UNIVERSITY OF WEST LONDON
The Career University

UWL REPOSITORY

repository.uwl.ac.uk

Artificial intelligence as a service for immoral content detection and eradication

Shah, Fadia, Anwar, Aamir, ul haq, Ijaz, AlSalman, Hussain, Hussain, Saddam and Al-Hadhrami, Suheer (2022) Artificial intelligence as a service for immoral content detection and eradication. Scientific Programming, 2022. pp. 1-9. ISSN 1058-9244

http://dx.doi.org/10.1155/2022/6825228

This is the Published Version of the final output.

UWL repository link: https://repository.uwl.ac.uk/id/eprint/8745/

Alternative formats: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk

Copyright: Creative Commons: Attribution 4.0

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy: If you believe that this document breaches copyright, please contact us at open.research@uwl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Rights Retention Statement:

Hindawi
Scientific Programming
Volume 2022, Article ID 6825228, 9 pages
https://doi.org/10.1155/2022/6825228

Research Article

Artificial Intelligence as a Service for Immoral Content Detection and Eradication

Fadia Shah,[1] Aamir Anwar,[2] Ijaz ul haq,[3] Hussain AlSalman,[4] Saddam Hussain,[5] and Suheer Al-Hadhrami[6]

[1]*Department of Computer Science, Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology, Islamabad, Pakistan*
[2]*School of Computing and Engineering, University of West London, London W5 5RF, UK*
[3]*Faculty of Education, Psychology and Social Work, University of Lleida, Lleida 25003, Spain*
[4]*Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia*
[5]*School of Digital Science, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam*
[6]*Computer Engineering Department, Engineering College, Hadhramout University, Hadhramout, Yemen*

Correspondence should be addressed to Saddam Hussain; saddamicup1993@gmail.com and Suheer Al-Hadhrami; s.alhadhrami1@gmail.com

Received 19 November 2021; Accepted 27 December 2021; Published 17 January 2022

Academic Editor: Punit Gupta

Copyright © 2022 Fadia Shah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Social media is referred to as active global media because of its seamless binding thanks to COVID-19. Connecting software such as Facebook, Twitter, WhatsApp, WeChat, and others come with a variety of capabilities. They are well-known for their low-cost, quick, and effective communication. Because of the seclusion and travel constraints caused by COVID-19, concerns, such as low physical involvement in many possible activities, have arisen. Depending on their information, knowledge, nature, experience, and way of behavior, various types of human beings have diverse responses to any scenario. As the number of net subscribers grows, inappropriate material has become a major concern. The world's most prestigious and trustworthy organizations are keenly interested in conducting practical research in this field. The research contributes to using Artificial Intelligence (AI) as a service (AIaaS) for preventing the spread of immoral content. As software as a service (SaaS) and infrastructure as a service (IaaS), AIaaS for immoral content detection and eradication can use effective cloud computing models to leverage this service. It is highly adaptable and dynamic. AIaaS-based immoral content detection is mostly effective for optimizing the outcomes based on big data training data samples. Immoral content is identified for semantic and sentiment evaluation, and content is divided into immoral, cyberbullying, and dislike components. The suggested paper's main issue is the polarity of immoral content that can be processed using an AI-based optimization approach to control content proliferation. To finish the class and statistical analysis, support vector machine (SVM), selection tree, and Naive Bayes classifiers are employed.

1. Introduction

Connecting various people, the Internet has no longer most effectively aided the folks around the arena, ultimately, a huge extent of users can specify their perspectives. It is not acquainted as a city village by way of [1] insisted that despite the range of software applications have their procedures, nevertheless, the public voice is heard to share their views. Shah [2] mentioned big variations in lives. There are various factors, and COVID-19 has remote human beings from social lifestyles. Plentiful people are distantly related to society. COVID-19 has badly reshaped the lives of people, and now, nearly every research-based domain is actively exploring its consequences by various means. Shan [3] referred that this has made humans specific to their opinions on the Internet because the most

effective manner for growth is to be truly mingled. The most popular social media platforms are YouTube, Facebook, TikTok, Twitter, etc.

Omar et al. [4] found that community-based software utilities encompass famous systems like Table 1 suggests the number of users improved in any type of social network from the last decade. The contemporary research directs that billions of users are connected through social media organizations. Because of the pandemic, the users prefer the Internet to live participation. Huynh [5] elaborated that the number of customers is countlessly accelerated. Several platforms are providing their specialized services to communiqué because physical presence is not obligatory. Furthermore, because of the hurdles of personal participation, the priority of human beings is the Internet because of its availability, easiness, and fast response. However, it is sometimes a problem to find the relevant information from the results provided by the Internet.

Another problem of these social platforms is that the use of these remote platforms is not always productive. Always, a certain group of people is responsible for creating trouble, humiliating, and wasting others' time. It is mentioned by researchers like [6] who elaborated that using these platforms creates troubles for a large number of users. Because of the community boom, this most probably these boards. The problem is crucial since it is now feasible for users to avail the message delivery to the rest of the users after simply signing into the system. The person who is a consumer of the software product can attain the attention of the other users by initiating a hot topic and involving several users. The comments can be used in two ways; constructively for the learning and problem-solving mode; or the manipulated and exaggerated theme by involving comments like racism, extremism, political dispute, or specific objectives.

