Artificial Intelligence & Robotics Development Journal
Volume 3, Issue 1, pp 214-228, January 2023, https://doi.org/10.52098/airdj.202366
ISSN: 2788-9696 Received: 1/2/2023 Revised: 12/2/2023 Accepted: 14/2/2023

A Method for SMS Spam Message Detection Using Machine Learning

Vaman Ashqi Saeed[1]
[1] 1Technical College of Administration, Information Technology Management Department, Duhok, Kurdistan Region of Iraq
* Corresponding author: Vaman Ashqi Saeed [1], vamanatroushi@gmail.com

Abstract

In recent years, it has become increasingly common for individuals to connect with their relatives and friends, read the most recent news, and discuss various social topics using online social platforms such as Twitter and Facebook. As a consequence of this, anything that is considered spam can quickly spread among them. Spam identification is widely regarded as one of the most significant problems in text analysis. Previous studies on detecting spam concentrated primarily on English-language content and paid little attention to other languages. The information gathered by the University of California; Irvine served as the basis for developing our spam detection technology (UCI). This study investigates the effectiveness of various supervised machine learning algorithms, such as the J48, K-Nearest Neighbors (KNN), and Decision Tree (DT), in identifying spam and ham communications. SMS spam is becoming more widespread as the number of internet users continues to rise, and many businesses disclose their customers' personal information. E-mail spam filtering is the progenitor of SMS spam filtering, which inherits many features. We evaluate the proposed method based on accuracy, recall, and precision. Experiments showed that the Decision Tree method obtained higher accuracy than other machine learning classifiers.

Keywords: Short Message Service (SMS); Data Cleaning; Spam Detection; Machine Learning.

1. Introduction

Short Message Service (SMS) on mobile devices is utilized by humans for the purposes of communication and business. SMS has recently become the data service that is utilized the most throughout the world. Given that 690 billion SMS messages are sent monthly and the world sent 8.3 trillion SMS messages in 2017, it is clear that SMS plays a significant role in business conversations. The development of new technologies has been a driving force behind the rise of social media platforms in the internet age. There are many different programs and websites that feature user-generated content, such as social networks, bulletin boards, blogs, instant messaging tools, news commentary, and so on (Bakshy et al., 2012). Due to the unfortunate events surrounding the COVID-19 quarantine, there has been a significant increase in the amount of user-generated content across all platforms (Luo et al., 2021; Yousif et al, 2021). The enormous amount of user-generated text material that is created each day includes, but is not limited to, illegal and harmful data. Some examples of this type of data include ads, fraud, phishing, and other forms of spam (Saini & Yousif, 2021). The large flood of various sorts of spam has generated a variety of serious difficulties, some of the most prominent of which are the polarization of viewpoints, decreases in users' online interaction time, and a decreased quality of the information that is delivered (Rao et al., 2021).

This occurrence is an indication that spam is overwhelming the whole network, which causes cyber citizens to experience a degree of annoyance. Single-modal spam filtering systems have a high detection rate for text spam and picture spam; nevertheless, spammers may put garbage material into the multi-modal section of an email, which we term hybrid spam, to lower the detection rate of single-modal spam filtering systems, and eventually achieve the aim of escaping detection. This is because hybrid spam is a combination of text spam and image spam (Gao et al., 2018). This is done in order to avoid being caught by single-modality spam filtering systems, which have a high detection rate for text spam and picture spam respectively. Hybrid spam is more destructive than regular spam because it includes more information than traditional spam and because it takes more network bandwidth and storage space for forwarding and delivery to mailbox servers. The reason for this is because hybrid spam, as opposed to typical spam, includes a greater amount of information. Additionally, viruses or undesired material that are supplied through hybrid spam are more difficult to discover, which offers tremendous information security risks for the communication of persons (Asaad & Saeed, 2022).). Therefore, it is of the utmost importance to acquire the knowledge necessary to correctly recognize hybrid spam (Jain et al., 2019).

