ISyE 6740 – Summer 2020
Final Report

Group Number: 83

Team Member Names: Caroline Coyle (903186626), Katrina Green (903369888), Anika Rahman (902909455)

Project Title: Modern Day Slavery and Machine Learning: Utilzing Supervised Learning to Identify Human Trafficking In Online Advertisements

Team Member Responsibilities:
Our team utilized an iterative model with weekly team check-ins in order to execute and complete our research. During our weekly team calls, we reviewed the work completed by each team member the week prior, troubleshooted challenges, and peer programmed our code to improve readability, efficiency, and to help each other grow programmatically. Every team member contributed substantially to the development of the problem statement, data source, methodology, evaluation, and final results.

Team Member	Key Items Contributed
Caroline Coyle	● Research and identification of key data sources and identification of feature variables to be included in models ● Analysis of country and continent of origin ● Analysis of weight variables ● Development of data graphs ● Research paper structure, review, and editing
Katrina Green	● Analysis of massage and URL variables ● Analysis of Transient Language and Unconventional Sex phrases ● Created dummy variables for non integer/float variables for modeling ● Implemented the following machine learning models: K-NN, SVM, Gaussian Naive Bayes, Adaboost, Random Forest ● Applied modeling on the YesBackpage data
Anika Rahman	● Analysis of Organized and Minor Indicator phrases ● Research on top cities known for Human Trafficking ● Scraped testing data from YesBackpage using the BeautifulSoup package ● Dataset reduction ● Implemented the Neural Network Model

Modern Day Slavery and Machine Learning:

Utilizing Supervised Learning to Identify Human Trafficking In Online Advertisements

Introduction

Human trafficking is often referred to as modern day slavery. Every year, thousands of men, women, and children are coerced, taken, or sold into slavery at the hands of human traffickers. Human trafficking consists of three main components: exploitative labor, harboring of victims, and coercion.[1] Victims are often deceived with promises of working abroad, promises of marriage, or other fraudulent means. Often, they are then smuggled across international borders to countries where they don't have any connections, stripped of their travel documents, and forced to work as domestics, sex workers, farm laborers, factory workers, and more. Coerced drug use, threats, fraud, and violence are an inextricable part of the human trafficking supply chain making it difficult for victims to seek help or try to escape.[2]

In attempting to address how to disrupt the human trafficking supply chain, it is important to be cognizant of the difficulties in collecting information. Since human trafficking is clandestine in nature, the availability of data is next to none. As mentioned by Renata, et al in "Human Trafficking Analysis", "there is a dearth of quantitative data regarding trafficking activity … data that [does] exist can be inaccurate, missing or worse: false". Additionally, since human trafficking is frequently an international crime, data collection is hindered by the need to collect information from multiple sources and jurisdictions, many of which may not have the capacity to collect sufficient data themselves. Another challenge arises in the face that victims may not want to share information about what is likely a traumatic experience even though data on victims who have escaped is likely more available than data on perpetrators.[3] This poses significant challenges when attempting to develop quantitative methodologies that can be used to identify, predict, and stop human trafficking operations.

Human trafficking is a multi-faceted problem with many sub-problems to address. In this analysis, we will focus on a key issue pertinent to ending human trafficking: identifying sex trafficking in online advertisements. In the digital age, human trafficking has also moved much of its work from neighborhood signs to online advertisements on backdoor internet pages such as Backpage.[4] where thousands of sex ads are posted every single day. Use of these sites allows human traffickers to advertise their victims alongside at-will sex workers, disguising the exploitation of their victims as sex work voluntarily engaged in of the victim's own volition. Identifying and distinguishing advertisements that involve human trafficking from advertisements that are posted by at-will sex workers is a crucial aspect of interrupting human trafficking operations. In this analysis, we aim to develop a machine learning model that can identify online sex advertisements with high likelihood of being related to human trafficking in order to identify potential victims and provide law enforcement with a focused group of ads for investigation.

A Note on Terminology and Data

In order ensure that we respect the rights of individuals to engage in sex work and given that we do not have conclusive evidence of human trafficking, we will distinguish between three separate groups of

---

[1]Renata, et al. "Human Trafficking Analysis." *OR/MS Today*, vol. 44, no. 2, Apr. 2017, https://www.informs.org/.
[2] "What Is Human Trafficking?" *United Nations: Office on Drugs and Crime*, www.unodc.org/unodc/en/human-trafficking/what-is-human-trafficking.html#What_is_Human_Trafficking.
[3] Renata, et al. "Human Trafficking Analysis." *OR/MS Today*, vol. 44, no. 2, Apr. 2017, https://www.informs.org/.
[4] Backpage is no longer operational as it was seized by the US government in April 2018. Though no one site has effectively replaced it, human traffickers are still using other similar sites to put out advertisements.

individuals: posters, sex workers, and victims. We will refer to the person posting the advertisement as the "poster", this may be the individual in the ad or, in the case of human trafficking, it may be a handler posting for the victim. Individuals who do willingly and legally engage in sex work of their own volition will be referred to as "sex workers". Finally, where discussing the realities, indications, and risks of human trafficking, we will refer to those individuals who may have been trafficked as "victims".

Since the Trafficking-10K and YesBackpage datasets include postings that involve real people, their lives, and, in many cases, their livelihoods, in an effort to protect their privacy, we have removed or blacked out individual's names, email addresses, phone numbers, and photos from our report and have not included direct links to any of the advertisements; these will be evident in all examples as black boxes.

Difficulties of Gathering Human Trafficking Data

The traumatic nature of human trafficking can make victims hesitant to share personal information. This means that there is very little data available on human trafficking victims and even less so on perpetrators. In an attempt to combat this data deficiency, researchers at Carnegie Mellon University compiled the Trafficking-10K dataset, which is comprised of approximately 12,000 sex work advertisements with text and images, labelled with their likelihood (on a scale from 0-6) of having been posted by human traffickers.[5] Sample advertisements with probabilities of 0 (Strongly Likely Not Trafficking), 3 (Some Possibility of Human Trafficking), and 6 (Strongly Likely Human Trafficking) in the Trafficking-10K dataset have been included in *Table 1* and was the starting point for feature engineering and model building for our machine learning models.

Table 1

Rank	**Title**	**Body**
0	🇺🇸🇺🇸🇺🇸🇺🇸❤ GO RED NOT TED❤🇺🇸🇺🇸🇺🇸🇺🇸🇺🇸 TEXT ME - 30	Openings today 3pm-1am:) Hello there my name is [redacted] :) I am a sweet girl that is INDEPENDENT. I am discrete friendly and can hold a good conversation. My sessions are thorough and never rushed. *Sensual Session I don't post anywhere else. I am independent and all pictures of me are real & recent. INCALL/OUTCALL ➡ 4.5-5 ⭑hotels Text first for appointment **NOTE I block ppl that act RIDICULOUS VIA TEXT. 😉 INCALL/OUTCALL ID: [redacted] Text me ....:)
3	Gεn†εℓmεn's Cho¡cε 🍯WaRm & §ti©Ky🍯👅💦🙈 FuN😘👯🐵 ↘eXotiicc✓💖❀💕 uℓtiMatε ℰ𝓍𝓅erĭeŋ¢e💖❀💕 - 24	GENTLEMEN💋 I'm Not What Most Call The Typical Girl... The Complete Package Is Right here 💎Sexy Sensual SWEETHEART. 💎No rush!! 💎Down-To-Earth And Aiming To Please At All Times!! You've Tried Average && Ordinary. Now Experience What The Best Is My amazing certified skills will leave you craving more [redacted]
6	👉🔥NeW iN TOWN💗ToP NoTcH InDiaN GoddESs👅👄👈Soft N Thick🍒🍬Flawless 😍✌FrEak 🍸 💋 - 21	PRETTY FACE 👸PeTiTe TIGhT lil bOdy ❥ & a FUN loving Open Minded Attitude! a Truly TOP NOTCH provider!!! My looks are all natural and absolutely delicious!!😯Trust me you will love your time with me!! 💯%Clean&Discrete✔️ 💯%Real&Independent💰💯%Satisfaction EveryTime👅 No rush & NO FUSS! ((I love what I do!!)) 🍒RESPECTFUL GENTS ONLY🍒 ☑SERIOUS INQUIRIES 📵NO