The researchers like [7–9] are those examples, including [10] quantified linguistic behavior o the communication method. The techniques of effective communication are verbal exchange and nonverbal communication. If the primary purpose is met, the communication is effective. The receiver understands exactly what the sender wants to say. While verbal communication is not always effective, it indicates that the sender is saying something the receiver does not understand. As virtual communication lacks physical engagement, there is a lack of body language, such as the tone of voice, eye contact, and others, which can lead to misleading impressions. Hu [11] said that this false impression creates conflicts. Because of this, aggressive and abusive language initiates. Even though there are various elements worried. It includes a lack of records, and the distinction makes it hard enough to make it understand. It creates doubts and many other problems. When an individual wants to win the comment, the individual can go up to any level of exchange of comments. Those are the types of individuals who do not hesitate to harass others. Often, the annoyed candidates annoy others with their remarks, scripts, and responses on shared media.

The further study is consistent with that of the academician [12]. Harassment is identified by offensive stated material. Mankind's isolation is more usually found in immoral stuff, which is a serious problem. It is increasing daily. Immoral content is produced by a specific group of individuals. It includes a common mentality from a specific period, gender, reputation, education, and religion. It is a problem that a large number of clients on social media are dealing with. Almost every social structure provides some sort of venue for recording or avoiding immoral content [13]. Reporting this person may result in a warning or, in the worst-case scenario, a permanent ban. Toxic information can be quite hazardous to even the most innocent minds.

Table 1: Increase in the number of people using social media each year.

year	Users in millions
2010	95
2011	120
2013	148
2015	210
2017	245
2019	280
2020	294
2021	326

The big populace browsing the net has a variety of kids who are new to the worldwide village. They are studying and making their minds apprehensive of the sector. A few researchers [14] are still looking into the content, while a few girls are housewives who spend their leisure time working on websites. The Immoral depend is being managed by social media web-based systems in a completely professional manner. There is a unique mechanism in place to deal with this type of information. The assessment committee and the concerned personnel are constantly on the lookout for complaints and inquiries. The pathetic content displayed on the Internet can be refined using an information technology architecture.

Oppressive language character is not just about as basic as a lump of cake. Sorting out exploitative substance material is a difficult task, especially when it needs to be extracted from large data. Oppressive language is of several kinds. The sentence structure is made up of a variety of expressions. The character of hazardous words in this series surely necessitates a handful of particular strategies. Table 1 illustrates that the number of net clients is steadily increasing. With a larger population, there are more opportunities to obtain diverse measures.

Governmental concerns have been raised, and strategies are being explored in every impacted district around the world to develop new mechanisms to restore normalcy and control coronavirus. Because the Coronavirus has boosted web traffic, determining the text's sequence is crucial. The prevalence of web material can be spread, and false content can be eliminated. Savelev et al. [15] expressed with regards to quick data sharing and spread that the web is the main source to associate, and similar data is divided among different clients. Some of the time the worry isn't about unwavering quality, however, individuals share content unexpectedly to make their darlings update about the ongoing occurring. Boksova et al. [16] referenced that individuals are classified as web clients and nonclients.

Gao et al. [17] made an intensive report on assessment investigation and the utilization of online media content. Social substance investigation is prevalently broken down by opinion examination. The different strategies of opinion examination are utilized, including deceptive substances. There are various methods, such as determining the uniqueness of senders, messages, and recipients. Krzewniak [18] reported that web-based clients have access to all data on the Internet, however, this might put the information of clients in grave danger. There has been research on how scientists have used information mining processes to characterize data in the past. Rule-based frameworks, solo learning procedures, administered learning strategies, and other well-known ways are examples of this dynamic. As shown in Table 1, the number of informal community clients is increasing every year. Gianfredi et al. [19] Table 2 shows the most well-known exercises [20] and the level to which they were carried out by individuals in the global information media status report.

The table above shows how frequently people use Internet hotspots for various activities in their daily lives. The web is considered the backbone by the majority of the dynamic. Considering these characteristics to be of top priority, the suggested research paper identifies two common issues. The first is the assurance of deceptive content from online media (utilizing AI classifiers). The subsequent issue is the assignment of classifications to untrustworthy substances based on their power degree, including simulated intelligence, to the point where the most dreadful class should be acknowledged for special arrangements.

The plausibility of the review is that the web cannot separate the moral and exploitative substance. The valid and bogus assertions cannot be recognized. [21] If questionable data is halted to engender and its ubiquity can be diminished, there are opportunities to guarantee the web is dependable. Likewise, there are opportunities to decrease the consuming circumstance and lead the course toward the quiet circumstance as the consistently famous substance is not a dependable substance as well. Man-made intelligence is so stretched out in different applications that now it is not only utilized for consistent programming but also as a policymaker and a device. Shrewd frameworks (K. Ghosh, 2019) for web-based media are fundamental these days for quality assistance arrangement.