SMS has become more popular as a method of communication between attackers and their targets because of its ease. The majority of people who fall prey to phishing and smishing scams are those who possess mobile devices, specifically smartphones. By delivering a link to victims or making direct contact with them through SMS messages, the attackers seek to obtain confidential information from users, such as credit card numbers, bank account data, and other information (Jain et al., 2020). The short messaging service (SMS), in which messages must be delivered in accordance with communication standard protocols, is currently one of the most popular methods of communication between large groups of people. As a consequence of this, there is a need for text classification algorithms that may be used as a component of the process of classifying the communications as either ham or spam messages. Ham communications, on the other hand, are the ones that are created by legitimate individuals. Spam messages are not appealing in any way. As a result, unwanted communications, also known as spam messages, have to be recognized and removed from the mobile station as soon as they arrive there. Some examples of spam messages are those produced by organizations that offer promotional services. In addition to being an annoyance, the extra effects of spam text messages are also time-consuming, costly, resource-intensive, and reduce network capacity. In any event, the availability of spam filtering software that can recognize SMS spam is restricted. In addition, there is the possibility of an additional misclassification problem occurring when ham communications are deleted and prohibited because they are considered spam (Almeida et al., 2011). SMS spam typically has an effect on a group of people and is spread over mobile networks; email spam, on the other hand, is transmitted through the World Wide Web. Both types of spam cause users to get irritated, which in turn leads to the humiliation of the service's performance. Despite this, a variety of approaches to email spam filtering and classification have been adapted into methods for detecting spam in SMS messages (Cormack G., 2008). Additionally, the researchers and analysts working on the identification of SMS spam are confronted with a variety of challenges, one of which is the limitation of the datasets that are freely accessible. On the other hand, not all of the approaches to filtering spam in emails are very effective when applied to the detection of spam in SMS (Ji & Zhang, 2015).

We created a method for the detection of spam making use of machine learning algorithms as part of the scope of this research project. This method would classify communications as either ham or spam. For the purpose of validating the current research, the SMS spam collection dataset was taken into consideration. The dataset was split into two groups: thirty percent for use in testing the prediction models, and seventy percent for use in training those models. During the process of analyzing the proposed study, the assessment measures for performance, which include

specificity, accuracy, and sensitivity, were thought about and deliberated upon. The results of the trials that were carried out offered conclusive evidence that the planned research had been successfully carried out, hence proving the truth of the aforementioned assertion.

2. Literature Review

The scope of these studies was confined to the use of machine learning and deep learning-based models only, and they were published throughout the course of the past several years by computer scientists. It first separates the texture features from the attribute features of the image part of the email, then classifies both sets of features using the SVM method to obtain two classification probability values. Then incorporates both sets of features into the model that was designed by the SVM method once more to determine whether or not the email should be considered spam. This method was proposed by Xu et al. (Xu et al., 2011) who were working on the design of a spam filtering system. These techniques can deal with picture spam, but not word spam or hybrid spam; only images can be handled by them. Almeida introduced a processing approach (Almeida et al., 2016) in order to improve the efficiency of classification techniques when applied to text messages, it is necessary to first normalize and then expand the text messages. This approach was developed with the intention of enhancing the effectiveness of classification algorithms. The technique that has been proposed relies heavily on lexicography, semantic dictionaries, and several other types of semantic analysis and disambiguation as its primary building blocks. The primary objective was to expand the original content while simultaneously cutting down on performance-degrading variables such as redundancies and inconsistencies by standardizing the words and developing new qualities. This was accomplished by standardizing the terms.

Almeida (Almeida et al., 2011) presented a brand-new SMS spam collection that was unencrypted and available to the public. They showed that SVM performed better than other classifiers that were evaluated, including Basic Naive Bayes, Flexible Bayes, Minimum Description Length, KNN, and others. On the other hand, traditional methods of machine learning often demand a considerable amount of feature engineering. A comprehensive exploratory data analysis performed on the dataset, followed by a straightforward process for dimension reduction, calls for a substantial amount of time upfront, and it is difficult to reuse this labor later on. In addition, the number of parameters that may be adjusted by machine learning approaches is restricted, which places a ceiling on how well the model can match the data. Detecting smishing messages using a feature-based technique was one of the authors' recommended methods Jain and Gupta (Jain & Gupta, 2019). The method utilizes the extraction of 10 attributes, all of which the

