Chapter 6
Predicting Cyber-Trafficking Websites Using a Naive Bayes Algorithm, Logistic Regression, KNN, and SVM

Aiza Jane Sulit, Risty Acerado, Ramon Christus Tomaquin, and Roselia Morco

Abstract Researchers and software developers continue discovering the best approach to combat the rising cyber-trafficking issues. However, most studies focus only on one platform or one of the gateways of cyber-trafficking. Thus, this paper introduces the development and comparison of the Naive Bayes Algorithm, Logistic Regression, k-nearest neighbor (KNN), and Support Vector Machine (SVM) classification models to predict trafficking and non-trafficking websites. In developing the supervised classification models, 37 keywords were used to scrape suspected trafficking websites. Thirty-five (35) websites were classified as trafficker out of 63; this data was used to create the models. Upon evaluating the accuracy rates of the models, the Naive Bayes Algorithm got ninety-one percent (91%), Logistic Regression got eighty-one percent (81%), KNN got sixty-four percent (64%), and SVM got sixty-four percent (64%). Thus, Naive Bayes can predict more accurately than the other classification algorithms. The result shows that the predictive model could be an effective tool for identifying different online platforms that are used in trafficking. Once the model is integrated into an application, this will be easier and faster for law enforcement agencies to monitor human trafficking in a fast-growing cyberspace community.

A. J. Sulit · R. Acerado (✉) · R. C. Tomaquin · R. Morco
Information Systems Department Technological Institute of the Philippines, Quezon City, Philippines
e-mail: racerado.is@tip.edu.ph

6.1   Introduction

Human trafficking cases have been increasing continuously. Human traffickers target no exceptions of age, gender, and race. However, most reported victims worldwide are mostly the vulnerable ones, women and girls, and they are mostly bound for sexual exploitation [1, 2]. In the Philippines, commercial sex exploitation usually occurs near offshore gaming operations and tourist destinations [3]. But now, sex trafficking is also prevalent in cyberspace, leading to many cybersex trafficking cases. The popularization and free access to many online platforms, such as social media and video-sharing sites, have opened many opportunities to traffickers. These online activities in chat rooms, social networking sites, online ads, and many other social media sites have enabled traffickers to target more victims [4].

As cyber trafficking is unstoppable, the data related to these cases are growing. However, there is no assurance that all the data stored in the government's database are "good" data as, according to [5], the data management process is poor. The challenge of analyzing good data paves the way for developing tools for information extraction, data mining, and machine learning. Experts can use these techniques to identify patterns with pertinent information on human trafficking on the internet [6]. In addition, many terminologies are used to identify human trafficking activities. For example, the terms "pimp" or "madam" are employed to indicate the probable trafficker in the situation, a "provider," who is the same as the person being sold. The word "johns" is also a known word referring to the customers of this online trafficking business [7]. Realizing the greater challenges faced by the government and seeing the opportunities to help, many researchers were interested in analyzing the activities and strategies of traffickers. Thus, many research studies were conducted using advanced technologies and innovations such as sentiment analysis [8–10] and natural language processing [11]. Most studies just focused on one social platform, such as social media messages [12], dark web [13], website advertisements [14, 15], open internet sources [16], and Twitter posts [17].

The data quality and the learning algorithms' efficacy all play a role in determining how successful a machine-learning solution will be [18]. Therefore, researching numerous machine learning algorithms enables one to determine which algorithms may be combined to provide the most accurate predictive model for monitoring websites. Studies that present classification predictions of cyber-trafficking websites are still limited. Most studies focus only on one platform or one of the gateways of cyber-trafficking. Therefore, this paper introduces the development and comparison of the Naive Bayes Algorithm, Logistic Regression, KNN, and SVM classification models to predict trafficking and non-trafficking websites.

6.2   Review of Related Studies and Literature

This section presents reviews of literature and studies related to trafficking and analytics.

6.2.1 Review on the Human Trafficking Cases

The data for the visualizations in this section was derived from the Counter-Trafficking Data Collaborative (CTDC), the world's first global data hub on human trafficking, which publishes standardized data from anti-trafficking groups worldwide [19, 20].

Figure 6.1 shows that not only the most vulnerable in the society are being targeted by traffickers.

Human trafficking comes from different forms. Trafficking is really active in online and offline global trading. Figure 6.2 shows that sexual exploitation is the major market for human traffickers. It reflects that traffickers gain more profit in this form of trafficking.

