



# Human Trafficking in Social Networks: A Review of Machine Learning Techniques

Mike Bermeo<sup>(✉)</sup> , Silvana Escobar , and Erick Cuenca 

School of Mathematical and Computational Sciences, Yachay Tech University, Hda.  
San José s/n y Proyecto Yachay, San Miguel de Urququí 100119, Ecuador  
{mike.bermeo,sescobar,ecuenca}@yachaytech.edu.ec

**Abstract.** Human trafficking is a severe problem worldwide and social media platforms have emerged as a potential tool to detect and prevent this crime. Machine learning (ML) algorithms have shown promise in identifying human trafficking activity on these platforms. This paper comprehensively reviews ML techniques for human trafficking detection on social media, including supervised, unsupervised, and semi-supervised approaches. We identify each approach's advantages, limitations, and challenges for human trafficking detection. Finally, we provide future directions for research in this field, including the need for more standardized datasets and the development of explainable machine learning models to increase transparency and accountability. Our review provides a better understanding of the potential of machine learning in combating human trafficking and to guide future research efforts in this field.

**Keywords:** social media · human trafficking · machine learning · Twitter

## 1 Introduction

Human trafficking is a global problem affecting millions of people worldwide [29]. It is a form of modern slavery that involves the exploitation of individuals for various purposes, including forced labor, sexual exploitation, and organ removal, among others [43]. In recent years, social networks have emerged as a key platform for human traffickers to operate and reach potential victims [39]. Researchers and policymakers are interested in the role that social networks and other internet-based platforms play as facilitators of this type of crime that affects millions of people globally. Efforts are being made to employ various methodologies to identify and prevent these forms of exploitation.

The misuse of technology for human trafficking in all stages is increasing. For example, traffickers use deception techniques to hide their identities and avoid detection [25]. They can also use social networks to lure potential victims, often using emotional manipulation and other tactics to gain their trust [43].

In addition, the vast amount of data generated on social networks can make it difficult to identify and trace traffickers and their victims [4].

However, social media can also serve as a valuable tool for raising awareness of human trafficking and gathering support for victims and survivors [35]. Through social media campaigns and sharing of information and resources, individuals and organizations are able to take action in order to combat human trafficking. Additionally, social media provide a space for survivors to share their stories and connect with others, helping to break down isolation and stigma. While social media functionalities have the potential to play a significant role in anti-trafficking efforts, it is important to recognize its limitations and the need for broader actions to address the root causes of human trafficking and provide support to the victims [18].

In this context, technological advances present new opportunities for the detection and prevention of human trafficking. For instance, data generated on social networks can be utilized to identify patterns and trends in human trafficking activity [39]. Also, Machine Learning (ML) techniques and other data-driven approaches can be employed to create algorithms capable of automatically identifying and flagging suspicious activity on social networks, helping to address this social problem [17]. These technologies can also aid in the detection of human trafficking advertisements and the identification of relevant keywords used in social media to facilitate these crimes. [47].

In this paper, we present a comprehensive review of the current work that explores the application of Machine Learning techniques in social networks and other platforms to tackle the issue of human trafficking. Additionally, we delve into the challenges and opportunities associated with using ML approaches to detect and prevent human trafficking on social networks. The article is organized as follows: Sect. 2 presents the fundamental knowledge about ML techniques. In Sect. 3, we describe the methodology to follow to collect the related work. Section 4 reviews the existing research on ML techniques that address human trafficking on social media. Finally, in Sect. 5, we conclude the paper and discuss future directions for research in this area.

## 2 Background Study

This section provides definitions of human trafficking, discusses how this phenomenon behaves in social networks and outlines the concepts of Machine Learning, the various types of learning, and the techniques employed in each.

### 2.1 Human Trafficking on Social Media

Human trafficking is characterized by using force or coercion to recruit, transport, and exploit individuals for various purposes, including sexual exploitation, forced labor, and organ removal. It involves the abuse of power or vulnerability and the offering or receiving of payments or benefits in exchange for consent from those in positions of authority. Exploitation can include sexual exploitation, such

as prostitution, and other forms of exploitation, including forced labor, slavery, servitude, and organ removal [43].

Social media platforms have become increasingly important in recruiting and grooming potential trafficking victims. For example, traffickers may use social media to lure individuals with false promises of employment or romantic relationships and then exploit them once they have been recruited [43]. In some cases, traffickers may use social media to advertise their victims for sexual exploitation or forced labor or to arrange transportation to different locations [3].