authors say are able to differentiate between genuine and fake communications. After that, the characteristics were applied to a dataset that had previously been used as a benchmark, and five different classification techniques were used in order to evaluate how well the suggested method worked. The results of the experimental assessment demonstrated that the model has a true positive rate of 94.20% and an accuracy rate of 98.74% overall when it comes to detecting smishing messages. CNN and LSTM are two examples of different types of deep neural network models. Ghourabi (Ghourabi et al., 2020) presented a hybrid model for recognizing text messages written in Arabic or English that is based on the combination of these two models. This model may be used to recognize communications written in either language. According to the findings, the CNN-LSTM model obtained an accuracy score of 98.37%, which is higher than the scores obtained by other methods such as the Support Vector Machine, K-Nearest Neighbors, Multinomial Naive Bayes, Decision Tree, Logistic Regression, Random Forest, AdaBoost, Bagging Classifier, and Extra Trees. Specifically, the CNN-LSTM model scored higher than the scores obtained by the Support Vector Machine, K-Nearest Neighbor In the research that was conducted out by Bassiouni (Bassiouni et al., 2018), many classifiers were tested out in an attempt to filter emails that were collected from the Spambase UCI dataset, which consisted of 4601 unique instances. They began by preprocessing the data, and then they selected the features using an algorithm called Infinite Latent Feature Selection (ILFS). They were able to classify emails with an accuracy of 95.45% by using the Random Forest (RF) algorithm, while the remaining classifiers (Artificial Neural Network (ANN), Logistic Regression, Support Vector Machine (SVM), Random Tree, K-Nearest Neighbors (KNN), Decision Tree (DT), Bayes Net, Naive Bayes (NB), and Radial Basis Function (RBF) scored 92.4%, 92.4%, 91.8%, and 91.5%.

3. Proposed Method

In this part, we will provide a comprehensive explanation of the model that we have presented. Processing the gathered text messages (SMSs) and using a machine learning algorithm in order to categorize them and determine which ones are regarded as spam or phishing communications is the primary objective of this detection system. In Figure 1, we show the overall structure of the model that has been suggested. Traditional machine learning algorithms, such as J48, KNN, and Decision Tree (DT) are the ones that we have decided to implement in our model. The purpose of this is to make an attempt, using a few different phases, to categorize SMS. The suggested system starts out by removing any superfluous information from the text messages that are received. Following this, a pre-processing job will be done on these messages in order to represent the textual data in a suitable form. After that, the data from this form will be sent to the machine learning algorithms that are going to be utilized. After the data preparation has been

finished, the classification algorithms will be performed on this data in order to make a distinction between messages that are considered spam and messages that are not considered spam.

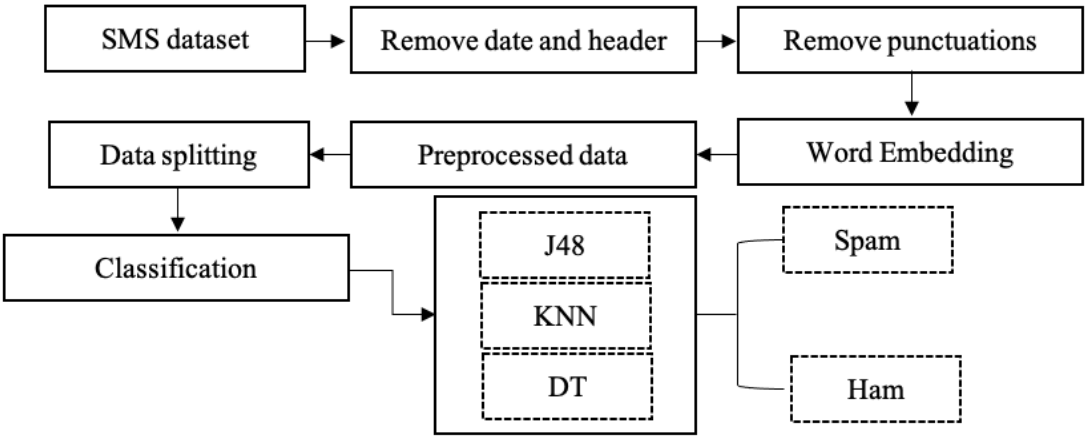

Figure 1: Proposed Framework for SMS Spam Detection

3.1 Dataset

In the experimental efforts, the dataset, which is made up of text messages written in English, is utilized. Downloadable copies of the dataset are available from the UCI repository (Almeida et al., 2011). These text messages were taken from a public forum that was located in the United Kingdom. According to the information shown in Table 1, the dataset contains a total of 5574 SMS messages, 747 of which are classified as spam and 4827 of which are classified as ham.