2. Related Work

The major purpose of the research is to create text-mining algorithms for detecting immoral content on social networking sites [22]. Elareshi et al. [23] have discussed and evaluated several areas of aspect mining using text in depth. Text mining is the process of analyzing content using a variety of machine learning algorithms established by experts [24]. Supervised learning, unsupervised learning, rule-based learning, pattern-based learning, and so on are common learning methodologies. As a result, there are three categories of literature on this subject.

Table 2: Proportion of Internet users activities.

Sr. No	Utility	Proportion
1	In contact with friends/Family	48.6
2	Spare time usage	36.3
3	News updates	35.2
4	Entertainment	30.9
5	Information in the air	29.3
6	Products browsing	26.1
7	Opinion sharing	24.9
8	Live stream watch	23.7
9	Make new contacts	23.6
10	Content browsing from favorite brands	22.9
11	Work-related networking and research	22.7
12	Finding communities and interests	22
13	Favorite sports	21.1
14	Celebrities & stuff	20.7

2.1. Artificial Intelligence as a Service (AIaaS). Artificial intelligence is a concept that refers to the scientific studies and practices aimed at improving the efficiency with which robots make decisions. The term "intelligence" encompasses a wide range of concepts. It includes factors, such as addressing the problem in a short amount of time, solving the problem correctly, providing the best answer, and so on, according to the experts [25]. Computers are used to efficiently solve complex problems. Machine learning, with or without the assistance of a person, strives to solve issues using computations and create proper outcomes (Jiwon Kang). Several machine learning techniques exist that combine great computational capabilities with low computing expenses. Li et al. [26] reported that supervised, unsupervised, and reinforcement learning are among the machine learning approaches employed. Shah and Li [27] AI's effects on jobs and society [25] management and strategic challenges associated with AI [25]. They all use different approaches, however, the end goal is the same: to find the most effective solution. AI is utilized as a third party because it strives to deliver better/improved reasoning. The computed results are employed as a third party, and the data is communicated and evaluated using these results in many research and technology domains. Machine learning and artificial intelligence are employed in a variety of platforms. It is for this reason that it is known as third-party evaluation. Artificial intelligence and machine learning have a big impact on information and communication systems in general. The machine is being programmed to calculate the solution automatically.

As a result, the ideal solution is not only quick but also offers precise results. The year Takeuchi and Yamamoto [28]. The more capable a system is of delivering accurate results, the more it is called a smart system [29]. It also offers the capability of halting processing in the event of an error or autonomously rectifying processing before output without contacting an external entity. The term "agents" is used to refer to intelligent systems. Agent decision-making has been a prominent issue in research for years, and now it can be used as a service to run a variety of systems.

Several analysts [30] devised various ways to incorporate AI into their projects to improve the quality and efficiency of their work. Real-time frameworks are those that initiate their actions at a specific interval or timeframe. Social media content is derived from a variety of sources. The majority of the design is built on a dispersed framework layout. The cloud underpins all-encompassing interfacing of workstations in various regions, linked through organized topologies, and ensures error-free data transfer. Such systems are well-prepared so that sending data from one source to a few kilometers away arranged recipient takes only a few seconds.

The Internet of Things (IoT) is a connected example of a socially connected distributed system. Gianfredi et al. [19] articulated around the data generated by a variety of devices and disseminated through social media applications. Because Yang et al. [31] is like mimicking human insights in information translation and applying a clever decision-making framework, it can be used as a supplement to circumvent social media's forceful content. An intelligent degree for social media substance isn't too brutal rough substance; rather, there should be distinguishing proof of brutal substance and destruction of that smudged stuff some time lately increasing it to various systems. Here, we use AIaaS for immoral content detection and eradication can use effective cloud computing models to leverage this service. It is highly adaptable and dynamic. AIaaS-based immoral content detection is mostly effective for optimizing the outcomes based on big data training data samples.

2.2. Machine Learning. Text mining techniques, which are applied under the platform of machine learning techniques, help identify immoral content material. Machine mastery for medically-linked textual material polarity is also based on strategies that are a systematic manner to apply some algorithms by teaching a computer to decide for itself.

Unsupervised, supervised, and reinforcement learning is used to apply machine learning-based text mining strategies. In addition to labeled datasets, the supervised learning-based skilled models [32] emphasized text mining in the biomedical dataset. The impacts already understood are included in the supervised learning of facts. The outcomes are evaluated using a supervised learning-based approach model. The data set was used for prediction by using a support vector machine (SVM) decision tree.

His ideality was a topic for churn, and he developed a neural network for audience-related literary material with the help of [33]. However, the findings of a survey-based dataset linear regression, which are well-known supervised researching approaches for determining immoral text, have been spectacular. The goal of this study is to use supervised learning to identify the outcomes from the data.

The application of supervised learning to understand the use of the SVM algorithm was originally done by [34]. It was previously multilingual and employed seven languages, as well as a variety of bootstrap sentiment analysis methodologies.

Once, a progressive B4MSA polarity classification was utilized. Forman, et al. [35] and Heri [36] introduced a the detection of negative consequences of publicly publishing humiliating content material on social media. Other multilingual sentiment assessment models employed were SENTIPOLC′14, SemEval'15–16, and TASS′15. It determined the harshness and disrespectfulness of the words. For classification-based tasks, the results were environmentally favorable.