There are several ways in which social media can facilitate human trafficking [43]. First, it allows traffickers to reach a broad audience and target specific demographics, such as young people or those vulnerable due to economic or social circumstances. Second, social media can provide anonymity and secrecy, allowing traffickers to operate without detection. Third, it can be used to obscure the true nature of the exploitation, for example, by presenting it as legitimate work or a consensual relationship.

Combating human trafficking on social media requires a multifaceted approach involving governments, law enforcement agencies, stakeholders, and technology companies. They must work together to combat human trafficking and protect victims' rights. In addition, efforts to raise awareness and educate the public about the risks of human trafficking on social media can help prevent individuals from falling victim to these crimes [43]. Human trafficking is a severe global problem that the proliferation of social media has exacerbated.

## 2.2 Machine Learning Techniques

Machine Learning (ML) techniques involve using algorithms and statistical models to analyze and identify patterns of large amounts of data [15]. These techniques have been used in various fields, including image recognition [11], speech recognition [21], sentiment analysis [30,31], and predictive modeling [22]. They can be used as well in the battle against human trafficking. For instance, examining the content of social media posts and finding any that might be about trafficking [24], identifying patterns in the movements of individuals that may indicate trafficking activity [5, 10, 12], detecting anomalies in employment records that could indicate the presence of forced labor [34], predicting the likelihood of individuals becoming victims of trafficking [38, 42], and identifying potential victims before they are exploited [27]. ML techniques are categorized into supervised, unsupervised, and semi-supervised learning.

- **Supervised Learning:** It is a subcategory of ML that uses labeled training data to predict outputs or classify data into predefined categories. The goal is to build a model to make accurate predictions or classifications based on input data. This is achieved by providing the model with a large set of labeled training examples consisting of input data and the corresponding correct outputs or classifications. The model is then trained to learn the relationship between the input data and the outputs or classifications, using this training data as a guide.

Considering human trafficking, several types of supervised learning algorithms have been used, including Linear regression, Logistic regression, K-Nearest Neighbors (KNN), Support Vector Machines (SVMs), Naive Bayes (NB), Random Forest (RF), Neural Networks (NN), Decision Tree (DT), AdaBoost, and so on.

- **Unsupervised Learning:** It is a type of ML in which a model is trained to discover patterns and relationships in a dataset without using labeled training examples [7]. In this learning, the goal is to find hidden patterns in the data rather than to make specific predictions or classifications [14]. This is achieved by providing the model with a large dataset and allowing it to learn the underlying structure of the data through techniques such as clustering [23] or dimensionality reduction [33].

It is also important to evaluate the model’s performance using appropriate evaluation metrics, such as silhouette scores (for clustering) or reconstruction errors (for dimensionality reduction) [36].

- **Semi-supervised Learning:** It is an ML type between supervised and unsupervised learning. It involves using labeled and unlabeled data to improve the accuracy of predictions or classifications where the goal is to leverage the available labeled data to make better predictions or classifications on the unlabeled data, using techniques such as self-training or co-training [44].

There are several semi-supervised learning algorithms, including self-training algorithms, co-training algorithms, and multi-view learning algorithms [44]. Self-training algorithms use a single model trained on labeled and unlabeled data [26]. Co-training algorithms involve using two or more models trained on different views of the data and used to label the unlabeled data iteratively [28]. Multi-view learning algorithms use multiple models trained on different data views and combined to make a final prediction or classification [48].

### 3 Methodology

The scope of this work is to review and compile previous approaches that assess human trafficking detection on social networks using ML techniques and extract the main hints and trending paths used to address this problem. To fulfill this, we define the following steps:

1. Select the most relevant papers based on a selection criteria to obtain the most relevant works in this area.
2. Provide a deep insight of the main aspects used to analyze Human Trafficking in social media with ML.

#### 3.1 Paper Selection

The selection criteria were based on the fact that the work should contain the following aspects:

- The focus of the work seeks to help the problem of human trafficking, such as sexual exploitation, forced labor, and modern slavery.

- One of the methods used to address the human trafficking problem must be a Machine Learning technique.
- The data used to work with must be a part of a social network or a website accessible to anyone.
- The work is at least from the year 2016.