Table 1. The statistics related to the dataset.

Class	Number of samples
Spam	747
Ham	4827
Total	5574

3.2 Preprocessing

The text pre-processing comes in as the second phase of the model. machine learning (ML) approaches can only work with numerical data and are unable to process textual input is a challenge for us that we are attempting to address

at this stage of the process. As a consequence of this, the objective is to transform the textual data obtained from the short messages into a format that is simple to understand and that can be comprehended by machine learning (ML) algorithms. Because of the unique qualities of the algorithms for machine learning that were employed in this tactic, this technique was successful, and we have come up with two distinct solutions for this particular challenge.

Pennington (Pennington et al., 2014) developed the word embedding approach, which is presently being utilized in this investigation. This method is used to transform the sequence of words that have previously been pre-processed into a vector representation known as embedding space. This space has more dimensions than the typical word data that is used to train classifiers. After adding padding and truncating the data, the next step is to utilize the word embedding approach to create extra dimensions for the data that is in sequence. The embedding size is set to 32. After the data had been preprocessed, we divided it into a training set consisting of thirty percent and a test set consisting of seventy percent. As a direct consequence of this, the training data set includes a total of 238 spam messages in addition to the 509 legitimate communications.

3.3 Machine Learning

In machine learning (ML), learning structure from data is accomplished by a combination of representing the data, evaluating the data, and optimizing the representation. It is anticipated that ML, which is also known as data mining or predictive analytics, will be the driving force behind the subsequent major wave of technological advancement (Abayomi et al., 2022). Its application is helpful in accomplishing tasks linked with artificial intelligence, for instance, identification, recognition of voice, diagnosis, recognition of face, robot control, named entity identification, ranking, and other similar activities. On the other hand, the representation of text documents is a very important factor for machine learning techniques for text categorization. According to the research that has been conducted, ML may be broken down into three distinct groups. These classes include supervised learning, unsupervised learning, and semi-supervised learning. According to (Learned-Miller, 2014) supervised learning is the process of learning from a derived function (classifier) based on training data that includes a collection of samples that serve as both the input object (vector) and the predicted result (supervisory signal). Additionally, supervised machine learning forecasts the correct output value for each valid input, allowing one to derive inferences from training data in a method that is both useful and manageable. Neural networks, decision trees, and multilayer perceptron are a few examples of models that fall under the category of being supervised (Khamis & Yousif, 2022). This learning strategy is particularly expensive since the bulk of the data is provided in an unlabeled format. As a result, the cost of labeling this data to account for past

knowledge is also quite expensive. Unsupervised learning is essentially the process of learning using unlabeled data, in which input classes are clustered only on the basis of their statistical features (Yousif & Saini, 2020). In order to accomplish the task at hand, semi-supervised learning makes utilize of both labeled and unlabeled data. It involves contrasting a substantial number of data points with no labels with a relatively small number of data points with labels. This particular kind of machine learning has garnered a lot of interest in a variety of study fields, such as web mining and so on and so forth. The purpose of this research is to evaluate the effectiveness of machine learning classifiers using MATLAB and WEKA as their respective environments. There is a need to examine some of the ML classifiers that were employed in this work, such as J48, K-Nearest Neighbor Decision (KNN), and Decision Tree, in order to provide a comprehensive overview of the relevant previous research (DT).