Human trafficking is not bounded by time and space, as presented in Fig. 6.3. There are many instances around the globe. Unfortunately, the Philippines has the highest proportion of trafficking occurrences among all nations, as shown in Fig. 6.3, with 11,365 instances, followed by Ukraine, which has 7,761 instances.

Figure 6.4 indicates that the vast majority of victims did not provide any specific information about their interaction with the traffickers. Unfortunately, even the victims' relatives are the primary reason victims suffer in misery in exchange for monetary compensation.

Fig. 6.1 Pie chart distribution of gender

Fig. 6.2 Tabulation of the different forms of human trafficking

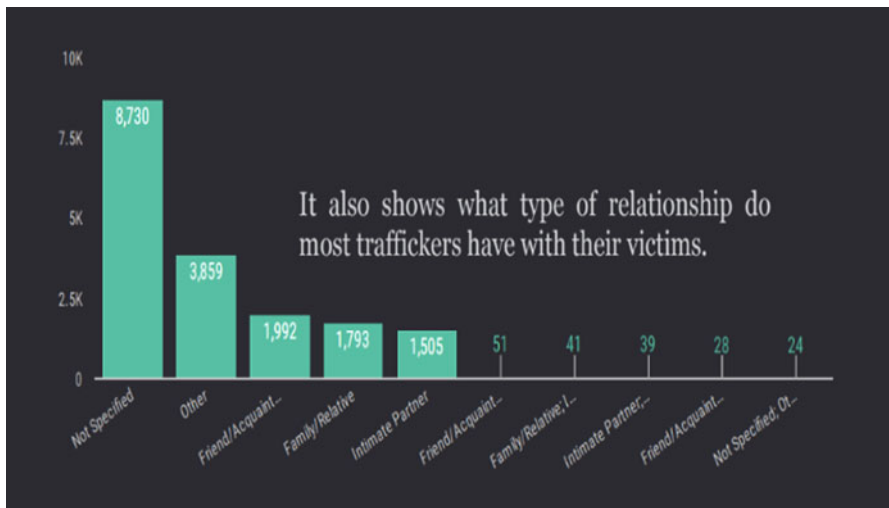**Fig. 6.3** Filled map chart for global trafficking cases

Fig. 6.4 Bar chart showing the recruitment relationship

6.2.2   Review on the Analytics to Combat Human Trafficking

Predictive analytics is the branch of advanced analytics used to predict unknown future events. This study used predictive analytics to determine a potential trafficking website using predetermined keywords. Like in the concept of Search Engine Optimization (SEO), keywords can be used to predict top search results. The study of [21] has integrated it with machine learning to predict SEO rankings. The same approach was applied in this study that a set of keywords were identified to determine a website that contains keywords most probably used by online traffickers.

Naive Bayes shows satisfactory results even in small-scale datasets [22]. Therefore, in this study, Naive Bayes was also used. In addition, Naive Bayes classifiers have better resilience to missing data than support vector machines [20]. Additionally, other machine learning algorithms, logistic regression, KNN, and SVM were also employed to evaluate their performance in predicting cyber trafficking.

6.3   Development of Cyber Trafficking Websites Classification Models

Naive Bayes classifier, Logistic Regression, KNN, and SVM were used to develop the model for classifying cyber-trafficking websites. In addition, it was used to forecast traffic on both trafficking and non-trafficking websites.

The data was extracted from scraped websites using a scraping algorithm written in the Python programming language and 37 keywords marked as red flags for traffickers. The websites were categorized into two types: trafficker and non-trafficker. The categorization process was done with the help of an expert who visited the websites with special tools. However, due to project time constraints, there were only 35 websites that were classified as trafficker out of the 63 total scraped websites. The data that was collected was formatted into a spreadsheet.

Then, to manage and analyze the data used for the analysis, the data were first cleaned by removing some of the less significant data and converting a portion of the data from text to numeric, since most of the data is string type. Encoding of string-type data is conducted so that the machine learning algorithm can execute its arithmetic operation to understand the data that must be analyzed. The algorithms performed for this study do not accept string values. Only those that are close to floating data types, which is why a label encoder is used. The data was separated into two categories: features and encodedClass. The feature category contains type and keyword data, whereas the encoded class category contains all other data. The following elaborates on the columns of the dataset:

- Url is the link to the website.
- Type indicates the type of the website or its general description such as a news article, eCommerce, educational, nonprofit, blogs, portfolio, portal, and search engines/job search engines.
- Keywords are one of the 37 keywords with the highest occurrences on the particular website.
- Keycounts corresponds to the number of occurrences of the keyword appearing on a website.
- Date is the date that the website was posted. When no date indicates the website or page that contains the keywords, the date was set to the date of the data collection.