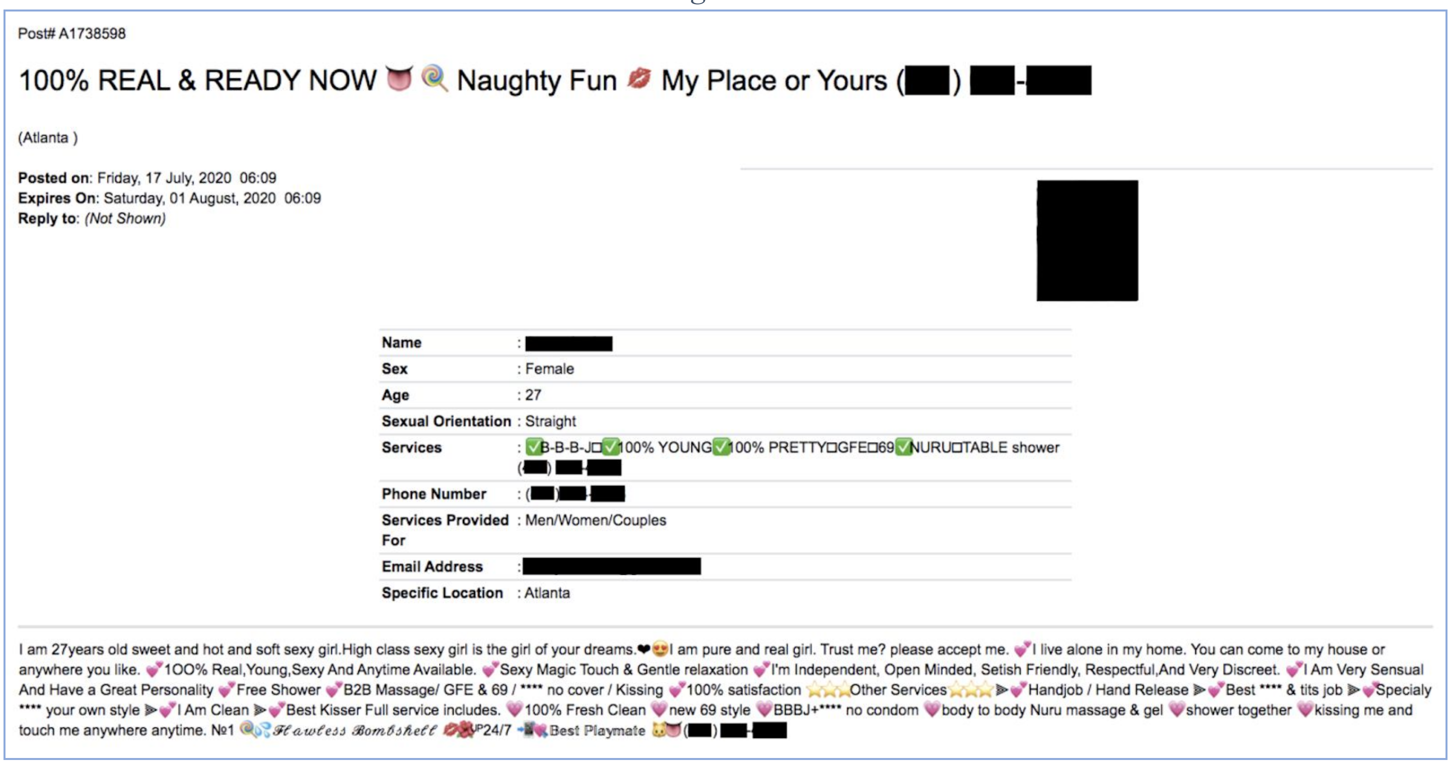
[5] E. Tong, A. Zadeh, C. Jones, and L-P. Morency. "Combating Human Trafficking with Deep Multimodal Models," Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1 Long papers), pp 1547-1556, 2017. [arxiv.org/pdf/1705.02735.pdf]

		BLOCK CALLS 🚫NO DRAMA PLEASE 🚫NO EXCESSIVE TEXTING 🌴420/FRIENDLY 📱 ......▌

In addition to the Trafficking-10K dataset, we also scraped our own data from *YesBackpage.com*, a website alternative to *Backpage.com* that rose up after the original Backpage which was seized by the law enforcement in 2018 for high volumes of crime, including human trafficking. [6]

Utilizing the National Human Trafficking Hotline's report on the top US cities for human trafficking, we scraped the Adult Services section YesBackpage for each of the 10 ten worst cities for human trafficking using BeautifulSoup.[7] Since many of the posts were duplicated or posted multiple times, we scraped the most recent 10 pages for each city as of July 6, 2020. This returned over 10,375 scraped ads. We then removed the duplicate ads which reduced the number of ads in our YesBackpage dataset down to 3,465 ads. A sample ad from YesBackpage is shown in *Figure 1*.

Figure 1


Post# A1738598

100% REAL & READY NOW 👅🍭 Naughty Fun 💋 My Place or Yours ([redacted]) [redacted]-[redacted]

(Atlanta)

Posted on: Friday, 17 July, 2020 06:09
Expires On: Saturday, 01 August, 2020 06:09
Reply to: *(Not Shown)*

[redacted]

Name	: [redacted]
Sex	: Female
Age	: 27
Sexual Orientation	: Straight
Services	: ✅B-B-B-J□✅100% YOUNG✅100% PRETTY□GFE□69✅NURU□TABLE shower ([redacted]) [redacted]-[redacted]
Phone Number	: ([redacted]) [redacted]-[redacted]
Services Provided For	: Men/Women/Couples
Email Address	: [redacted]
Specific Location	: Atlanta

I am 27years old sweet and hot and soft sexy girl.High class sexy girl is the girl of your dreams.❤😍I am pure and real girl. Trust me? please accept me. 💕I live alone in my home. You can come to my house or anywhere you like. 💕1OO% Real,Young,Sexy And Anytime Available. 💕Sexy Magic Touch & Gentle relaxation 💕I'm Independent, Open Minded, Setish Friendly, Respectful,And Very Discreet. 💕I Am Very Sensual And Have a Great Personality 💕Free Shower 💕B2B Massage/ GFE & 69 / **** no cover / Kissing 💕100% satisfaction ⭐⭐⭐Other Services⭐⭐⭐⫸💕Handjob / Hand Release ⫸💕Best **** & tits job ⫸💕Specialy **** your own style ⫸💕I Am Clean ⫸💕Best Kisser Full service includes. 💜100% Fresh Clean 💜new 69 style 💜BBBJ+**** no condom 💜body to body Nuru massage & gel 💜shower together 💜kissing me and touch me anywhere anytime. №1 🍭💦Flawless Bombshell 💋🌺24/7 ➡📲💘Best Playmate 🐱👅([redacted]) [redacted]-[redacted]