a) J48

The C4.5 algorithm developed by Quinlan implements J48 in order to produce a pruned C4.5 decision tree. Every part of the information is broken down into more manageable subsets before a determination is made. J48 has a look at the standardized data gain that, in reality, results in the splitting of the information by selecting a characteristic. In a nutshell, the data gathered via rigorous standardization is put to use. The method will yield the subsets that are considered minor. When a subset can be placed in a class that is comparable in all cases, the split techniques come to an end. J48 constructs a decision node by making use of the anticipated estimations provided by the class. A J48 decision tree is equipped to handle specific features, lost or incomplete attribute estimations of the data, as well as variable attribute prices. Here, increasing the precision may be accomplished by pruning (Venkatesan & Velmurugan, 2015). The J48 iteration of the C4.5 technique has numerous more features, some of which include the capacity to create rules, the ability to account for missing data, the ability to prune decision trees, continuous attribute value ranges, and so on. You may find an open-source Java implementation of the C4.5 method in the WEKA data mining program. This implementation is known as J48. Users of J48 are given the opportunity to categorize data either utilizing decision trees or rules that are generated from those trees.

The Algorithm

1. If the examples correspond to the identical class, therefore the class will be labeled with the identical class as the instances if they correspond to the same class.

---

2. The potential data will be calculated for each attribute, and the gain in the data will be determined based on the results of the test performed on the attribute.

3. In the end, the characteristic with the most favorable current selection parameter will be selected as the winner.

b) K-Nearest Neighbor Decision (KNN)

An instance-based learning technique that classifies elements in the resource space based on the k-training samples that are the closest to them is the KNN algorithm, which was proposed by Aha (1997). In addition to serving as the primary adjustment parameter for the KNN algorithm, these data are also crucial in the process of spatial forecasting. K-Nearest Neighbors (KNN) is a typical classification method used in applications involving remote sensing and data mining. It is also commonly used for mapping burnt regions. KNN is a non-parametric method of principal component analysis (PCA) that does not presume anything about the primary data set. When identifying processes of change in territory, such as floods and fires, for which there is little to no prior knowledge of the distribution of data, this is a crucial aspect to take into consideration. In KNN, a pixel's membership in a class is determined by the pixels that are spectrally nearest to it and whose class identities are already known (Thanh Noi & Kappas, 2017). A pixel's membership in a class is determined even if its class is unknown. The basic structure of the KNN algorithm is shown in Figure 2.

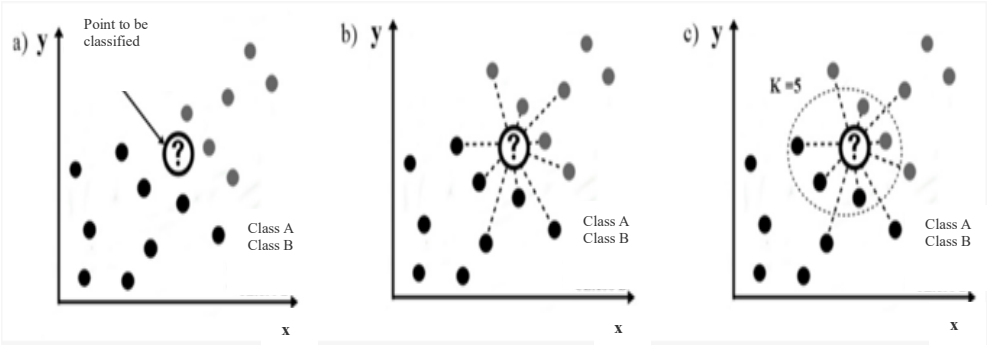

Figure 2. Classification-based K-Nearest Neighbor (Thanh Noi and Kappas, 2017)

In the beginning, the parameter k, which stands for the number of neighbors that are the most immediate, needs to be chosen. The number of neighbors is determined by the value of this option. In a binary group with k equal to five, the Euclidean distance is used to determine which five points are the closest to one another. It is feasible to determine which class an individual point most closely resembles by calculating the smallest distance between the point that is

being categorized and each of the other points in the data set. This is the point that is being classified. As a result, the classification is finished, and the previously unclassified point is now classified. As the primary KNN adjustment parameter, the parameter k is critically important to the operation of the KNN and plays a significant role in its performance ruined by the value of this option. In a binary group with k equal to five, the Euclidean distance is used to determine which five points are the closest to one another. It is feasible to determine which class an individual point most closely resembles by calculating the smallest distance between the point that is being categorized and each of the other points in the data set. This is the point that is being classified. As a result, the classification is finished, and the previously unclassified point is now classified. As the primary KNN adjustment parameter, the parameter k is critically important to the operation of the KNN and plays a significant role in its performance. In this work, we investigated a variety of k values, ranging from five to twenty, in order to identify the best parameter for the KNN classifier based on the lowest estimate of the root mean square error (RMSE) (Pacheco et al., 2021). We used a variety of data subsets to conduct these tests.