The databases and repositories employed for this investigation included Scopus, ScienceDirect, ArXiv, IEEE XPLORE, and SpringerLink. This study aims to comprehensively review existing research on human trafficking by employing these resources and combining relevant variables. Specifically, the fixed variable “*human trafficking*” was identified and combined with three descriptive variables, namely “*social media*”, “*social networks*”, and “*machine learning*” to create multiple search queries. The utilization of these variables allowed for the identification of a diverse range of relevant papers. At the end of this search, 23 papers between 2016 and 2022 were selected.

### 3.2 Main Aspects

This work has considered six general aspects to analyze in each reviewed paper to identify the most relevant contributions. These aspects include the type of data used, the number of classes, the model or algorithm employed, the dataset utilized, the number of observations in the dataset, and the metrics that the paper considered to evaluate the performance of the ML algorithm.

## 4 Main Findings

This section analyzes the various sub-aspects based on the general aspects to be considered when addressing human trafficking in social networks through ML algorithms.

### 4.1 Supervised Approaches

Table 1 summarizes the works using supervised algorithms. Supervised ML algorithms are commonly used for detecting human trafficking. However, obtaining labeled data for human trafficking is challenging, as it involves sensitive information and can potentially put individuals at risk. Despite this challenge, several approaches have been taken to obtain labeled data, including using pre-existing data or manually labeling data pulled from social media platforms like Twitter.

The data extraction typically involves using web scraping techniques or social network APIs, followed by pre-processing steps such as removing links and non-relevant characters such as emoticons and emojis. The labeled data is then fed into the supervised ML algorithm for further analysis. Once the data is ready, it is commonly split into training and testing subsets, or in some cases, into three subsets: training, validation, and testing. The proportion of the split varies depending on the researcher, but a common split is 80% for training and 20% for testing.

**Table 1.** Supervised Machine Learning Approaches to Address Human Trafficking on Social Media.

Authors	Input Feature	Classes	Model	Dataset
Vieira et al. [45]	9Tweets: 13.6k Images: 11.1k	NA	OCR	TwitterPolice Data
Shishira et al. [38]	9Tweets: 155.1k Images: 89k	(suspicious, not suspicious), (man, woman), (under14, over14)	SVMVGG16	TwitterIMDB-WikiOther DS
Hernández and Granizo [13]	9Tweets: 55.1k Images: NA	(suspicious, not suspicious), (man, woman), (under14, over14)	SVM	Twitter
Wiriyakun et al. [47]	9Ads: 10k	(trafficking, not trafficking)	LIMENBDT ANN	Trafficking-10k
Granizo et al. [10]	9Tweets: 100k Images: 820	(suspicious, not suspicious), (man, woman), (under14, over14)	SVMNBCNN	Twitter
Diaz et al. [8]	9Reviews: 456k	(illegal activity, not illegal activity)	SVMNBRFANN	Yelp reviewsRubmaps DS
Hernandez [12]	9Tweets: 55.1k	(suspicious, not suspicious)	SVMNB	Twitter
Goist et al. [9]	9Text: 178 country reports	(sending relationship, receiving relationship)	LR SVMNBRF	TIP reports
Zhu et al. [49]	9Ads: 10k	(high-risk, low-risk)	KenLMLRSVM	Trafficking-10k
Tong et al. [40]	9Ads: 10k	(certainly no, likely no, weakly no, unsure, weakly yes, likely yes, certainly yes)	RFLRSVMHTDN	Trafficking-10k

Supervised learning algorithms in human trafficking detection include Support Vector Machine (SVM) [1, 5, 8–10, 12, 13, 38, 40, 49], Logistics Regression (LR) [1, 5, 9, 40], Random Forest (RF) [1, 5, 8, 9, 40], Gaussian Naive Bayes (NB) [8–10, 12, 47], and Artificial Neural Networks [8, 47]. These algorithms form the basis for supervised learning in classification, detection, and regression tasks.

To evaluate the performance of the classifiers, metrics such as Precision, Recall, F1-score, Accuracy, and AUC are commonly used. These metrics pro-

vide insights into the algorithm’s effectiveness in detecting human trafficking and can guide future improvements in the methodology. Overall, supervised learning approaches have proven effective in detecting human trafficking. Despite the challenges of obtaining labeled data, various approaches have been taken to acquire it, such as using pre-existing datasets and manually labeling data extracted from social media platforms.