In *Table 2*, we have included the raw format of the Yesbackpage dataset which was our starting point for feature engineering. Similar to the Trafficking-10K dataset, each YesBackpage ad has an ad title, body text, and photo (intentionally excluded here). The "age" feature was simple to extract from YesBackpage since it was already a populated field that the website requires from users on each YesBackpage ad. As shown in *Table 2*, emojis and Leetspeek text are very common on both datasets. YesBackpage does censor words considered in

[6] United States, Congress, Office of Public Affairs. *Www.justice.gov*, Department of Justice, 9 Apr. 2018. www.justice.gov/opa/pr/justice-department-leads-effort-seize-backpagecom-internet-s-leading-forum-prostitution-ads.
[7] "100 Most Populous US Cities." *Www.humantraffickinghotline.org*, National Human Trafficking Hotline, 2017, humantraffickinghotline.org/sites/default/files/100%20Most%20Populous%20Cities%20Report.pdf.

typical American culture to be profane; these words are replaced with asterisks, as seen in the Houston ad in *Table 2*.

Table 2

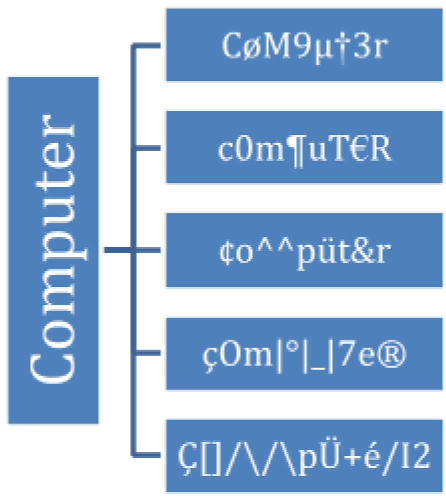
City	Title	Text	Age
Houston, TX	I'm Available for both Incall and outcall [redacted] 💝🍆🍆 [redacted]	Hey My Name is [redacted] 🍬🍭😋 Age: 29💖 indulge in the perfect Package ✨🍭 that I guarantee you'll love come back running for more 💋💋 ✨ ✔ Real 💯✔ Sexy 💯✔ Discreet 💯💋 Available For Men who desire the best in Companionships !! Serious Inquiries only 💋 Playful 🌟 5 star enjoyment 💯 Professional Safe 💯 💕 Beautiful Smooth Body A Mesmerizing face 💕 Ill leave you satisfied from the moment we meet 💋💕✈ CATCH ME WHILE YOU CAN ✈ ✔ ✔ ✨ 100% independent 💎 ✨ CleanClassy 🎀👑 ✨ 100% Real Pictures 📸🎞👀 ✨ Great Hygiene 😛🍭 DONT HESITATE TO CALL or Text [redacted] Services: Incall, Outcall, Car Date, 2some, 69, ****, FWB, ****, and lots more. Am submissive so you can try style with me. I can travel any where in Houston	30
Los Angeles, CA	💋Sⓦⓔⓔⓣ 🐝❥ Lⓘⓚⓔ 🐝❥ HⓄ NⓔY • 🍯★ NŝŧY *LIKE YOU ωANŧ ME TO	✔💯 %Independent 💕💕 ✔💯%Real & Recent Pictures Need a Break away from your regular routine, indulge in something Extraordinary & Addicting give me a call you are my main focus during our time together. I would love to be your greatest distraction. Explicit talk will get BLOCKED ! Call or Text and ask for HONNEY 🍯💛 NO CHEAP MEN 🚫 QUALITY IS IMPORTANT !! YOUR DEALING WITH THE BEST ⭑	21
Atlanta, GA	▃▅💦 Private Nude 🅱🅾□□ □🅰□□🅰□□ 💦▅▃💰 Cash Only💰	Hi Everybody ❖ We have Chinese ,Korean Girls. ❖ Age No Matter ❖No C-ondom. ❖Young Sexy Hot Beautiful Sweet Nice ❖Safe have fun, open your mind, You will enjoy Session, you will be satisfied. ❖ Happy Massage/ Nude Body slide/ More style ❖Private Clean Comfortable Rooms, Soft Music. ❖ FOR OUR LOCATION & MORE INFORMATION MAIL US❖ 🍇 Open 7days/ Week, ♣OPEN 24 hours ♣ 🍇 No link ...❌ .No Credit Card ❌ ♣ Only Cash ♣ 🍇 First mail me here,then i will give you my phone number with in 2 minutes I call you.OK 🍇 🍇 Need privacy - please first contract My E-mail- [redacted]@gmail.com » 🍇 To prove r u real please type "Okay" subject line. 🍇 No law or affiliation!No games No busy ♣ Thank You	28

In our analysis, we utilize the Trafficking-10K dataset to build supervised machine learning models to identify online ads that are likely incidents of human trafficking. We then use our models on the YesBackpage ads to predict the likelihood that these ads are human trafficking advertisements.

Challenges in Identifying Human Trafficking Advertisements

One major issue worth addressing that could arise while utilizing the machine learning models to identify trafficking ads is the use of "leet" or "leetspeak" and emojis. Leetspeak is a method of replacing alphabetic characters with symbols that, to humans, read the same as a word in the English dictionary, but utilizes different characters as visualized in *Figure 2*.

Figure 2



Leetspeak and emojis are commonly used in advertisements for sex work and human traffickers utilize similar structures in order to prevent law enforcement and from identifying ads selling sex online. By using a high volume of leetspeak and emojis, which can confuse trafficking classification models, traffickers are able to evade detection since simple language recognition programs will not account for the variability in replacing alphabetic characters with symbols. The widespread use of leetspeak would make it more difficult for the model to correctly classify ads as the model would likely not interpret the word the way it would be read by a human or could disregard it altogether as relevant when classifying.

In order to address the use of emojis and Unicode characters in an ad, before running an ad through the Machine Learning model, it could first be filtered through a deep learning model that has been trained to identify leetspeak characters that are commonly used to replace alphabetical characters and replace them with an English language counterpart. Consider the characters in *Table 3* and their English language counterparts:

Table 3

Unicode	Emoji Text Translation	Visual Representation
ðŸ'¶ðŸ¾	B_baby4	·
ðŸ˜	A_loveeyes	😍

While the “B” in the first box is quite clear to someone reading it, a computer would have a hard time classifying it as such based on the Unicode it is receiving. Before a machine learning model can be applied, these issues need to be addressed to ensure the accurate classification of advertisements.

<u>Data Preparation and Feature Engineering</u>

Despite having a labeled dataset, there was extensive data cleaning and preparation that came into play before modeling could begin. As seen in *Table 1* and *Table 2*, both sets consist of text-based entries; which meant in order to have features to train on, all data had to be extracted from the text. Given the nature of the data, we operate under the assumption that outliers do not exist for the purposes of the model. If they did, they would be extremely difficult to identify and distinguish from the rest of the data. There were however duplicates that existed in the YesBackpage dataset. Where we were able to locate these instances, we removed exact duplicates. There were some ads that were partial duplicates where all information but the poster’s age and city, these ads were left in as we felt that a mostly duplicate ad in multiple cities could be indication of a human trafficking ring.