c) Decision Tree (DT)

A decision tree technique is a decision-making aid that takes the shape of a tree structure and is formed with the help of characteristics that are entered. This kind of decision-making aid is also sometimes referred to as a DT classifier. The fundamental objective of this specific classifier is to build a model that, on the basis of a number of input characteristics, is able to generate predictions concerning the variables that are of interest to us. This classifier is suitable for a broad number of various sorts of applications (Acharya et al., 2012) because of the ease with which decision rules may be derived from any particular collection of input data.

The DT technique is a method for nonparametric supervised learning that may be used for the purpose of finding solutions to problems. The DT technique is a method for nonparametric supervised learning that may be used for problem-solving, including regression and classification. This method has the potential to be applied in both of these areas. The DT model is shown in Figure 3 to illustrate how it may be interpreted as a representation of a structure. This paradigm consists of three nodes: the root node, the division node, and the leaf node. The root node is the most fundamental of the three. Each and every internal node is, in reality, nothing more than a test that is being carried out on some property. Every division that results from that test is the final product, and every leaf node remembers the class label that belongs to its parent. The building of the tree really starts at the root node, which is the point at which

it is named. At the outset of the procedure, an attribute is chosen to be transferred to the root node, where it will remain for the entirety of the operation. After that, a division is carried out for each of the possible values that may be used. This causes the development of subgroups inside the dataset, one for each of the potential values that may be discovered in the property. These subgroups are named after the properties themselves. The process of the tree is carried out in a recursive manner for each division, and the only instances that are taken into consideration are the ones that make it to the branch. When all of the instances on a node have the same classification, it is feasible to halt the advancement of the tree and come to a complete stop. Entropy and classification error are often the two measures that are considered to be the best options when trying to identify which tree partition is the best (Lipinski et al., 2020).

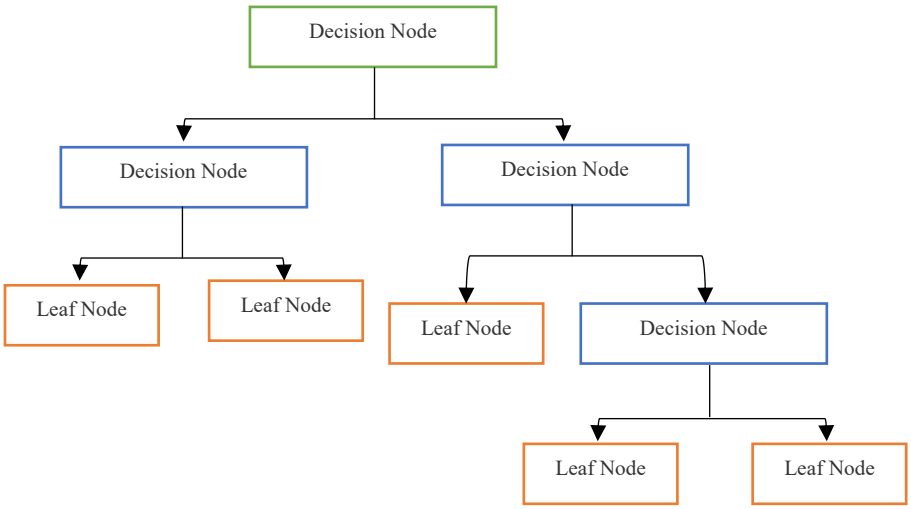

Figure 3. Decision Tree-based Classification (Lipinski et al., 2020)

4. Experimental Results

This section analyzes the performance of the model in identifying spam messages acquired from the UCI repository and compares the findings with machine learning classifiers. In addition, the usefulness of the model in identifying spam messages is investigated in this part of the article. Initially, the dataset was split into two unique subsets. Seventy percent of the messages were assigned for utilization in the training phase, while the remaining thirty percent were kept for utilization in the validation and testing stages. For the purpose of performance evaluations, this research makes use of the widely used dataset from the UCI repository. This imbalanced dataset includes a total of 5574 English SMS messages, 747 of which are considered spam while the remaining 4827 are considered ham. These messages were obtained from Grumbletext, which is a public forum located in the United Kingdom (www.grumbletext.co.uk) (Tagg C, 2009).

