be able to run on \mathcal{F}_1 , the Laplacian \mathcal{L}' has to be constructed using \mathcal{F}_1 while inherently is supposed to be made using \mathcal{F}_2 . This is because C_r is essentially associated with \mathcal{F}_1 , and C_s corresponds to \mathcal{L}' and correspondingly \mathcal{F}_2 . The same holds for $\{\mathcal{F}_1, \mathcal{F}_2\}$, where we need to construct a new feature space by concatenating \mathcal{F}_1 and \mathcal{F}_2 as

the Laplacian SVM does not inherently use \mathcal{F}_1 at all. The new feature space is then used to construct the Laplacian \mathcal{L}' .

- Since our approach inherently incorporates both of the Laplacian matrices corresponding to the two feature spaces \mathcal{F}_1 and \mathcal{F}_2 , all other baselines were also run using the concatenation of these two feature spaces for the sake of fair comparison. Unlike our approach which used the wise combination of \mathcal{F}_1 and \mathcal{F}_2 , other methods do not gain high AUC by simply combining the feature spaces.

Blind evaluation

For the next set of experiments, we first run our method on the *entire* filtered dataset and without cross-validation. Recall from the previous sections that this is to make a better use of the unlabeled examples. Then the following *control* experiment was conducted. Our learner was tested on the whole set of the unlabeled examples. Out of 3343 instances, our approach identified two sets of positive and negative instances. The positive set contained 394 advertisements which were likely to be of interest to law enforcement, whereas the negative set included the remaining 2962 unlabeled advertisements of probably less interest to law enforcement. Next, to precisely determine the correctly identified fractions of these two sets, we randomly picked two subsets (control groups) of 100 examples from each set for further validation by our expert.

We passed these two control groups to our expert for further verification. The expert-validated results demonstrated that all of the examples in the positive group were

of interest to law enforcement, while only two examples from the negative group were not correctly classified as of not being of any interest to law enforcement. Thus, both results support the effectiveness of our framework in identifying highly human trafficking advertisements. Using the same two control groups and AUC metric, we now perform so-called blind evaluation (see Table 7) of other baselines. Note, we call this blind since actual labels are not provided and the expert-generated labels might convey uninformative information. In general, supervised methods failed to achieve good results in the blind evaluation compared to most of the semi-supervised methods.

Hyperparameter sensitivity

Here, we discuss how altering the hyperparameters C_l , C_r and C_s may affect the performance of $S^3VM - R$. We start off by fixing the value of C_l to 0.6, which was empirically found to work well in our experiments. Also, recall from the previous sections that one typical choice for C_s is $\frac{\gamma_l}{(l+u)^2}$ [14]. Here, we set $C_s = 0.2$ and varied the values of C_r as $\{0, 0.0002, 0.0006, 0.2, 1.0\}$ and plotted the results

Table 7 Blind evaluation of the baselines on the two control groups

Learner	AUC		
	\mathcal{F}_1	\mathcal{F}_2	$\{\mathcal{F}_1, \mathcal{F}_2\}$
Laplacian SVM	0.9	0.92	0.93
LabelSpreading (RBF)	0.75	0.85	0.87
LabelSpreading (KNN)	0.7	0.82	0.79
Co-training (SVM)	0.8	0.9	0.91
SVM	0.8	0.65	0.69
KNN	0.74	0.62	0.77
Gaussian NB	0.77	0.51	0.52
Logistic regression	0.76	0.62	0.75
AdaBoost	0.77	0.74	0.74
Random forest	0.8	0.8	0.8

The best performance is in italic

in Fig. 5. We used the same tenfold cross-validation setting from the previous section.

We made the following observation. With the slight increase of C_r , the performance of our approach increased, peaked and then stabilized, i.e., further increase of C_r did not change the performance. This suggests significance of deploying the additional information from our first feature space \mathcal{F}_1 , over \mathcal{F}_2 and its corresponding smoothness penalty parameter C_s which is used by $S^3VM - R$ and the standard Laplacian SVM.

Next, to see the impact of C_l on the performance, we set $C_r = 0.2$ and varied C_l as $\{0.2, 0.4, 0.6, 0.8, 1.0\}$. The results are depicted in Fig. 5. We note that setting $C_l = 0$ is meaningless and thus we do not have any performance corresponding to that—otherwise each β_i in Eq. 19 would be zero. In general, the performance was not particular sensitive to this parameter—varying by 0.2 for values of 0.4 and greater.

Finally, having fixed $C_l = 0.6$ and $C_r = 0.2$, we also tried other values for C_s including $\sum_{i,j=1}^{l+u} W_{ij}$ suggested by [14] and depicted the results in Fig. 5. The results suggest that our approach is less sensitive to this parameter compared to C_r and C_l .