Other challenges that arose with the data included the aforementioned “leetspeak” and other confounding separation characters. The confounded characters were more common in the data scraped from YesBackpage as it included some html break characters, such as “
” that had to be removed from the text entries.

In preparation for feature engineering and extraction, we conducted an extensive review of published articles, papers, and books on the topic of human trafficking, machine learning, and building data-based models that can help authorities identify and disband human trafficking operations. Our primary focus in the review was to analyze which features other researchers had utilized (and why) in building their models so that we could better focus our features to reflect and build on prior research. A full review of all features selected for model inclusion, their data types, feature engineering strategy, and justification for selection are included below. We have included below an overview of the features included in our models and our justification for inclusion.

Age

The majority of the ads in the Trafficking-10K dataset and the YesBackpage dataset listed the Age of the poster. Although these ages might not necessarily be true, we extracted age as a feature to see if there was a relationship between age and likelihood of trafficking. The age of the poster was explicitly listed on the web scraped data, whereas it was embedded in the Title text on the Trafficking-10K dataset. Some entries in the Trafficking-10K dataset did not include age and we dropped these entries from the dataset.

Weight [8]

Two features were extracted using the weight of the poster. The first feature is the weight value itself, since many ads listed the weight of the poster. The second feature is a Binary value (0 or 1). If the weight value is less than 115 lbs, then the binary value is 1 otherwise it is a 0. The Centers for Disease Control states in a National Health Statistics Report that the average weight for a woman in America in her twenties is 167.5 lbs.[9] Ads with

[8] Alvari, Hamidreza, et al. “Semi-Supervised Learning for Detecting Human Trafficking.” *Security Informatics*, vol. 6, no. 1, 11 May 2017, doi:10.1186/s13388-017-0029-8.
[9] Fryar, Cheryl D., et al. “Mean Body Weight, Height, Waist Circumference, and Body Mass Index Among Adults: United States, 1999–2000 Through 2015–2016.” *National Health Statistics Reports*, no. 122, 20 December 2018.. Hyattsville, MD: National Center for Health Statistics. https://www.cdc.gov/nchs/data/nhsr/nhsr122-508.pdf.

lower weights, under 115 lbs, could indicate underage victims, and therefore, a high indication of human trafficking. In the YesBackpage dataset, the weight is not as commonly found as age.

The distributions of age and weight from the Trafficking 10-K and YesBackpage, respectively, are shown below in *Figure 3* and *Figure 4*. In the Trafficking-10K dataset, we can clearly see that younger individuals and lower weights dominate the dataset. In the YesBackpage dataset, younger ages still dominate the ads, however, far fewer posters include their weight.

Figure 3

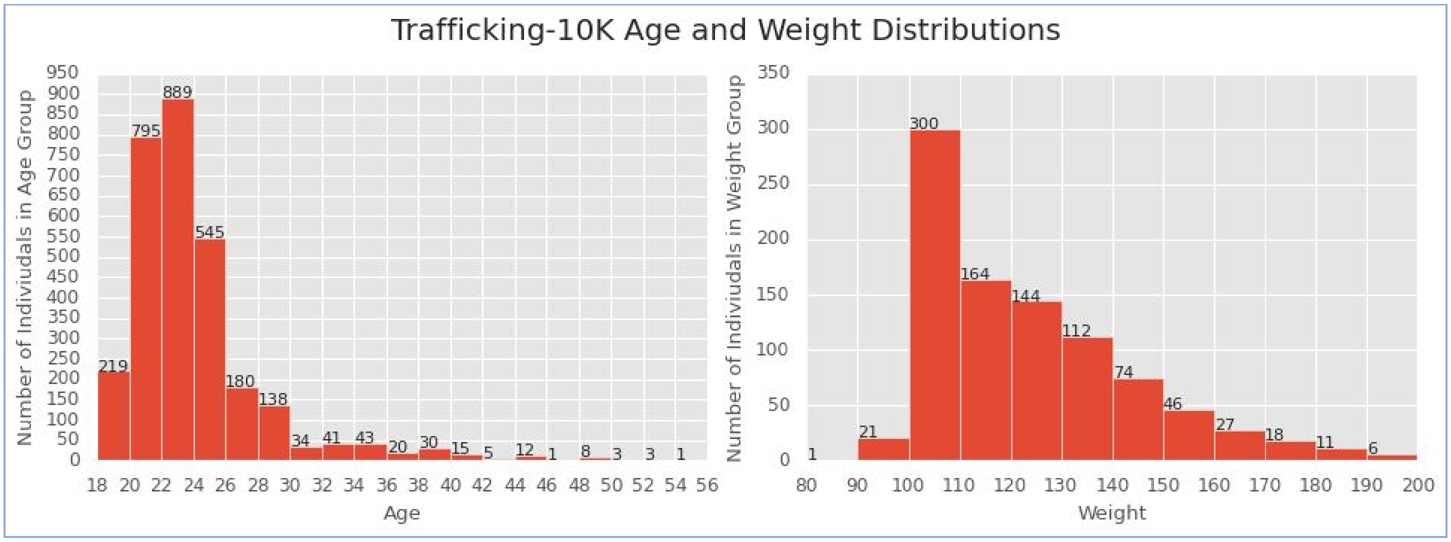
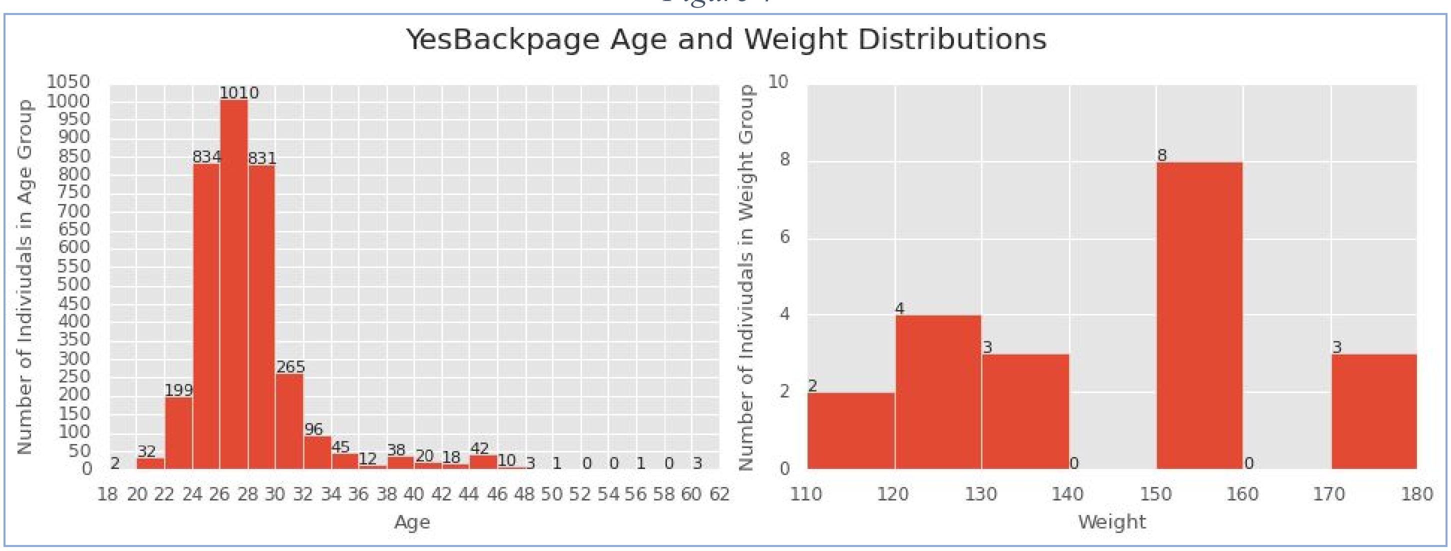