The accuracy of models created using Naive Bayes Algorithm, Logistic Regression, KNN, and SVM was compared to determine know which is best suited for the data on hand. The algorithm examines and calculates occurrences of keywords found on trafficking or non-trafficking websites.

6.4   Evaluation of the Model

Using the train-test split library in the sklearn package and the 80/20 ratio of the training and validation data sets, the dataset is split to produce the training and testing sets. The evaluation yields two classes: class 0 for non-traffickers and class 1 for traffickers.

Figure 6.5 shows the classification report of the logistic regression model. Out of all the non-traffickers, the model predicted that 100% of them were correctly classified. Similarly, classifying traffickers also accumulated a 100% correct classification rate. The value result showed that the model does a great job of predicting whether or not it is a trafficking or non-trafficking website.

Precision and recall indicate a missing value in the accuracy column. This is because data accuracy should be as high as possible, and comparing the two models becomes difficult if the data has a low precision but a high recall, or vice versa. Therefore, to evaluate the results of the accuracy test, the F-score is used to evaluate both the precision and the recall of the data.

Fig. 6.5 Classification report for logistic regression

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3
1	1.00	1.00	1.00	4
accuracy			1.00	7
macro avg	1.00	1.00	1.00	7
weighted avg	1.00	1.00	1.00	7

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3
1	1.00	1.00	1.00	4
accuracy			1.00	7
macro avg	1.00	1.00	1.00	7
weighted avg	1.00	1.00	1.00	7

KNN

Fig. 6.6 Classification report for k-nearest neighbors

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3
1	1.00	1.00	1.00	4
accuracy			1.00	7
macro avg	1.00	1.00	1.00	7
weighted avg	1.00	1.00	1.00	7

Fig. 6.7 Classification report for Naive Bayes

|	precision	recall	f1-score	support
0	0.75	1.00	0.86	3
1	1.00	0.75	0.86	4
accuracy			0.86	7
macro avg	0.88	0.88	0.86	7
weighted avg	0.89	0.86	0.86	7

Fig. 6.8 Classification report for SVM

Figure 6.6 shows the classification report of the K-Nearest Neighbors model. The model showed a similar output to the logistic regression model with 100% correctly identified positives as well as the same for the accuracy of each positive prediction, while the support showed a balanced dataset consisting of three (3) non-traffickers and four (4) traffickers.

Figure 6.7 shows that in the same manner as with the K-nearest neighbors and logistic regression, Naive Bayes model has 100% correctly identified positives. The same goes for the accuracy of each positive prediction as well as the support.

Figure 6.8 shows the classification report of the SVM model. The model predicted that out of all the non-traffickers, 75% of them were correctly classified. On the other hand, classifying traffickers accumulated a 100% correct classification

Fig. 6.9 Confusion matrix
for logistic regression,
K-nearest neighbors (KNN),
and Naive Bayes
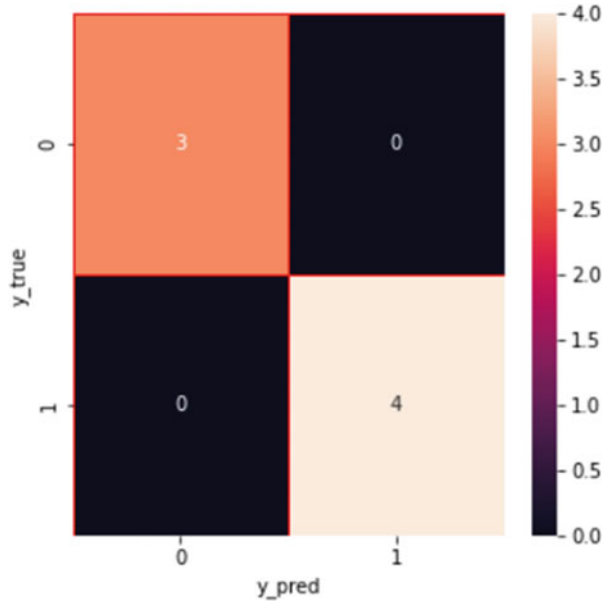
rate. The F1 score value appears to be near 1, indicating that the model is good at
predicting both non-traffickers and traffickers while also maintaining support at an
acceptable range.