2.3. Big Data, A Source of Unethical Content. Social media platforms are the sources of revenue gains for various regions. Various small and medium financial industries are now applying technology as a mandatory component of their procedures. Shan et al. [37] emphasized the vitality of platforms for data reforms for the future growth of a region. The increase in data during the last decade is increasing, however, the pandemic since 2019 has isolated the social activities of mankind. Technically, domestic isolation has raised behavioral complications. The vitality of the content generated during a pandemic is more polluted with the combination of less ethical and more unethical statements. People found no other way to communicate, For online social networking applications, where it is recommended to proceed with all the activities while staying at home, people have no alternative ways to establish social networking activities [38]. Only the online social media platforms facilitate them to meet new people virtually. Platforms such as Twitter, Facebook, Netflix, Yahoo, Whatsapp, Wechat, and similar applications can help one to find new people and discuss his ideas fearlessly.

Another research by [39] illustrated the sufferings raised because of more use of gadgets, especially by the youngsters. Big data is saturated with more impact of poor communication, inappropriate words, lack of belief, and particularly, it is saturated with the impact of cyberbullying impression. He used the classification approach for unethical text determination. Most of the platforms justify there are more proportions of immoral content. The proposed study emphasizes more on Twitter data set because of its diversity in the same domain, availability, and rapid user interactions.

2.4. Datasets. The dataset is made of Twitter, Kaggle, and survey-based information. There was no uniformity in the columns, information format, or aesthetic style because the data was acquired from a variety of sources. Although this unstructured data had previously been in text format, it was no longer appropriate for the next stage of processing. To advance to the next level, the authors converted heterogeneous data into homogeneous data with a standardized form. Twitter also offers an API (Twitter Stream API) that may be used to retrieve data from the website, such as tweets, comments, and likes, but only with authentication.

In social networks, the detection of abusive content, cyberbullying, and harassment is typically framed as a classification problem. The homogeneity of the data set was not the only concern. There was also the issue of multiclass imbalanced datasets, which resulted in a varied distribution of situations into instances.

Using the approach of oversampling, the problem was attempted to be minimized up to the lowest stage. Rapidminer is used to process the records to simulate models quickly and accurately.

2.5. Model. The primary goal of the proposed study is to determine which records contain immoral content. The aggregated records have been reduced to a set of documents containing 13,000 tuples and seven features. AI's utility is machine learning. After consolidation, the factual units were in a homogeneous state. This data was once sophisticated enough to be used as input in the mannequin. The flowchart of the AIaaS-based model for immoral content detection and eradication is shown in Figure 1.

The figure indicates there are two sections of the proposed system. The top approach of the flow chart shows the processing of data from the initial stage till the identification of unethical content. The process continues for the entire data chunks. It is sure that if the data set is big enough, the methodology is applicable for the data segments until the entire data is examined. The results are stored and the same logic is applied as service-based architecture. It can be possibly achieved for cloud architecture. The dedicated unethical content identification cache can encounter the text with segmentwise results calculation until the condition is stable. Both types, i.e., semantic and sentiment analysis, are possible in the above model.

2.6. Data Preprocessing. The preprocessing of records is the initial stage in the suggested paradigm. The facts collection is heterogeneous with little homogeneity and a large number of information elements. The preprocessing of facts is usually the first stage for each mannequin. It is utilized to turn it into a more homogeneous state. The information that has resulted is now of high quality. Data preparation is required to remove noise from the data and to improve the accuracy of the effects purchased once purified records have been utilized for mannequin training.

There were numerous abnormalities in our dataset, including incomplete records with missing values, incorrect values, and special data sorts for a variety of attributes. The information gathered from a variety of sources was also incorrect. It is critical not to remove any useful information from the content while preparing the data. Anomalies were present in the data sets used in our research.

It is important that often preprocessing and data cleansing deletes the missing and abnormal or incorrect values. However, in the proposed study, preprocessing is applied to the data items that contain minimum null values. The missing values were replaced with the most possible/nearly estimating possible values. The impure data was found to be approximately 0.002%, which is quite less in proportion. The problem of overfitting and underfitting were carefully observed so that the data quality should be consistent. The data after preprocessing is quality-oriented.