Significance of features

To examine how much discriminative our feature groups in \mathcal{F}_1 are, we further conducted an analysis using the labeled examples and the standard feature selection measure χ^2 to find the top features—only half of the features with scores greater than a given threshold (0.5) were selected (see Table 8 for the complete set of features and their corresponding χ^2 scores).

From this list, we noticed that ‘countries of interest’ and ‘reference to spa massage therapy’ were the most discriminative feature groups, while ‘advertisement language pattern’ group (with 3 important features) appeared to be the most dominant feature group.

Figure 6 compares the top features against the less important subset of the features (denoted by $\overline{\mathcal{F}}_1^*$) in the filtered dataset, in terms of frequency values. Note for clarity, we have removed from this figure, the features with frequency less than 20. According to this figure, our

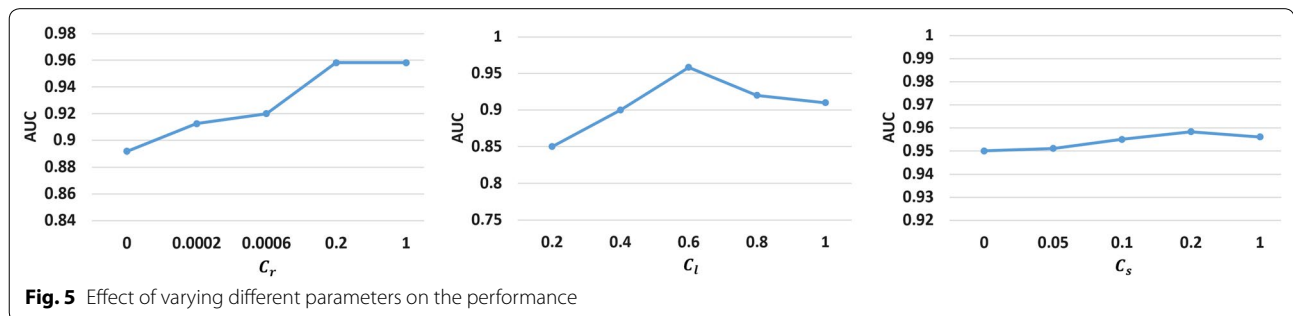


Fig. 5 Effect of varying different parameters on the performance

Table 8 Significance of the features in \mathcal{F}_1

No.	Feature group	χ^2	Selected
1	Advertisement language pattern		
	Third person language	8.4	✓
	First person plural pronouns	9.5	✓
	Kolmogorov complexity	0.7	✓
	<i>n</i> -grams (1)	0.4	
	<i>n</i> -grams (2)	0.0	
	<i>n</i> -grams (3)	0.4	
2	Words and phrases of interest	0.0	
3	Countries of interest	59.3	✓
4	Multiple victims advertised	14.1	✓
5	Victim weight	0.2	
6	Reference to website or spa massage therapy		
	Reference to website	0.1	
	Reference to spa massage therapy	33.5	✓

The check-marked features show the top features

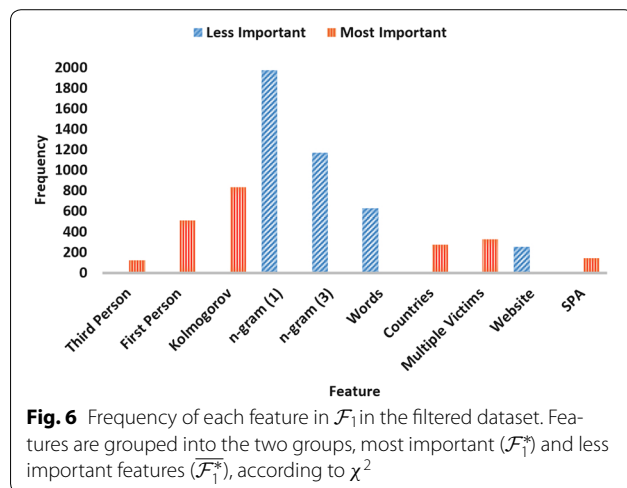


Fig. 6 Frequency of each feature in \mathcal{F}_1 in the filtered dataset. Features are grouped into the two groups, most important (\mathcal{F}_1^*) and less important features ($\overline{\mathcal{F}_1^*}$), according to χ^2

most discriminative features are not necessarily those that appear more often.