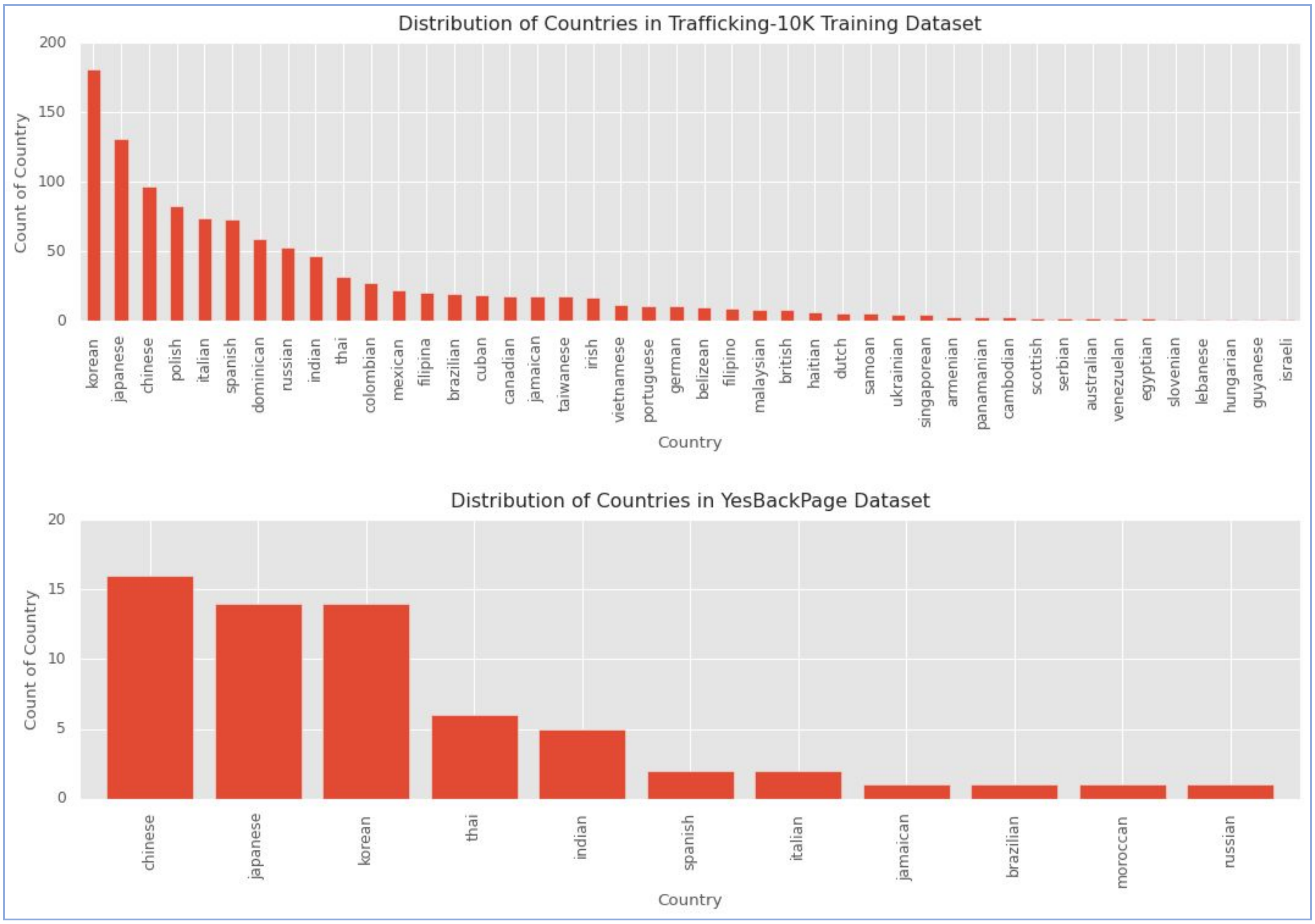
Figure 4



***Country*[10]**

Identified any mention of Countries or Nationalities within the advertisements. Particularly, references to countries in Southeast Asia can indicate human trafficking as many victims that originate in this region. The distribution of references to Countries in the datasets are shown in *Figure 5*.

Figure 5



Continent

In many cases, an ad may not include an individual's country of origin, but may include a more general indicator, such as their race or continent. For example, an ad may use "Asian" rather than "Korean." In order to capture this as a place of origin, in addition to the Country feature, we identified any mention of Continents within the advertisements.

[10] Alvari, Hamidreza, et al. "Semi-Supervised Learning for Detecting Human Trafficking." *Security Informatics*, vol. 6, no. 1, 11 May 2017, doi:10.1186/s13388-017-0029-8.

Massage[11]
References to massage and spa therapy in the arena of sex work often indicates organized crime. The importance of this feature is particularly simple to grasp when we consider the "massage parlors" references in pop culture and the news and how they are often heavily associated illegal sex work, often performed by women from Asian countries. We compare each ad to a list of massage related words, and count how many references there are.

Minor Indicators[12]
The ads don't mention children explicitly due to laws against pedophilia, but many ads use words, phrases, or emojis associated with children such as 'sweet', 'candy', 'new girl', and 'fresh'. We scan the ads to count the number of times these phrases are used within the body and title text. Ads with high counts indicate a high likelihood of underage human trafficking or young victims.

Organized Work[13]
Ads may be written in plural first person using words such as 'we' or 'our'. Ads may also mention multiple individuals at the same time such as '2some' or 'girls'. These words and phrases can often indicate that the individual is part of a larger sex work operation and, given that most sex workers operate independently and not with other individuals, it can often be indicative of organized human trafficking.

Unconventional Sex[14]
While sexual experimentation with a known partner can be safe and fun, language indicating sexual fetishes or BDSM (bondage, dominance, sadism, and masochism) can often be an indication of human trafficking as many sex workers will not engage in potentially dangerous sexual activity with partners they are not familiar with. With this in mind, individuals who seek out kink and fetish activity from sex workers may do so to fulfill a desire to abuse or harm the sex worker. We count the number of times these words are used within the body and title text.

Transient Language[15]
Transient language indicates availability for a short time. Phrases such as 'just arrived', 'in town for the weekend', and 'catch me while you can' are common within the ads indicative of human trafficking. Human traffickers move their victims between cities in order to keep them from making connections that could lead to their rescue. This strategy is especially effective on children as it ensures that the trafficker is the only person that they have a relationship with and that they view as "providing" for them.