## 4.2 Unsupervised Approaches

Table 2 summarizes the works using unsupervised algorithms. These techniques have also been used to detect human trafficking, as they do not require labeled data to train the algorithms as in supervised learning. For instance, Clustering algorithms [19, 20] are commonly used to similar group data together based on similarities in their features. One of the most used clustering algorithms in human trafficking detection is k-means clustering. Clustering [19, 20] and anomaly detection [6, 37] can also be effective tools in identifying instances of human trafficking since they can be used when labeled data is unavailable and can be used in conjunction with supervised learning algorithms to improve their accuracy.

Once the unsupervised techniques have been applied, the researcher can further refine the results by manually examining the data points flagged as potential instances of human trafficking. This manual approach can eliminate false positives and increase the accuracy of the results.

**Table 2.** Unsupervised Machine Learning Approaches to Address Human Trafficking on Social Media.

Authors	Input Feature	Classes	Model	Dataset
Li et al. [20]	Ads: 10M	(positive, negative)	HDBSCAN	Backpage
Burbano and Hernández [6]	Tweets: 167k	NA	NLP	Twitter
Lee et al. [19]	Tweets: +18M	(certainly no, likely no, weakly no, unsure, weakly yes, likely yes, certainly yes)	InfoShield	Twitter
	Ads: 10k			Trafficking-10k
Shelke et al. [37]	Images: +3M	NA	K-Means	FaceScrub <sup>a</sup> VGG <sup>b</sup> FPA <sup>c</sup>

<sup>a</sup><http://vintage.winklerbros.net/facescrub.html>

<sup>b</sup>[https://www.robots.ox.ac.uk/~vgg/data/vgg\\_face/](https://www.robots.ox.ac.uk/~vgg/data/vgg_face/)

<sup>c</sup><https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>

### 4.3 Semi-supervised Approaches

Table 3 summarizes the works using semi-supervised techniques that combine unsupervised and supervised techniques. They have shown promising results in human trafficking detection. One such hybrid approach is to use unsupervised techniques for word embedding and supervised techniques for the detection task.

Word embedding is a technique used to represent words in a vector space, where words with similar meanings are closer to each other. Word embedding techniques such as a bag of words [41], word2vec, TF-IDF, FastText, and Skip-

**Table 3.** Semi-supervised Machine Learning Approaches to Address Human Trafficking on Social Media.

Authors	Input Feature	Classes	Model	Dataset
Alvari et al. [1]	Ads: 20k	(positive, negative)	S3VM-R SVM KNN NB LR Adaboost RF	Backpage
Wang et al. [46]	Ads: 20kEmojis	(certainly no, likely no, weakly no, unsure, weakly yes, likely yes, certainly yes)	ORNN Skip-gram	Trafficking-10k
Alvari et al. [2]	Ads: 20k	NA	K-Means KNN RBF	Backpage
Kejriwal et al. [16]	Text	NA	Flexible and adaptive generation of indicators	Web Resources
Ramchandani et al. [32]	Posts: 13.5M	(suspicious, not suspicious)	CBOW	Deep Web
Burbano and Hernández [5]	News: 814	(related to human trafficking, not related)	SVM LR RF DT Perceptron SGD KNN	News
Tudis et al. [41]	Tweets: 122k	(suspicious, not suspicious)	RF BoW	Twitter



grams [46] are commonly used. In the context of human trafficking detection, word embedding can be used to represent words and phrases that are commonly associated with human trafficking, such as “sex trafficking” or “forced labor”.

Once the word embedding is generated, it can be used as input to a supervised learning algorithm, such as an ANN or an SVM [1, 5]. The supervised algorithm is trained on labeled data to learn the relationship between the word embedding and the presence of human trafficking activity. The labeled data can be obtained using earlier methods, such as manual labeling or web scraping.

The advantage of using a semi-supervised approach is that it leverages the strengths of both unsupervised and supervised techniques. Unsupervised techniques can generate high-quality word embedding, while supervised techniques can be used to learn the relationship between the embedding and human trafficking activity. This approach can also be effective when labeled data is limited, as it can augment the available labeled data with word embedding generated from a larger, unlabeled dataset.

Moreover, it is used for a small amount of labeled data to train a model and then used the trained model to label a larger amount of unlabeled data as in the work of [16]. The labeled and unlabeled data can be trained in a new model. This iterative process can be repeated until the model’s accuracy is satisfactory.