To further investigate the importance of each of the top features, we performed classification using the labeled examples and the previous setting, on basis of these two subsets of the features and their combination, i.e., \mathcal{F}_1^* , $\overline{\mathcal{F}_1^*}$ and \mathcal{F}_1 . The classification results are shown in Table 9. We made the following observations:

Table 9 Classification results (AUC) using tenfold cross-validation and different subsets of the features on the labeled data

Name	Value		
	$\overline{\mathcal{F}_1^*}$	\mathcal{F}_1^*	\mathcal{F}_1
$S^3VM - R$	0.82	0.87	0.91

- Considering only the feature space \mathcal{F}_1 , our approach achieved higher performance compared to all other baselines by either using the whole feature space or the most discriminative features \mathcal{F}_1^* .
- Deploying only the features from \mathcal{F}_1^* , we were able to achieve comparable results as if we used the whole feature space \mathcal{F}_1 .

Conclusion

Readily available online data from escort advertisements could be leveraged in favor of fight against human trafficking. In this study, having focused on textual information from the available data crawled from Backpage.com, we identified if an escort advertisement can be reflective of human trafficking activities. In particular, we first proposed an unsupervised filtering approach to filter out the data which are more likely involved in human trafficking. We then proposed a semi-supervised learner, namely $S^3VM - R$, and trained it on a small portion of the data which was hand-labeled by a human trafficking expert. We used the trained model to identify labels of unseen data. Results suggested our approach is effective at identifying potential human trafficking related advertisements.

Our future plans include replicating the study by integrating more interesting features especially those supported by the criminology literature. Also, since hand-labeling unlabeled examples is expensive, an interesting research direction would be to deploy active learning to enable iterative supervised learning to actively query the user for labels. We also note that real-world data is often more imbalanced compared to our data, and the reason is that number of negative samples usually outweigh positive ones. We would thus like to apply the proposed framework on a more realistic dataset which contains much less suspicious posts than normal posts.

Authors' contributions

HA developed and implemented the human trafficking detection approach and drafted the manuscript. PS provided guidance through the whole project and revised the manuscript. JS was in contact with an expert from law enforcement who was responsible for hand-labeling portions of the data. All authors read and approved the final manuscript.

Author details

¹ Arizona State University, Tempe, Arizona, USA. ² Find Me Group, Tempe, Arizona, USA.

Acknowledgements

This work was funded by the Find Me Group, a 501(c)3 dedicated to bring resolution and closure to families of missing persons. The authors would like to thank anonymous reviewers for their valuable suggestions to improve the quality of the paper.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 20 December 2016 Accepted: 9 April 2017