In order to ensure we built a robust model founded on research, we addressed each of the features above by performing an extensive analysis of the text data in the advertisements in order to identify and extract the

---

[11] Alvari, Hamidreza, et al. "Semi-Supervised Learning for Detecting Human Trafficking." *Security Informatics*, vol. 6, no. 1, 11 May 2017, doi:10.1186/s13388-017-0029-8.
[12] Alvari, Hamidreza, et al. "Semi-Supervised Learning for Detecting Human Trafficking." *Security Informatics*, vol. 6, no. 1, 11 May 2017, doi:10.1186/s13388-017-0029-8.
[13] Alvari, Hamidreza, et al. "Semi-Supervised Learning for Detecting Human Trafficking." *Security Informatics*, vol. 6, no. 1, 11 May 2017, doi:10.1186/s13388-017-0029-8.
[14] Whitney, Jessica, et al. "Identifying Victims of Human Sex Trafficking in Online Ads." *Encyclopedia of Criminal Activities and the Deep Web*, edited by MEHDI KHOSROW-POUR, vol. 1, IGI GLOBAL, 2020, pp. 497–506.
[15] Whitney, Jessica, et al. "Identifying Victims of Human Sex Trafficking in Online Ads." *Encyclopedia of Criminal Activities and the Deep Web*, edited by MEHDI KHOSROW-POUR, vol. 1, IGI GLOBAL, 2020, pp. 497–506.

features and prepare the dataset for machine learning use. In *Table 4*, we have included the data types and feature engineering strategy for each of the features we selected for our model.

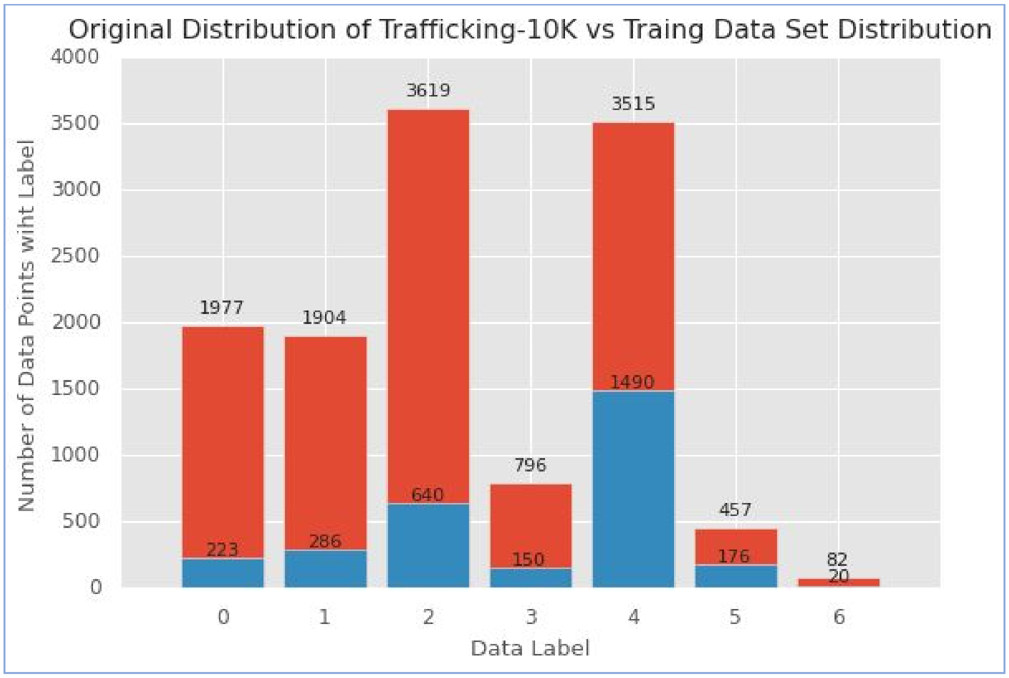
Table 4

Feature	**Data Type**	**Feature Engineering Strategy and Details**
Age	Int	Used regex to extract the age
Weight	Int and Binary	Used regex to extract integers followed by “lbs” or “kg”. Weights were then converted to a Binary value. Weights less than 115 lbs are 1, else they are set to 0.
Country	Char	Text match to a full list of countries or nationalities using pandas.str.join
Continent	Char	Text match to a list of continents using pandas.str.join
Massage (Body Text and URL)	Int	Count of references to massage and spa therapy related words using pandas.str.join Example: ‘massage’, ‘deep tissue’
Minor Indicators (Body and Title)	Int	Count of words related to trafficking of underage girls using pandas.str.join Example: ‘sweet’, ‘candy’, ‘new girl’, ‘fresh’
Organized Work (Body and Title)	Int	Count of words related to organized crime using pandas.str.join Example: ‘2some’, ‘girls’, ‘we’, ‘our’
Unconventional Sex (Body and Title)	Int	Count of words indicating fetishes, risky, not mainstream language using pandas.str.join Example: ‘open minded’, ‘fetish’
Transient Language (Body and Title)	Int	Count of words indicating availability for a short time using pandas.str.join Example: ‘just arrived’, ‘new in town’

Supervised Learning Methodology

The full Trafficking-10K dataset has a total of 12,350 ads, and, after performing feature engineering, we reduced the number of ads to 2,985 ads to build our machine learning models. The dataset reduction eliminated ads that do not have any of the three features: weight, country or continent. *Figure 6* displays the label distribution within the full and reduced datasets. By reducing the dataset, we kept the majority of the ads that have our chosen features present while maintaining the overall label distribution. Reducing the dataset improved modeling accuracy by ensuring that we only included ads with enough features that our model could accurately train on.

Figure 6



After performing feature engineering and data reduction, we begin by building supervised learning models on the Trafficking-10K dataset. Our approach for labeling our scraped data as human trafficking included the following models: SVM, KNN, Neural Network, Gaussian Naive Bayes, Adaboost, and Random Forest. Prior to running the models, we split our dataset into 80% training data and 20% testing data. Included below is an overview of each model included for evaluation and its performance.

Support Vector Machines

Support Vector Machines, or SVM, is a type of supervised learning model that is typically used for regression or classification problems. The model works by maximizing the distances between the nearest data points. This allows SVM to find a boundary or hyper plane that separates the data into the appropriate classes. SVM can be very accurate in its classifications and works well for datasets that are clean and smaller in size. However, SVM can be less than ideal for larger datasets due to the amount of time it takes to train. SVM also can be less accurate with datasets that are noisier and do not have a clean definition of separation. To apply the SVM model on our 10-K dataset, we ran through the following four kernels: linear, poly, rbf, and sigmoid. Applying the SVM model with the linear kernel on our 10-K dataset produced the best accuracy of 59.794% over a training time of 0.813 seconds.

K-Nearest Neighbors

K-NN is a clustering model known as K–Nearest Neighbors also used for both regression and classification problems. This type of model is non-parametric as it doesn't make any assumptions regarding the underlying data distribution pattern. It is a simple algorithm which can be easy to implement, but can become a slower algorithm with larger datasets. While the model can easily evolve in classifying as we are able to provide more training data, it does not work well with imbalance data and will naturally give preference to the most common label. In addition to this, unfortunately K-NN does not have the capability to deal with missing values in the data so it works best when the data is clean with minimal missing data. To implement K-NN on our data, we used the K-NN model from the scikit-learn library. This package assumes 5 neighbors by default, but can be adjusted depending on the data. In order to find the best results, we tested the model through a range of 1 to 5 neighbors. With this, we found that 5 nearest neighbors for the model gave us the best accuracy of 54.124% and a training time of 0.013 seconds.