#### 4.4 Datasets

For data extraction, public web pages are used where there may be indications of human trafficking, such as pages of adult services and pornographic pages where sex work is offered [16]. Another common avenue for recruiters is social media. They can more easily contact potential victims by posing as friends, acquaintances, or job recruiters. Traffickers most widely use microblogging platforms because they allow for more meaningful interaction between strangers, and since they share everyday thoughts, they can more easily identify potential victims and form a relationship.

One of the most commonly used sites for data extraction using web scraping techniques is Backpage.com [1]. It was a classified ads website founded in 2004 which allowed users to post ads in categories such as personals, automotive, rentals, jobs, and adult services. The latter type was used by human trafficking rings to recruit potential victims for their network. For this same reason, it is widely used to extract advertisements related to human trafficking as used in [1].

Other sites used for data mining are news and advertisements [6, 8]. These are known to contain job offers that are methods used by traffickers to recruit new victims. Likewise, YELP reviews are also used to detect places where sexual services are provided based on keywords such as massage, spa, and so on.

Finally, social media contain a wealth of data that can be used to detect this problem. One of the most widely used data extraction services is Twitter [5, 6, 10, 12, 13, 19, 38, 41, 45], which has a freely available API that facilitates the retrieval of tweets using queries and keywords, without using web scraping techniques.

**Data Labeling.** Due to the complexity of the problem of human trafficking, it is very complicated to find publicly labeled data since it may contain sensitive information and data of persons who may or may not be involved, such as phone numbers, addresses, names, etc. Therefore, to request this data, it is usually necessary for the author to explain that it is for research and explain the area and what you plan to do with the data.

Manual data labeling relies on hand-checking and assigning a label to each piece of data, whether or not it is related to human trafficking. This process is highly dependent on the individual’s judgment, and there may be bias in assigning the label. Therefore, this task typically requires a person with human trafficking experience to review and label the data. This process requires manual review where the amount of final data is limited. There aren’t many people with experience in human trafficking willing to go through thousands or millions of pieces of data and label it. Furthermore, not all universities, research centers, or research groups have people with the ideal characteristics to carry out this task.

In human trafficking, there is a public dataset (upon request to authors) with labeled data that is used to train ML models to detect ads indicating human trafficking. This data is called Trafficking-10k and comprises 10 thousand ads and seven different labels: Certainly not, Probably not, Weakly no, Unsafe, Weakly yes, Probably yes, Certainly yes. Several papers have used this data set in their work [18, 40, 46, 47, 49].

## 4.5 Pre-processing

Several preprocessing steps are commonly followed by researchers when text is used in order to classify it. These steps are aimed at cleaning and transforming the raw tweet data to a form suitable for further analysis. Some of the most commonly used preprocessing steps are:

- **Tokenization:** It is breaking a text into individual words or tokens. In tweets, tokenization can be challenging due to the presence of emoticons, hashtags, and mentions. Therefore, specialized tokenization techniques. The works of [5, 10, 13] tokenize the text from the tweets to better understand the model and gain performance.
- **Stopword Removal:** They are common words that do not carry much meaning, such as “the,” “a,” and “an.” Removing stopwords can reduce the dimensionality of the data and improve the efficiency of subsequent analysis steps. However, the effectiveness of stopword removal in a text has been debated in the literature, with some studies suggesting that it may harm the performance of classification models. In the case of the works [8, 9, 38] extract the data from social media platforms, so they remove stopwords less dimensionality to the model, and gain meaningful data for the model.
- **Stemming/Lemmatization:** Stemming and lemmatization are techniques for reducing words to their root form. This can help reduce the data’s sparsity and improve the accuracy of subsequent analysis steps. However, stemming

can also result in the loss of information, and lemmatization can be computationally expensive. Therefore, the choice of stemming/lemmatization technique may depend on the specific task and dataset.

- **Removing URLs, emojis, mentions, and hashtags:** Tweets often contain URLs, mentions, emojis, and hashtags, which can be irrelevant or misleading for classification tasks. Therefore, these elements are often removed before analysis. The works [12,41] extract the data from Twitter, so they come with too much poor information, such as URLs, emoticons, mentions, and hashtags, that are affecting the performance of the ML models.
- **Spell correction:** Text from social media often contains misspellings and abbreviations, making it challenging to analyze the data accurately. Therefore, spell correction techniques, such as spell-checker, can improve the accuracy of subsequent analysis steps.
- **Normalization:** Normalization refers to standardizing text data, which typically involves converting text to lowercase, removing punctuation, and replacing numbers with their word equivalents. The main goal of normalization is to reduce the variability in the text data and make it easier to process.