Accuracy, recall, and precision is the three well-known and convincing evaluation measures for categorization. The parameters that they use for the computation of metrics are shown in Table 2. Accuracy is the most essential component to consider with regard to these measurements (Forman G, 2003). The proportion of spam SMS messages that can be properly predicted among all SMS is referred to as accuracy. The portion of the total number of relevant instances that are successfully recovered is referred to as sensitivity, and it is the same thing as recall. The percentage of relevant instances within the total number of retrieved instances is referred to as precision.

Table 2: Confusion matrix

Actual	**Predicted**	
	Negative	**Positive**
Negative	**True Negative (TN)**	**False Negative (FN)**
Positive	**False Positive (FP)**	**True Positive (TP)**

$$\text{Accuracy} = \frac{\text{TN+TP}}{TN+TP+FN+FP} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{TP+FN} \quad (3)$$

Text pre-processing techniques, for example, the removal of stop words, and removing punctuation, were applied for the purpose of cleaning the dataset in order to apply machine learning classifiers such as J48, K-Nearest Neighbors (KNN), and Decision Tree (DT). This was done in order to apply these machine-learning classifiers.

According to what is shown in Table 3, machine learning classifiers have varying degrees of precision, recall, and accuracy when it comes to classifying messages as spam or not spam.

Table 3: Achieved Results based on Using Traditional Machine Learning classifiers

Classifier	Class	Accuracy	Recall	Precision
J48	Spam	81.32%	86.71%	71.09%
	Harm	93.33%	84.56%	95.24%
KNN	Spam	88.43%	82.42%	73.16%
	Harm	97.08%	93.52%	93.91%
DT	Spam	95.71%	93.11%	92.17%
	Harm	98.40%	97.87%	97.65%

This study compares the results obtained by the proposed method with those achieved by other ML approaches, such as naive Bayes (NB), support vector machines (SVM), Random Forest (RF), and Decision Tree (DT), in order to demonstrate the efficacy of the method (CNN). The comparison of the results is presented in Table 4. As was previously said, the dataset from the UCI is used to evaluate each of the approaches that are stated in Table 4. (Almeida et al., 2011), the results of the NB, SVM, and NMF techniques as well as the LDA method are shown. In terms of the accuracy %, it is noteworthy that RF gets the best performance and surpasses the other algorithms that were analyzed in Ref (Sjarif et al., 2019). It obtains an accuracy of 97 while correctly classifying 97.5% of the samples.

Table 4: Comparison between Proposed Method and Previous Studies

Method	Classifier	Accuracy	Recall	Precision
(Almeida et al., 2011)	NB	84.2	97.2	95
	SVM	93.6	97.7	97
	NMF	91.7	97.6	96
	LDA	90.4	97.6	96
(Sjarif et al., 2019)	NB	97.06	-	97
	KNN	91.19	-	89

	SVM	87.49	-	82
	DT	96.57	-	97
	RF	97.5	-	97
Proposed	DT	97.05	95.49	94.91

5. Conclusion

Within the scope of this investigation, we offer classification models for SMS spam that are derived from machine-learning methods such as J48, KNN, and DT. In order to pre-process the SMS text data, we made use of a variety of techniques, including data cleaning and the word embedding approach. In addition to that, we assign categories to the dataset using several machine learning methods. In the end, we assessed the models using a test set that was taken from the SMS spam dataset. According to the findings, the performance of the DT model is superior to that of other models, achieving an accuracy of 97.05%. This research was a case study that focused on constructing a model for the categorization of SMS spam based on machine learning methods. In upcoming projects, one of our primary goals is to find ways to improve the performance of the model by amassing additional data drawn from a wider variety of sources. Our goal is to create a model that can be implemented in the real world to provide assistance to individuals.