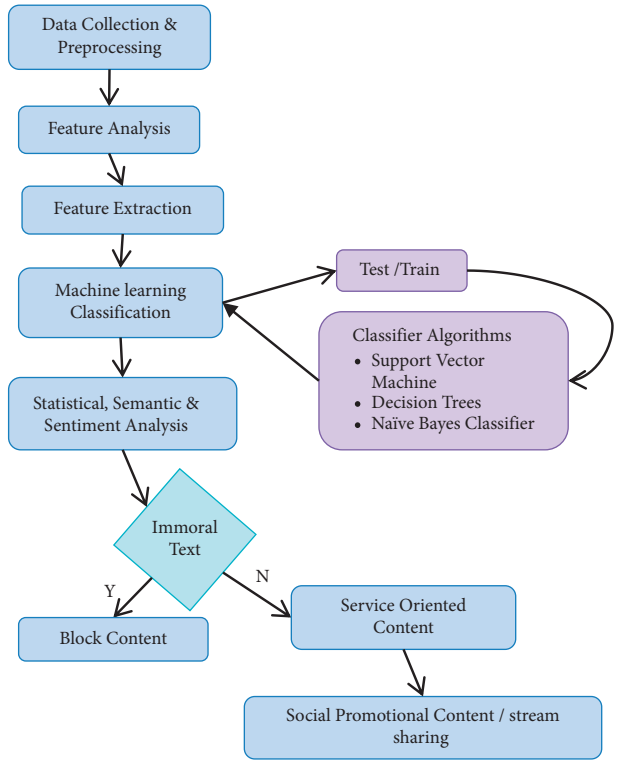

Figure 1: A flowchart of AIaaS-based model for immoral content detection and eradication.

2.7. Opinion Mining NLP. The data set is ideal for linguistic research. The discoveries that are reached following the information analysis are referred to as features. The two most important types of features are determined in this study. The facts are a collection of social behavior statistics. In this study, two categories of points are specifically examined. Sentiment points and semantic characteristics are what they are called. In the model for detecting abusive language, feature extraction is crucial. These considerations will aid in the detection of abusive phrases and context-based abuse. Preprocessed data aids in the extraction of particular elements, such as sentiment features, semantic features, unigram features, and pattern features, to detect abuse and subtypes, such as aggression, dislike, misbehavior, cyberbullying, and vulgar language in the material. The sentiment feature determines whether a tweet or remark has a sentiment thing, whereas the semantic feature aids in the detection of contextual base abuse by the usage of a specific letter, symbol, or word in the tweet.

2.7.1. Semantic Analysis. Semantic analysis is used to determine the relationship between the sentences. It can distinguish the sentence's class. It is the type of sentence that is employed in the sentence. It can distinguish the sentence's class. It is the type of sentence that is employed in the sentence which means the clear theme of the context is expressed in terms of semantic analysis [40]. Particularly, it is the comment on the expression's context. The planned research also assesses whether the statement is

straightforward or contains some hidden meaning. In semantic analysis, the selection of words and closures is critical. It is used for approximation analysis in machine learning. The theme of predicate information analysis is used to complete it. The proposition is a collection of predicates and quantifiers. These assertions are used to complete a sentence's structure. A minimum of one item of information should be included in each proposal. They result in variables, which are then utilized to form a variety of functions that provide useful data to the systems. When analyzing a sentence for semantic analysis, the same common sense is employed. In semantic analysis, letters, symbols, and quantifiers are used.

2.7.2. Sentiment Analysis. Another feature is sentiment evaluation, which is the determination of the sentence's polarity. It is utilized for opinion mining, with the ability to look at things in three ways: positively, negatively, or neutrally. It can tell whether the text is polarised in a very positive, positive, neutral, negative, or extremely poor way. Doaa Mohey El-Din Mohamed Hussein, 2021 identified issues in analyzing social media content (Shambhavi Dinakar). The texts utilized for sentiment analysis include a variety of languages (Rastislav Krchnavy) and negative and slang phrases, hashtags, and emoticons (Dr.Pappu Rajan). A lot of lookup scholars have extended this kind of view for herbal language processing. Microblog textual content analysis (Fotis Aisopos).

Table 3 states that out of 4458 tuples of datasets 1,2 and three The category of kind of content is decided by sentiment and semantic analysis. Hence, the classification of subcategories is made viable to in addition follow the model.

2.8. Feature Extraction. The facts set consisted of 12 features. The computation was once processed by way of using.

The key function of function extraction is to determine the most influential features that take part in result generation. In the chosen information set, the aspects that conclude their content as immoral, cyberbullying, or dislike text were text, content, category, and ..... Particularly, the reachable information units have awesome facets called textual content and effects as two foremost chosen facts.

2.9. Optimization and Training. This section includes the model optimization and training for the dataset so that the results could be precise and valid.

2.10. Classification. Additionally, divide the dataset into a teach or look at dataset to train a model for detecting abusive language. As the mannequin is trained using a 70% dataset, the records set is partitioned into the educated and checkfacts sets. The proposed lookup is classified using the supervised computer to get to know the method [41]. The ramifications of the statistics units are already recognized in this case. The ability to classify data is such that the results of trained statistics sets can be compared to the outcomes of checking out records.

Table 3: Opinion mining classification.

	Semantic data	Sentiment data	Total
Data set 1	2517	1941	4458
Data set 2	1375	1433	2808
Data set 3	1341	2307	3648
Tuples	5233	5681	10,914

There are a variety of classifiers [42]. Classification techniques are even important on encrypted data [43] and [44]. The accuracy measurement determines the overall performance of the mannequin. The three classes that are the consequence of categorization are immoral, cyberbullying, and hatred [45]. The classification process is divided into three stages. The classification is done by three classes. It incorporates binary classification, which divides the records into two categories. The outcome of binary categorization distinguishes the two major groups. Binary categorization has the advantage of displaying the straightforward distribution of statistics into two main groupings.