#### 4.6 Machine Learning Tasks

The horrible crime of human trafficking, which affects millions of individuals globally, uses social media as a significant recruiting and victimization tool. Machine learning has become a potent weapon in the fight against human trafficking on social media, giving researchers the power to sift through massive volumes of data and spot possible instances of trafficking activity. Machine learning algorithms automate and streamline the detection and examination of probable human trafficking activity. In this context, ML can handle the following tasks:

- **Text Classification:** ML algorithms are trained to automatically classify social media posts, comments, and messages as potential cases of human trafficking. For example, a study by [49] used supervised machine learning to classify online escort ads as either indicative of sex trafficking or not.
- **Entity extraction:** ML is used to extract entities related to human trafficking, such as locations, names, and phone numbers, from social media posts. This can help to identify potential victims or perpetrators of trafficking. For example, a study by [47] used machine learning to extract entities related to human trafficking from ads.
- **Network analysis:** ML techniques analyze the connections and interactions between individuals and groups involved in human trafficking on social media. For example, a study by [9] used network analysis and machine learning to identify the most influential countries in the human trafficking network.
- **Image analysis:** ML is used to analyze images and identify potential instances of human trafficking. For example, a study by [10] used machine learning to analyze online images and identify potential victims of sex trafficking.

- **Topic modeling:** ML techniques are employed to identify and analyze the topics and themes on social media posts related to human trafficking. This can help to identify patterns and trends in trafficking activity, as well as to understand the experiences and perspectives of victims and survivors. For example, a study by [6, 12] used topic modeling to analyze Twitter data related to human trafficking.
- **Sentiment analysis:** ML techniques are used to analyze the sentiment of social media posts related to human trafficking, such as whether they express positive or negative emotions. This can help to identify potential victims or perpetrators of trafficking, as well as to understand the public perception of human trafficking. For example, a study by [5] used sentiment analysis to identify behavioral patterns related to human trafficking from social media posts.
- **Predictive modeling:** ML techniques are employed to predict the likelihood of human trafficking activity based on social media data and identify potential victims and perpetrators. For example, a study by [19] used machine learning to predict Twitter bots and human trafficking activity with language-independent based on online escort ads.

## 5 Conclusions

Using machine learning techniques for assessing human trafficking in social media has shown promising results. Researchers have utilized supervised, unsupervised, and semi-supervised learning methods to analyze extensive data and datasets from social media platforms, intending to identify potential victims, traffickers, and understand the patterns and networks of trafficking activity. These methods have demonstrated high accuracy and efficiency in detecting potential cases of human trafficking and have the potential to assist law enforcement agencies in their efforts to combat this horrible crime. However, challenges remain in ensuring the ethical use of data, the sites where the data is extracted, how the corresponding label is assigned to the data, and developing models that can adapt to the dynamic and evolving nature of human trafficking networks. Nevertheless, the use of machine learning in this field has opened up new ways for understanding and combating human trafficking and holds great potential for further advancements in the future. Moving forward in this area, there is much work to do. One potential area is the development of hybrid models that combine multiple Machine Learning techniques to improve the accuracy and efficiency of trafficking assessments.

**Acknowledgment.** The authors express their gratitude to the Data Science and Analytics (DataScienceYT) group at Yachay Tech University for their assistance during the development of this work.