Neural Networks

Neural Networks are supervised or unsupervised models with a set of algorithms that are designed to recognize patterns within your data similar to the human nervous system. In supervised learning, neural networks use a mix of both input and output signals in order to properly train the model. Opposite from SVM and K-NN, Neural Networks are better suited when you have large amounts of data and can take a very long time to run regardless of the size of data. In addition, when viewing the results of the model, there is limited data provided as to how the network decided the outcome. Labeling our data as human trafficking is very much about recognizing patterns that occur in the text and titles of our datasets, but our dataset is not of ideal size for the amount of time it takes. For our dataset, the neural network model took a training time of 13.865 seconds and produced an accuracy of 58.591%.

Naïve Bayes

Naïve Bayes is a classification model based on the Bayes' Theorem that can be used to find a probability given the probability of a related event. With the Naïve Bayes algorithm, the classifier assumes that a particular feature in a class is unrelated to other present features. This algorithm is especially useful because it is easy to implement and can run very fast. However, a limitation of the algorithm is that it assumes that variables are completely independent which is not ideal in real-life data. Given that our data has several predictors that are most likely connected, we used the label encoder to preprocess our data before running the model. Unfortunately this did not help improve our model as expected and only gave an accuracy of 30.06% with a training time of 0.002 seconds.

AdaBoost

AdaBoost, or Adaptive Boosting, is an algorithm that solves classification problems by transforming weak predictors to strong predictors. An advantage of AdaBoost is that it is easy to implement and fast to run. The model is flexible in that it can be combined with other machine learning algorithms easily such as decision trees. With data, it has the ability to be used with more than just binary data and can be used on text and numeric fields. While it helps boost the ability of weak classifiers, if those classifiers are too weak, AdaBoost can end up with lower margins and potentially overfitting of the data. Applying the AdaBoost classifier on our data resulted in an accuracy of 46.73% over a training time of 0.192 seconds.

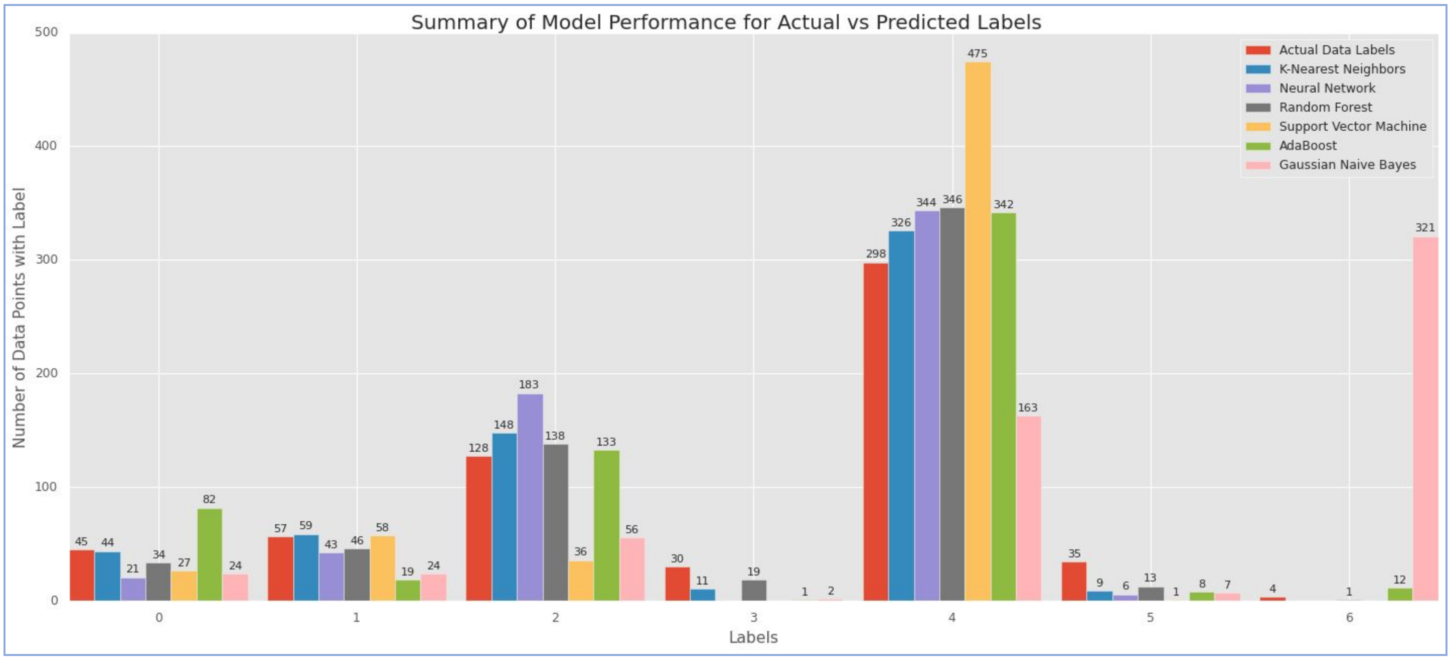
Random Forest

Random Forest is a classifier that utilizes multiple decision trees to make a prediction on the data. The performance of random forest can be competitive in accuracy with other supervised learning models even for large amounts of data, but does take time and resources especially when training a large number of deep decision trees. When implementing random forests, the model does a good job with handling missing data and still produces a good prediction regardless. Unfortunately, random forests, like neural networks, are difficult to interpret and explain as much of the machine learning process is not observable by the programmer. This makes it particularly difficult to understand how the model came up with a prediction. We implemented the Random

Forest classifier on the Trafficking10-K data and the result was an accuracy of 55.66% and a training time of 0.388 seconds.

Of the models we tested, the SVM and Neural Network models resulted in the highest accuracies, followed closely by K-NN and Random Forest. The SVM model only had a training time of less than a second while Neural Network took just over 10 seconds to run. In *Figure 7*, we can see the distribution of labels predictions for each model on the Trafficking-10K dataset compared to the actual data labels.

Figure 7



After we tested each of the chosen models on our labeled Trafficking-10K dataset, we applied the best performing models to the YesBackpage dataset we scraped to see if we could produce accurate labels on unlabeled data. As we are not law enforcement professionals training in identifying human trafficking, “accurate” is a bit of a wide interpretation here. A few samples of the predictive labeling done on the YesBackpage dataset using the highest performing models are shown in *Table 5*.