Figure 6.9 shows the confusion matrix for logistic regression, SVM, and Naive
Bayes. These three models appeared to share a classification report, resulting in the
same confusion matrix. The confusion matrix shows no misclassifications that
appear in the model. Instead, the confusion matrix demonstrates that the classifica-
tion report's three non-trafficker websites and four trafficker websites appeared to be
correctly classified. There are seven correctly classified websites for these models.

A confusion matrix is utilized because it allows visitors to examine the outcomes
of an algorithm at a glance. The confusion matrix presents analytical findings in the
form of a straightforward table, which effectively condenses the outputs into a
perspective that is easier to understand.

As shown in Fig. 6.10, there are six misclassifications. Three on the
non-trafficking website and three on the trafficking website. The correctly classified
website is only one (1).

Figure 6.11 shows the result of the cross-validations for all the models presented.
The KFold validation was used in determining the accuracy of the data in all the
models. This shows that although three of the models have the same classification
report, they still differ when it comes to the accuracy of their data. Cross-validation is
conducted to get more information about the algorithm performance. The validation
showed that Naive Bayes appears to be the highest among all the models.

Fig. 6.10 Confusion matrix for SVM
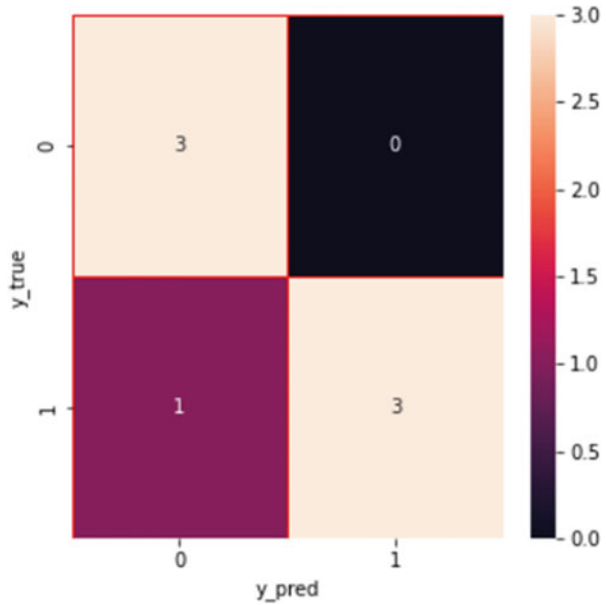
Fig. 6.11 Cross-validation results

```
Logistic Regression:  0.8181818181818182
SVM:   0.6363636363636364
KNN:   0.6363636363636364
Naive Bayes:  0.9090909090909091
```

6.5 Conclusion

The presented predictive algorithm using four machine learning algorithms, namely, the Naive Bayes Algorithm, Logistic Regression, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM), was able to predict trafficking and non-trafficking websites. The evaluation shows that the model performs well, having an accuracy result of 91%, 81%, 64%, and 64%, respectively, in classifying trafficking and non-trafficking websites. Furthermore, as observed from the classification report, some machine learning algorithms reflect the same value, because it reflects the measurement's proximity to the actual value, but they are not identical. Accuracy indicates how close a measurement is to a known or accepted value, regardless of how far it deviates from the accepted value. Both precise and accurate measurements are repeatable and close to the true values.

Moreover, Naive Bayes stands out with the highest accuracy rate, making it an ideal model to be used, also taking into consideration the results from SVM, which is acceptable enough to be used along with the Naive Bayes, which could garner a more effective outcome for predictions. The model can be integrated into a tool to help law enforcement agencies to analyze transactions on the web and identify

possible cyber traffickers. It is an excellent line of defense to act on early signs of human trafficking before it has taken a victim. In this way, law enforcement can develop a task force that will watch over red flags identified by the analytical models presented. Moreover, scrutinizing each online transaction would be much easier for law enforcement agencies, allowing them to act on the concerns of a larger community in need of their assistance. It shows that the model can be helpful in proactively combating traffickers that are roaming around in cyberspace.