Ternary classification is the next step after categorizing the statistics into binary labeled results. As the proposed study identifies three classes, namely immoral, cyberbullying, and dislike, the tertiary classification is used. The research completes its classification, however, if it needs to be extended to more than three classes, a multivalued classification is used.

The three most common and popular classifiers used in this study were Naive Bayes, SVM, and Decision Tree, which were used to divide the content material into categories, such as aggression, misbehaving, dislike, cyberbullying, vulgar message, and ordinary. In Table 4, the categorization consequences are discussed. The commonplace class represents natural language. If a tweet or remark is not harsh, it will be listed here. By modifying these factors, the optimal parameter will help improve the overall performance of classifiers. The accuracy of each classifier will be displayed one by one for each class in the result. This accuracy will exhibit how precisely to discover the abusive language from the content. The exponential growth of big data might increase the chance of abusive and unethical content compared to ethical content [46]. The proportion of neutral content is much less than the other components. Thus, there are more aspects the people behave unethically over the Internet. It may be because of reasons, such as their identity being unknown and them being remotely available, or concerns, such as they can approach various platforms free of cost and they have freedom of speech. However, there are more traces of nonvaluable text.

3. Result and Discussion

The binary classification process produces two main lessons and has a 91.2 percent accuracy rate. The third category is considered neutral. Hence, if one runs the classification model on these three instructions, one gets an accuracy of 85.70%.

Table 5 indicates the accuracy of the proposed approach. It means the identification of unethical content via machine

Table 4: Sentiment analysis classification.

	Immoral	Dislike	Cyberbullying	Neutral	Total
Data set 1	1257	1327	794	1080	4458
Data set 2	736	1434	43	595	2808
Data set 3	1454	1324	462	408	3648
Tuples	3462	3354	2965	1133	10,914

Table 5: Social dataclassification

	Sentiment analysis			Semantic analysis		
	Precision	Recall	F1 score	Precision	Recall	F1 score
Immoral	0.57	0.32	0.46	0.42	0.71	0.46
Cyberbullying	0.43	0.41	0.53	0.43	0.12	0.36
Dislike	0.41	0.31	0.41	0.44	0.31	0.32

learning algorithms is promising, particularly for the datasets that are of larger size. The sentiment analysis and semantic analysis evaluations provide noticeable results. This approach can be applied for the larger data sets by selecting the segments of huge chunks of data, and the repetition of the process at various intervals can be deterministic for the data available.

3.1. Content Oriented AI as a Service. After passing through the model's elements, the outcomes disclose three key parameters. The statistical output is once again used as a source for hospitality analysis [47]. From banking transactions [48]. For content show processing, the impacts obtained after sentiment and semantic analysis are critical. Instead of displaying before the readers, items that are more infected with traces of immoral, cyberbullying, and dislike classes with high accuracy (meaning are more corrupt) can be prevented. Low-accuracy communications may provide a risk of displaying content material on the Internet. The most situation column in Table 1 has a high F1 score. In any event, the recall, precision, and F1 rating factors would no longer be prioritized.

3.2. AI Influencial Content Control. The content material with excellent polarity is ideal for displaying in front of the reviewers. Negative polarity and a poor F1 score are the examples of material with negative impact. Alternatively, if the content material is displayed, it may harm the users. After the model has been applied, the cumulative results can be separated into different labels and the time stamp can be lowered. If the content is very obnoxious, it may be prohibited at an early stage. The dreadful effect can be mitigated in this way. This method is ideal for websites that have a great reputation and only provide excellent service. These social media platforms are often at high popularity.

4. Conclusion

The suggested investigation is completed to find and remove immoral content from social media networks. Misbehaving, cyberbullying, and use of immoral language; in the statement is unethical content. Textual content mining is done using a supervised learning approach. To obtain reliable results, firstly, unethical content identification is done. Then, based on these results, the content containing illegal text can be prevented. The use of a multiclass imbalanced dataset is refined with resampling, undersampling, and oversampling techniques. Then, sentiment and semantic analysis methods are applied to find the severity of immoral content. Decision Tree, SVM, and NaveBayes are used for classification. The content polarity and unethical sensitivity are determined. The negative content is limited to the social media display. The feasibility of this study is extremely important for better text-based component decision-making. New policies pushed decision-making, social content delivery, and display, as well as the permissibility and prohibition of ethical writing on reputable and genuine websites. The proposed study's social advantages are measured in terms of the amount of content that can be exhibited, regional characteristics from the community, and many more.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest associated with publishing this paper.

Acknowledgments

This research was supported by the Researchers Supporting Project number (RSP-2021/244), King Saud University, Riyadh, Saudi Arabia.

References

[1] C. Forman and N. van Zeebroeck, "Digital technology adoption and knowledge flows within firms: can the Internet overcome geographic and technological distance?" *Research Policy*, vol. 48, no. 8, p. 103697, 2019.

[2] N. Shah, "From global village to global marketplace: metaphorical descriptions of the global Internet," *International Journal of Media and Cultural Politics*, vol. 4, no. 1, pp. 9–26, 2008.