Table 5

Model	**Label**	**Title**	**Body**
K-NN	5	Hot juice for you to **** and drink up🔥 🔥	🔥Hey My Name is Luckyb 🍬🍭😋 Age: 29💖 indulge in the perfect Package ✨🍭 that I guarantee you'll love come back running for more 💋💋✨ ✔ Real 💯✔ Sexy 💯✔ Discreet 💯💋 Available For Men who desire the best in Companionships ‼ Serious Inquiries only 💋 Playful 🌟 5 star enjoyment 💯 Professional Safe 💯 💕 Beautiful Smooth Body A Mesmerizing face 💕 Ill leave you satisfied from the moment we meet 💋💕✈ CATCH ME WHILE YOU CAN ✈ ✔ ✔ ✨ 100% independent 💎 ✨ CleanClassy 🎀👑 ✨ 100% Real Pictures 📸🎞👀 ✨ Great Hygiene 😛🍭 DONT HESITATE TO CALL or Text ([redacted]) [redacted] Services: Incall, Outcall, Car Date, 2some, 69, ****, FWB, ****, and lots

			more. Am submissive so you can try style with me. I can travel any where in Las Vegas
SVM	5	🔞💦I am a screamer and I love to squirt.Ready for Incalls and outcalls.BDSM,69	Hello I am [redacted],I am clean and smell nice. std/virus free(i have std test results for verification) my top priority and primary service are ===>> giving you top-notch and total attention, anonymity and discretion, identity safety and security, post-escort-services, maximum satisfaction and total guarantee on getting your money's worth. i can give you 69(69 **** position) Prostrate massage bdsm (giving) being filmed (costs a little extra) cim (come in mouth) cob (come on body) cof (come on face) couples dfk (deep french kissing) dinner dates erotic massage extraball (having **** multiple times) face sitting french kissing gang bang gfe (girlfriend experience) humiliation (giving) lap dancing lt (long time; usually overnight) massage mmf 3somes o-level (oral ****) role play & fantasy **** toys spanking (giving) swallow (at discretion) swinging threesome beach parties attending corporate parties travelling companion preparing a meal....\ncell # ([redacted])[redacted] [redacted]\nP.S.!!!!!No psychopaths, No weirdos....just fun and happy moments❤❤. DO NOT contact me to ask for money or FREE-PHOTO-FAVORS as i wont oblige!!! thanks
NN	5	🐙 35 Year_0lder Korean Discreet Oral B-j 420 Style 💦이리 와줘 🐙	**Hi,** I am [redacted]. 35 years 0lder Korean divorced, horny, mom looking for a discreet guy for fun. I am open to almost everything so is your imagination. I can host night or day Full Service charge $70 _ If there is a mutual attraction lets meet _나는 당신을 기다리고 있습니다 **Just verify with your email. If you registered this site then search me now.I'm waiting for U tonight.:::) My Personal E-mail-** [redacted]@gmail.com » please type "**Korean**" subject line. **After text me i will give u my number n deal with video call.** ♣ No busy ♣ Thank You ♣

From a human perspective, we can clearly see why the models would label these three listings as high likelihood of human trafficking. It is also worth mentioning all of the ads below are women and that, though we do not display them here, all of the photos included with the ads in *Table 5* are taken by someone other than the poster; in other words, they are not selfies. As sex work is usually fairly private for individuals who engage in it freely, the fact that the photos are taken by someone other than the poster indicates that another person or party is involved. Though we did not included images in our modeling, this reinforces what we can observe as humans when reading the ads identified as high-likelihood of human trafficking.

The K-NN posting highlights that the poster is "playful" and utilizes candy emojis which is often used by traffickers to indicate a child or very young girl as these are traits/images associated with children. The use of "catch me while you can" is also an indication of transient behavior, a tactic traffickers use to keep victims moving and unable to make connections with people or get help. In addition, explicitly listing the poster's "services" and including phrases like "Am submissive so you can try style with me" indicates that the buyer can engage in BDSM with the poster which, as mentioned before, can be quite dangerous for the poster if the buyer decides to take the kink in dangerous directions.

In the SVM posting, we again see an explicit list of the poster's services, including "being filmed", "BDSM, and "gang bang"; these are activities that are generally risky to the poster's anonymity and safety. Its also worth noting that though the poster is listed as age 29 on YesBackpage (not shown here), they also go through the trouble to include an emoji that indicates that the poster is not 18. It is interesting that the poster felt the need to call out in a separate space that they are not 18 when that bit of information is also included in their profile.

In the Neural Network posting, it is worth noting that the poster is Korean, a region of the world that has a high volume of human trafficking victims. In our research, we found that the poster including a price in their ad is a very high indicator of human trafficking as their is already a rate set for the poster regardless of what the buyer wants from them. The ad also notes that the poster is "open to almost everything", which again sets very few boundaries for their safety and well-being in terms of kink and BDSM.

Evaluating Model Success and Opportunities for Improvement

As the world, technology, and language are all ever-changing, deploying this model would require a level of maintenance to ensure that data and models are continuously validated, updated, and re-validated due to changes in speech patterns, avoidance strategies utilized by human traffickers, and the ever evolving online posting community. This would help to ensure its ongoing accuracy and relevance. Since Backpage was eliminated by the US government, no one site has taken its place. So while YesBackpage was a good choice for us to scrape, that doesn't indicate that it would be a good choice in the future. Human input will be needed to determine where and how online advertisements for human trafficking are evolving.

The Trafficking-10K dataset has seven different possible labels ranging from values of 0-6. The highest training accuracy was ~60% across all of our models, meaning there is lots of room for improvement. We have included below a few options that could be used to create a more robust and accurate model.

n-grams[16]

For each advertisement, we can split our text into a list of n-grams to find common patterns. We can then rank the full list of n-grams across all advertisements, and then compare each ad's n-grams to the full list. The top three ranked n-grams for each ad would add three new features to our models. This would allow us to directly use the text in each ad as opposed to just relying on features we are able to extract and is likely the option that would result in the greatest improvement in accuracy.

Kolmogorov Complexity[17]

The Kolmogorov complexity of an ad is the length of the shortest program that can reproduce the ad. For each ad, we can remove the stop words, analyze it for its complexity, and calculate an entropy value using the formula below. We expect complex ads that have leetspeek, emojis, and foreign language will have higher entropy values indicating a high likelihood of human trafficking.

$$K(X) = -\sum P(x_i) log_2 P(x_i)$$

[16] Li M, Vitnyi PM (2008) "An introduction to Kolmogorov Complexity and its Applications", 3rd edn. Springer, New York

[17] Li M, Vitnyi PM (2008) "An introduction to Kolmogorov Complexity and its Applications", 3rd edn. Springer, New York

Phone Number Area Codes vs Location[18]

Most ads list a contact phone number. On the YesBackpage data, there is a field to populate the phone number, and on the Trafficking-10K data, the phone number was embedded in the text. We can compare the phone number's area code to the city or state the ad is listed in. If the phone number's area code location is far from the location of the ad, then this could indicate possible movement of trafficking victims.

In Conclusion

Human trafficking is a widespread and invasive problem. While we have just addressed one small part of human trafficking and how machine learning models can be used to approach it and solve some of the most pressing issues, human trafficking is not an issue that can be solved with just analytics. It requires the active participation of not just law enforcement and government, but also corporations, nonprofits, and individuals too. We all have an opportunity to step up and be a force for good in the world. While we focused on sex trafficking, there are many forms of human trafficking, the sex trade, forced labor, and domestic servitude. You as an individual and consumer can help fight human trafficking by voting with your dollar and choosing to purchase products that are certified not to engage in exploitative human labor. So the next time you buy clothes from a fast fashion store or coffee from your local grocery, do your research. Find companies who make an effort to ensure that the individuals involved in their supply chains are not being exploited and, together, we can all help raise one another up and cut human trafficking out.

[18] Whitney, Jessica, et al. "Identifying Victims of Human Sex Trafficking in Online Ads." *Encyclopedia of Criminal Activities and the Deep Web*, edited by MEHDI KHOSROW-POUR, vol. 1, IGI GLOBAL, 2020, pp. 497–506.