References

1. Stop The Traffik. Definition and Scale.Learn more about human trafficking and how prevalent this fast-growing crime is. 2022. Retrieved in February, 2023. https://www.stopthetraffik.org/what-is-human-trafficking/definition-and-scale/
2. United Nations Office on Drugs and Crime (UNODC). Global report on trafficking in persons. 2009. Retrieved in February 2023. https://www.unodc.org/unodc/en/press/releases/2023/January/global-report-on-trafficking-in-persons-2022.html
3. U.S. Department of State. Trafficking in persons report: Philippines. 2022. Retrieved in February 2023. https://www.state.gov/reports/2022-trafficking-in-persons-report/philippines__trashed/
4. M. Latonero. Human trafficking online: the role of social networking sites and online classifieds. Center on Communication Leadership & Policy. Research Series: September 2011. USC. Annenberg School for communication & Journalism https://doi.org/10.2139/ssrn.2045851
5. J. Brunner. Getting to good human trafficking data. Assessing the landscape in Southeast Asia and promising practices from ASEAN governments and civil society. 2018. Retrieved in February 2023. https://humanrights.stanford.edu/sites/humanrights/files/good_ht_data_policy_report_final.pdf
6. D. Burbano, M. Hernandez-Alvarez. Identifying human trafficking patterns online, in IEEE Second Ecuador Technical Chapters Meeting (ETCM) (2017), pp. 1–6, doi: https://doi.org/10.1109/ETCM.2017.8247461
7. M. Ibanez, D. D. Suthers. Detection of domestic human trafficking indicators and movement trends using content available on open internet sources, in 47th Hawaii International Conference on System Sciences, pp. 1556–1565, 1, 2014
8. K.S.A. Rose, Application of sentiment analysis in web data analytics. Int. Res. J. Eng. Technol. **07**(06), 2395 (2020)
9. D. Liu, C.Y. Suen, O. Ormandjieva. A novel way of identifying cyber predators. ArXiv, abs/1712.03903 (2017)
10. A. Mensikova, C.A. Mattmann Ensemble sentiment analysis to identify human trafficking in web data, in Workshop on Graph Techniques for Adversarial Activity Analytics (GTA). Marina Del Rey, 2018, pp. 5–9
11. Maria Diaz, Anand V. Panangadan. Natural language-based integration of online review datasets for identification of sex trafficking businesses, in IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI) (2020), pp. 259–264
12. W. Chung, E. Mustaine, D. Zeng. Criminal intelligence surveillance and monitoring on social media: Cases of cyber-trafficking, in 2017 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE Press (2017), pp. 191–193. https://doi.org/10.1109/ISI.2017.8004908
13. C.A. Murty, P.H. Rughani, Sentiment & pattern analysis for identifying nature of the content hosted in the dark web. Indian J. Comput Sci Eng **12**(6) (2021). https://doi.org/10.21817/indjcse/2021/v12i6/211206142

14. A. Volodko, E. Cockbain, B. Kleinberg, "Spotting the signs" of trafficking recruitment online: Exploring the characteristics of advertisements targeted at migrant job-seeker. Trends Organ Crim **23**, 7–35 (2020). https://doi.org/10.1007/s12117-019-09376-5
15. M. Latonero. Human trafficking online: The role of social networking sites and online classifieds. Center on Communication Leadership & Policy. Research Series: September 2011. USC. Annenberg School for communication & Journalism https://doi.org/10.2139/ssrn.2045851
16. H. Wang, C. Cai, A. Philpot, M. Latonero, E.H. Hovy, D. Metzler. Data integration from open internet sources to combat sex trafficking of minors, in Proceedings of the 13th Annual International Conference on Digital Government Research (2012), pp. 246–252. https://doi.org/10.1145/2307729.2307769
17. M.H. Alvarez, D.B. Acuña. (2017). Identifying human trafficking patterns online, in Conference: 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), Salinas Ecuador (2017). DOI:https://doi.org/10.1109/ETCM.2017.8247461
18. I.H. Sarker. *Machine Learning: Algorithms, Real-World Applications and Research Directions – SN Computer Science*. SpringerLink. (2021). Retrieved May 29, 2022 from https://link.springer.com/article/10.1007/s42979-021-00592-x
19. Counter-Trafficking Data Collaborative. https://respect.international/counter-trafficking-data-collaborative/
20. Hongbo Shi, Yaqin Liu. Naïve Bayes vs. Support vector machine: resilience to missing data. Lecture Notes in Computer Science (2011), p. 8 https://doi.org/10.1007/978-3-642-23887-1_86
21. Michael Weber. Machine learning for SEO – How to predict rankings with machine learning (October 26, 2017). Retrieved June 2021 from https://www.searchviu.com/en/machine-learning-seo predicting-rankings/
22. Yuguang Huang, Lei Li. Naive Bayes classification algorithm based on a small sample set, in IEEE International Conference on CloudComputing and Intelligence Systems (September 2011), p. 6. DOI:https://doi.org/10.1109/CCIS.2011.6045027