## References

1. Alvari, H., Shakarian, P., Snyder, J.: Semi-supervised learning for detecting human trafficking. *Secur. Inform.* **6**(1), 1–14 (2017)
2. Alvari, H., Shakarian, P., Snyder, J.K.: A non-parametric learning approach to identify online human trafficking. In: 2016 IEEE Conference on Intelligence and Security Informatics (ISI), pp. 133–138 (2016)
3. Andrews, S., Brewster, B., Day, T.: Organised crime and social media: detecting and corroborating weak signals of human trafficking online. In: Haemmerlé, O., Stapleton, G., Faron Zucker, C. (eds.) ICCS 2016. LNCS (LNAI), vol. 9717, pp. 137–150. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-40985-6\\_11](https://doi.org/10.1007/978-3-319-40985-6_11)
4. Belcastro, L., Cantini, R., Marozzo, F.: Knowledge discovery from large amounts of social media data. *Appl. Sci.* **12**(3), 1209 (2022)
5. Burbano, D., Hernandez-Alvarez, M.: Identifying human trafficking patterns online, vol. 2017-January, pp. 1–6 (2018)
6. Burbano, D., Hernández-Alvarez, M.: Illicit, hidden advertisements on Twitter. In: International Conference on eDemocracy & eGovernment (ICEDEG), pp. 317–321. IEEE (2018)
7. Celebi, M.E., Aydin, K.: *Unsupervised Learning Algorithms*, vol. 9. Springer, Cham (2016). <https://doi.org/10.1007/978-3-319-24211-8>
8. Diaz, M., Panangadan, A.: Natural language-based integration of online review datasets for identification of sex trafficking businesses. In: 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), pp. 259–264 (2020)
9. Goist, M., Chen, T.H.Y., Boylan, C.: Reconstructing and analyzing the transnational human trafficking network. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019, pp. 493–500 (2019)
10. Granizo, S., Caraguay, A., Lopez, L., Hernandez-Alvarez, M.: Detection of possible illicit messages using natural language processing and computer vision on twitter and linked websites. *IEEE Access* **8**, 44534–44546 (2020)
11. Hafiz, A.M., Hassaballah, M., Binbusayyis, A.: Formula-driven supervised learning in computer vision: a literature survey. *Appl. Sci.* **13**(2), 723 (2023)
12. Hernandez-Alvarez, M.: Detection of possible human trafficking in twitter, pp. 187–191 (2019)
13. Hernández-Álvarez, M., Granizo, S.: Detection of Human Trafficking Ads in Twitter Using Natural Language Processing and Image Processing, vol. 1213. AISC (2021)
14. James, G., Witten, D., Hastie, T., Tibshirani, R.: Unsupervised learning. In: James, G., Witten, D., Hastie, T., Tibshirani, R. (eds.) *An Introduction to Statistical Learning*. STS, pp. 497–552. Springer, New York (2021). [https://doi.org/10.1007/978-1-0716-1418-1\\_12](https://doi.org/10.1007/978-1-0716-1418-1_12)
15. Kamalov, F., Cherukuri, A.K., Sulieman, H., Thabtah, F., Hossain, A.: Machine learning applications for COVID-19: a state-of-the-art review. In: Tyagi, A.K., Abraham, A. (eds.) *Data Science for Genomics*, pp. 277–289. Academic Press (2023)
16. Kejriwal, M., Ding, J., Shao, R., Kumar, A., Szekely, P.: Flagit: a system for minimally supervised human trafficking indicator mining (2017)
17. Kleinberg, J., Ludwig, J., Mullainathan, S.: A guide to solving social problems with machine learning. *Harv. Bus. Rev.* **8**, 2 (2016)