[3] Y. Shah, Y. Liu, Y. Liu, F. Shah, and F. Shah, "Challenges of small- and medium-sized businesses in Pakistan due to COVID-19 pandemic," *R-ECONOMY*, vol. 6, no. 3, pp. 222–226, 2020.

[4] H. Omar, S. Haggag, J. Grundy, and M. Abdelrazek, "COVID-19 vs social media apps: does privacy really matter," in *Proceedings of the 2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS) 2021*, IEEE, Madrid, ES, May 2021.

[5] G. T. T. Huynh, N. T. B. Chung, and T. T. Phung, "How to purchase an order from brick and mortar retailers during COVID-19 pandemic? A rise of crowdshipping," *International Journal of Data and Network Science*, pp. 439–450, 2021.

[6] M. Han, X. Y. Tan, R. Lee, J. K. Lee, and R. Mahendran, "Impact of social media on health-related outcomes among

older adults in Singapore: qualitative study," *JMIR Aging*, vol. 4, no. 1, p. e23826, Article ID e23826, 2021.

[7] P. Walaski, "Social media: powerful tools for SH&E professionals," *OnePetro*, pp. 40–49, 2013.

[8] Y. Katz, "Social media is powerful, but can it change policies of institutionalized organizations such as the Israeli army?" *International Journal of Humanities and Social Sciences*, vol. 6, no. 6, pp. 29–34, 2017.

[9] D. R. Admane, "**Social media powerful way to bring positivity during the covid-19 pandemic**," *Social media during covid-19 pandemic*, vol. 8, no. 1s, 2021.

[10] D. G. Taylor, "Putting the "self" in selfies: how narcissism, envy and self-promotion motivate sharing of travel photos through social media," *Journal of Travel & Tourism Marketing*, vol. 37, no. 1, pp. 64–77, 2020.

[11] W. Hu, "Innovative university subject teaching mode under the information technology environment," *Panyapiwt international conference on social sciences and management*, vol. 81, pp. 571–578, 2020.

[12] N. Peimani and H. Kamalipour, "Online education and the COVID-19 outbreak: a case study of online teaching during lockdown," *Education Sciences*, vol. 11, no. 2, p. 72, 2021.

[13] F. Shah, J. Li, J. D. Khan, F. Shah, and Y. Shah, "Hybrid compression for medical big data," *ICCWAMTIP*, vol. 2018, 2018.

[14] S. Raj Sangwan et al., "D-BullyRumbler: a safety rumble strip to resolve online denigration bullying using a hybrid filter-wrapper approach," *Multimedia Systems*, 2021.

[15] A. O. Savelev, A. Y. Karpova, D. V. Chaykovskiy et al., "The high-level overview of social media content search engine," *IOP Conference Series: Materials Science and Engineering*, vol. 1019, no. 1, p. 012097, Article ID 12097, 2021.

[16] J. Boksova, M. Bokša, J. Horák, K. Pavlica, J. Strouhal, and S. Saroch, "E-government services and the digital divide," *Journal of Telecommunications and the Digital Economy*, vol. 9, no. 1, pp. 27–49, 2021.

[17] Z. Gao, S. Fujita, N. Shimizu et al., "Measuring public concern about COVID-19 in Japanese internet users through search queries: infodemiological study measuring public concern about COVID-19 in Japanese," *Internet Users Through Search Queries: Infodemiological Study*, vol. 7, no. 7, Article ID e29865, 2021.

[18] D. Krzewniak, "Selected elements of the internet users' information security culture," *Roczniki Nauk Społecznych*, vol. 13(49), no. 2, pp. 121–140, 2021.

[19] V. Gianfredi, S. Provenzano, and O. E. Santangelo, "What can internet users' behaviours reveal about the mental health impacts of the COVID-19 pandemic? A systematic review," *Public Health*, vol. 198, pp. 44–52, 2021.

[20] B. Carlson, "Indigenous internet users: learning to trust ourselves," *Australian Feminist Studies*, vol. 36, no. 107, pp. 9–25, 2021.

[21] Y. Song, K. H. Kwon, Y. Lu, Y. Fan, and B. Li, "The "parallel pandemic" in the context of China: the spread of rumors and rumor-corrections during COVID-19 in Chinese social media," *American Behavioral Scientist*, vol. 65, no. 14, pp. 2014–2036, 2021.

[22] S. Chen, X. Guo, T. Wu, and X. Ju, "Exploring the online doctor-patient interaction on patient satisfaction based on text mining and empirical analysis," *Information Processing & Management*, vol. 57, no. 5, p. 102253, Article ID 102253, 2020.

[23] M. Elareshi, A.-K. Ziani, and A. Al Shami, "Deep learning analysis of social media content used by Bahraini women: WhatsApp in focus," *Convergence: The International Journal of Research Into New Media Technologies*, vol. 27, no. 2, pp. 472–490, 2020.

[24] D. Antons, E. Grünwald, P. Cichy, and T. O. Salge, "The application of text mining methods in innovation research: current state, evolution patterns, and development priorities," *R & D Management*, vol. 50, no. 3, pp. 329–351, 2020.