Published online: 11 May 2017

References

- UNODC on human tracking and migrant smuggling. <https://www.unodc.org/unodc/en/human-tracking/>. Accessed 20 Dec 2016
- (2000) Trafficking victims protection act of 2000. <https://www.state.gov/j/tip/laws/61124.htm>. Accessed 20 Dec 2016
- (2015) Trafficking in persons report. <https://www.state.gov/j/tip/rls/tiprpt/2015/>. Accessed 20 Dec 2016
- Desplaces C (1992) Police run 'Prostitution' sting; 19 men arrested, charged in Fourth East Dallas operation. Dallas Morning News
- Kristof ND (2012) How pimps use the web to sell girls. New York Times. <http://www.nytimes.com/2012/01/26/opinion/how-pimps-use-the-web-to-sell-girls.html>. Accessed 20 Dec 2016
- Kennedy E (2012) Predictive patterns of sex trafficking online. Dietrich College Honors Theses, B.S. thesis, Carnegie Mellon University
- Mitchell TM (2006) Learning from labeled and unlabeled data. Mach Learn 10:701
- Beigi G, Tang J, Liu H (2016) Signed link analysis in social media networks. In: 10th international conference on web and social media, ICWSM 2016, AAAI Press, Cologne, Germany
- Backstrom L, Leskovec J (2011) Supervised random walks: predicting and recommending links in social networks. In: Proceedings of the fourth ACM international conference on Web search and data mining. WSDM '11, Hong Kong, China. ACM, New York, pp 635–644. doi:10.1145/1935826.1935914
- Alvari H, Hajibagheri A, Sukthankar G, Lakkaraju K (2016) Identifying community structures in dynamic networks. Soc Netw Anal Min SNAM 6(1):77. doi:10.1007/s13278-016-0390-5
- Beigi G, Tang J, Wang S, Liu H (2016) Exploiting emotional information for trust/distrust prediction. In: Proceedings of the 2016 SIAM international conference on data mining (ICDM), SIAM, Miami, FL, USA
- Mitchell et al TM (1997) Machine learning, I–XVII. McGraw-Hill Education. <https://www.bibsonomy.org/bibtex/2e1eee0d0daef20092093d6643b53c4f/machinelearning>. Accessed 20 Dec 2016
- Beigi G, Jalili M, Alvari H, Sukthankar G (2014) Leveraging community detection for accurate trust prediction. In: ASE international conference on social computing, Palo Alto, CA
- Belkin M, Niyogi P, Sindhvani V (2006) Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J Mach Learn Res 7(Nov):2399–2434
- Alvari H, Shakarian P, Snyder J (2016) A non-parametric learning approach to identify online human trafficking. In: IEEE intelligence and security informatics (ISI) conference, pp. 133–138
- Hughes DM et al (2005) The demand for victims of sex trafficking. Womens Studies Program, University of Rhode Island, RI
- Hughes DM (2002) The use of new communications and information technologies for sexual exploitation of women and children. Hastings Women's Law J 13(1):129–148
- Latonero M (2011) Human trafficking online: the role of social networking sites and online classifieds. Available at SSRN 2045851
- Roe-Sepowitz D, Gallagher J, Bracy K, Cantelme L, Bayless A, Larkin J, Reese A, Allbee L (2015) Exploring the impact of the super bowl on sex trafficking. <https://www.mccainstitute.org/exploring-the-impact-of-the-super-bowl-on-sex-tracking-2015/>. Accessed 20 Dec 2016
- Miller K, Kennedy E, Dubrawski A (2016) Do public events affect sex trafficking activity? <https://arxiv.org/abs/1602.05048>. Accessed 20 Dec 2016
- Szekely PA, Knoblock CA, Slepicka J, Philpot A, Singh A, Yin C, Kapoor D, Natarajan P, Marcu D, Knight K, Stallard D, Karunamoorthy SS, Bojanapalli R, Minton S, Amanatullah B, Hughes T, Tamayo M, Flynt D, Artiss R, Chang S-F, Chen T, Hiebel G, Ferreira L (2015) Building and using a knowledge graph to combat human trafficking. In: International semantic web conference (2), ser. Lecture notes in computer science, vol 9367. Springer, Switzerland, pp 205–221
- Nagpal C, Miller K, Boecking B, Dubrawski A (2015) An entity resolution approach to isolate instances of human trafficking online. <https://arxiv.org/abs/1509.06659>. Accessed 20 Dec 2016
- Dubrawski A, Miller K, Barnes M, Boecking B, Kennedy E (2015) Leveraging publicly available data to discern patterns of human-trafficking activity. J Hum Traffick 1(1):65–85. doi:10.1080/23322705.2015.1015342
- Zhou D, Bousquet O, Lal TN, Weston J, Scholkopf B (2004) Learning with local and global consistency. In: Advances in neural information processing systems 16. MIT Press, Cambridge, MA, pp. 321–328
- Li M, Vitnyi PM (2008) An introduction to kolmogorov complexity and its applications, 3rd edn. Springer, New York
- Kanaris I, Kanaris K, Stamatatos E (2006) Spam detection using character n-grams. In: Antoniou G, Potamias G, Spyropoulos C, Plexousakis D (eds), Advances in Artificial Intelligence. SETN 2006. Lecture Notes in Computer Science, vol 3955. Springer, Berlin, pp 95–104
- Hetter K (2012) Fighting sex trafficking in hotels, one room at a time. Cable News Network. <http://www.cnn.com/2012/02/29/travel/hotel-sex-tracking/>. Accessed 20 Dec 2016
- Lloyd R (2012) An open letter to jim Buckmaster. Change.org Inc. http://www.hungtonpost.com/rachel-lloyd/an-open-letter-to-jim-buc_b_570666.html. Accessed 20 Dec 2016
- Dickinson Goodman J, Holmes M (2011) Can we use RSS to catch rapists. Grace Hopper Celebration of Women in Computing, Portland, OR. <http://jessicadickinsongoodman.com/2011/10/30/can-we-use-rss-to-catch-rapists-poster-nished/>. Accessed 20 Dec 2016
- Average height to weight chart—babies to teenagers. <http://www.disabled-world.com/artman/publish/height-weight-teens.shtml>. Accessed 20 Dec 2016
- Shannon CE (2001) A mathematical theory of communication. ACM SIGMOBILE Mob Comput Commun Rev 5(1):3–55
- van der Maaten L, Hinton G (2008) Visualizing high-dimensional data using t-SNE. Journal of Machine Learning Research 9:2579–2605
- Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
- Belkin M, Niyogi P, Sindhvani V (2004) On manifold regularization. Technical report tr-2004-05. The University of Chicago
- Grigoryan A (2006) Heat kernels on weighted manifolds and applications. Contempl Math 398:93–191
- Blum A, Mitchell T (1998) Combining labeled and unlabeled data with co-training. In: Proceedings of the eleventh annual conference on Computational learning theory. ACM, pp 92–100

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com