18. Lee, M.: *Human Trafficking*. Routledge (2013)
19. Lee, M.C., et al.: Infoshield: generalizable information-theoretic human-trafficking detection. In: 2021 IEEE 37th International Conference on Data Engineering (ICDE), pp. 1116–1127 (2021)
20. Li, L., Simek, O., Lai, A., Daggett, M., Dagli, C.K., Jones, C.: Detection and characterization of human trafficking networks using unsupervised scalable text template matching. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 3111–3120 (2018)
21. Liu, C., Shangguan, Y., Yang, H., Shi, Y., Krishnamoorthi, R., Kalinli, O.: Learning a dual-mode speech recognition model via self-pruning. In: 2022 IEEE Spoken Language Technology Workshop (SLT), pp. 273–279 (2023)
22. Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., Neubig, G.: Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.* **55**(9) (2023)
23. Madhulatha, T.S.: *An overview on clustering methods* (2012)
24. Mahesh, B.: Machine learning algorithms-a review. *Int. J. Sci. Res. (IJSR)* **9**, 381–386 (2020)
25. Mazza, M., Cola, G., Tesconi, M.: Ready-to-(ab)use: from fake account trafficking to coordinated inauthentic behavior on twitter. *Online Soc. Netw. Media* **31**, 100224 (2022)
26. McClosky, D., Charniak, E., Johnson, M.: Effective self-training for parsing. In: *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, pp. 152–159 (2006)
27. Motseki, M., Mofokeng, J.: An analysis of the causes and contributing factors to human trafficking: a South African perspective. *Cogent Soc. Sci.* **8** (2022)
28. Ning, X., et al.: A review of research on co-training. *Concurr. Comput. Pract. Exp.* **E6276** (2021)
29. Okech, D., Choi, Y.J., Elkins, J., Burns, A.C.: Seventeen years of human trafficking research in social work: a review of the literature. *J. Evid.-Inf. Soc. Work* **15**(2), 103–122 (2018). PMID: 29265959
30. Pijal, W., Armijos, A., Llumiquinga, J., Lalvay, S., Allauca, S., Cuenca, E.: Spanish pre-trained catrbeto model for sentiment classification in twitter. In: 2022 Third International Conference on Information Systems and Software Technologies (ICI2ST), pp. 93–98. IEEE (2022)
31. Quelal, A., Brito, J., Lomas, M.S., Camacho, J., Andrade, A., Cuenca, E.: Identifying the political tendency of social bots in twitter using sentiment analysis: a use case of the 2021 ecuadorian general elections. In: Abad, K., Berrezueta, S. (eds.) *DSICT 2022*. CCIS, vol. 1647, pp. 184–196. Springer, Cham (2022). [https://doi.org/10.1007/978-3-031-18347-8\\_15](https://doi.org/10.1007/978-3-031-18347-8_15)
32. Ramchandani, P., Bastani, H., Wyatt, E.: Unmasking human trafficking risk in commercial sex supply chains with machine learning. *SSRN Electron. J.* (2021)
33. Reddy, G.T., et al.: Analysis of dimensionality reduction techniques on big data. *IEEE Access* **8**, 54776–54788 (2020)
34. Reynolds, M.: Teaching ai to find forced labour camps. *New Sci.* (3132), 14 (2017)
35. Rodríguez-López, S.: (De)constructing stereotypes: media representations, social perceptions, and legal responses to human trafficking. *J. Hum. Traffick.* **4**(1), 61–72 (2018)
36. Shahapure, K.R., Nicholas, C.: Cluster quality analysis using silhouette score. In: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), pp. 747–748. IEEE (2020)

37. Shelke, V., Mehta, G., Gomase, P., Bangera, T.: Searchious: locating missing people using an optimised face recognition algorithm, pp. 1550–1555 (2021)
38. Shishira, S.S., Patil, M.J.S.: Detection of illicit messages in twitter using support vector machine and VGG16. *Inf. Technol. Ind.* **9**(3), 794–804 (2021)
39. Sierra-Rodríguez, A., Arroyo-Machado, W., Barroso-Hurtado, D.: La trata de personas en twitter: finalidades, actores y temas en la escena hispanohablante. *Grupo Comunicar* **30**, 79–91 (2022)
40. Tong, E., Zadeh, A., Jones, C., Morency, L.P.: Combating human trafficking with deep multimodal models (2017)
41. Tundis, A., Jain, A., Bhatia, G., Muhlhauser, M.: Similarity analysis of criminals on social networks: an example on twitter, vol. 2019-July (2019)
42. Um, M., Rice, E., Palinkas, L., Kim, H.: Migration-related stressors and suicidal ideation in North Korean refugee women: the moderating effects of network composition. *J. Trauma. Stress* **33**, 939–949 (2020)
43. United Nations: Office on Drugs and Crime.: Human-Trafficking (2020)
44. Van Engelen, J.E., Hoos, H.H.: A survey on semi-supervised learning. *Mach. Learn.* **109**(2), 373–440 (2020)
45. Vieira, C.C., Alburez-Gutierrez, D., Nepomuceno, M., Theile, T.: Desaparecidos: characterizing the population of missing children using Twitter, pp. 185–190 (2022)
46. Wang, L., Laber, E., Saanchi, Y., Caltagirone, S.: Sex trafficking detection with ordinal regression neural networks (2019)
47. Wiriyakun, C., Kurutach, W.: Feature selection for human trafficking detection models. In: *Proceedings - 2021 IEEE/ACIS 21st International Fall Conference on Computer and Information Science, ICIS 2021-Fall*, pp. 131–135 (2021)
48. Zhao, J., Xie, X., Xu, X., Sun, S.: Multi-view learning overview: recent progress and new challenges. *Inf. Fusion* **38**, 43–54 (2017)
49. Zhu, J., Li, L., Jones, C.: Identification and detection of human trafficking using language models. In: *2019 European Intelligence and Security Informatics Conference (EISIC)*, pp. 24–31 (2019)