[25] M.-H. Huang and R. T. Rust, "A strategic framework for artificial intelligence in marketing," *Journal of the Academy of Marketing Science*, vol. 49, no. 1, pp. 30–50, 2021.

[26] M. Li, D. Yin, H. Qiu, and B. Bai, "A systematic review of AI technology-based service encounters: implications for hospitality and tourism operations," *International Journal of Hospitality Management*, vol. 95, p. 102930, 2021.

[27] F. Shah and J. P. Li, "Broad big data domain," in *Proceedings of the 2017 4th International Conference on Systems and Informatics (ICSAI)*, IEEE, Hangzhou, China, Nov. 2017.

[28] H. Takeuchi and S. Yamamoto, "Method for assessing the applicability of AI service systems," *Human Centred Intelligent Systems*, vol. 189, pp. 323–334, 2021.

[29] X. Lv, Y. Liu, J. Luo, Y. Liu, and C. Li, "Does a cute artificial intelligence assistant soften the blow? The impact of cuteness on customer tolerance of assistant service failure," *Annals of Tourism Research*, vol. 87, p. 103114, 2021.

[30] S. Robinson, C. Orsingher, L. Alkire et al., "Frontline encounters of the AI kind: an evolved service encounter framework," *Journal of Business Research*, vol. 116, pp. 366–376, 2020.

[31] Y. Yang, L. Li, and J. Jiang, "The impact of COVID-19 pandemic on emerging country stock markets: evidence of the value effect," *Emerging Markets Finance and Trade*, vol. 58, no. 1, pp. 70–81, 2021.

[32] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, and J. Kim, "COVIDSenti: a large-scale benchmark twitter data set for COVID-19 sentiment analysis," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 4, pp. 1003–1015, 2021.

[33] J. Mell, S. Jang, and S. Chai, "Bridging temporal divides: temporal brokerage in global teams and individual performance," *Academy of Management Proceedings*, vol. 2020, no. 1, p. 13037, Article ID 13037, 2020.

[34] C. P. D. Cyril, J. R. Beulah, N. Subramani, P. Mohan, A. Harshavardhan, and D. Sivabalaselvamani, "An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM," *Concurrent Engineering*, vol. 29, no. 4, pp. 386–395, 2021.

[35] C. Forman, A. Goldfarb, and S. Greenstein, "How did location affect adoption of the commercial Internet? Global village vs. urban leadership," *Journal of Urban Economics*, vol. 58, no. 3, pp. 389–420, 2005.

[36] H. Heri, E. D. Minarti, and L. Linda, "Mathematical difficulties of 7th grade MTs students analysis in online learning based on the realm bloom's taxonomy," *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang*, vol. 5, no. 2, p. 201, 2021.

[37] F. Shah, fnm au, Y. Liu, Y. Shah, and F. Shah, "Influence of shareholder equity on trade credit demand: the study of non-financial firms in Pakistan," *R-Economy*, vol. 7, no. 1, pp. 61–67, 2021.

[38] P. Xhelili, E. Ibrahimi, E. Rruci, and K. Sheme, "Adaptation and perception of online learning during COVID-19 pandemic by Albanian university students," *International Journal on Studies in Education*, vol. 3, no. 2, pp. 103–111, 2021b.

[39] N. Radwan, *Big Data Ethics*, vol. 19, no. 1, 2021.

[40] X. Yang et al., "Frontline encounters of the AI kind: an evolved service encounter framework," *Journal of Business Research*, vol. 116, pp. 366–376, 2020.

[41] O. FYFY, fnm au, A. JET et al., "Supervised machine learning algorithms: classification and comparison," *International Journal of Computer Trends and Technology*, vol. 48, no. 3, pp. 128–138, 2021.

[42] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, no. 3, pp. 159–190, 2006.

[43] C. Buffington, J. Fields, and L. Foster, "Measuring the impact of COVID-19 on businesses and people: lessons from the census bureau's experience," *Aea papers and proceedings*, vol. 111, pp. 312–316, 2021.

[44] S. M. Weiss and I. Kapouleas, "An empirical comparison of pattern recognition, neural nets, and machine learning classification methods," *Proceedings of the 11th international joint conference on Artificial intelligence*, vol. 1, 1989.

[45] D. Michie, D. J. Spiegelhalter, and C. Taylor, "Machine learning, neural and statistical classification," *Does acute artificial intelligence assistant soften the blow? The impact of cuteness on customer tolerance of assistant service failure*, vol. 37, no. 4, 1994.

[46] Bo Pang, L. Lee, and S. Vaithyanathan, "Sentiment classification using machine learning techniques," *Proceedings of EMNLP*, vol. 79, 2002.

[47] M. Mariani and M. Borghi, "Customers' evaluation of mechanical artificial intelligence in hospitality services: a study using online reviews analytics," *International Journal of Contemporary Hospitality Management*, vol. 33, no. 11, pp. 3956–3976, 2021.

[48] E. H. Manser Payne, A. J. Dahl, J. Peltier, and J. Peltier, "Digital servitization value co-creation framework for AI services: a research agenda for digital transformation in financial service ecosystems," *The Journal of Research in Indian Medicine*, vol. 15, no. 2, pp. 200–222, 2021.