Acknowledgment

The research leading to these results has received no Research Project Grant Funding.

References

[1]. Abayomi-Alli, O., Misra, S., & Abayomi-Alli, A. (2022). A deep learning method for automatic SMS spam classification: Performance of learning algorithms on indigenous dataset. *Concurrency and Computation: Practice and Experience*, e6989.
[2]. Acharya, U. R., Molinari, F., Sree, S. V., Chattopadhyay, S., Ng, K. H., and Suri, J. S., Automated diagnosis of epileptic EEG using entropies. Biomed. *Signal Process. Control* 7(4):401–408, 2012.
[3]. Aha, D.W. Artificial Intelligence Review. *Lazy Learn*. 1997, 11, 1–6.
[4]. Almeida, T. A., Hidalgo, J. M. G., & Yamakami, A. (2011, September). Contributions to the study of SMS spam filtering: new collection and results. *In Proceedings of the 11th ACM symposium on Document engineering* (pp. 259-262).
[5]. Almeida, T. A., Hidalgo, J. M. G., & Yamakami, A. (2011, September). Contributions to the study of SMS spam filtering: new collection and results. *In Proceedings of the 11th ACM symposium on Document engineering* (pp. 259-262).
[6]. Almeida, T. A., Silva, T. P., Santos, I., & Hidalgo, J. M. G. (2016). Text normalization and semantic indexing to enhance instant messaging and SMS spam filtering. *Knowledge-Based Systems*, 108, 25-32.
[7]. Asaad, R. R., & Saeed, V. A. (2022). A Cyber Security Threats, Vulnerability, Challenges and Proposed Solution. Applied computing Journal, 227-244.

[8]. Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012, April). The role of social networks in information diffusion. *In Proceedings of the 21st international conference on World Wide Web* (pp. 519-528).
[9]. Bassiouni, M., Ali, M., & El-Dahshan, E. A. (2018). Ham and spam e-mails classification using machine learning techniques. *Journal of Applied Security Research*, 13(3), 315-331.
[10]. Cormack, G. V. (2008). Email spam filtering: A systematic review. *Foundations and Trends® in Information Retrieval*, 1(4), 335-455.
[11]. Forman, G. An Extensive Empirical Study of Feature Selection Metrics for Text Classification George. J. Mach.Learn. Res. 2003, 1, 1289–1305.
[12]. Gao, J., Lanchantin, J., Soffa, M. L., & Qi, Y. (2018, May). Black-box generation of adversarial text sequences to evade deep learning classifiers. *In 2018 IEEE Security and Privacy Workshops (SPW)* (pp. 50-56). IEEE.
[13]. Ghourabi, A., Mahmood, M. A., & Alzubi, Q. M. (2020). A hybrid CNN-LSTM model for SMS spam detection in Arabic and english messages. *Future Internet*, 12(9), 156.
[14]. Jain, A. K., & Gupta, B. B. (2019). Feature-based approach for detection of smishing messages in the mobile environment. *Journal of Information Technology Research (JITR)*, 12(2), 17-35.
[15]. Jain, A. K., Yadav, S. K., & Choudhary, N. (2020). A novel approach to detect spam and smishing SMS using machine learning techniques. *International Journal of E-Services and Mobile Applications (IJESMA)*, 12(1), 21-38.
[16]. Jain, G., Sharma, M., & Agarwal, B. (2019). Optimizing semantic LSTM for spam detection. *International Journal of Information Technology*, 11(2), 239-250.
[17]. Ji, H., & Zhang, H. (2015). Analysis on the content features and their correlation of web pages for spam detection. *China Communications*, 12(3), 84-94.
[18]. Khamis, Y., & Yousif, J. H. (2022). Deep learning Feedforward Neural Network in predicting model of Environmental risk factors in the Sohar region. Artificial Intelligence & Robotics Development Journal, 201-2013.
[19]. Learned-Miller, E. G. (2014). Introduction to supervised learning. I: Department of Computer Science, University of Massachusetts, 3.
[20]. Lipinski, P., Brzychczy, E., & Zimroz, R. (2020). Decision tree-based classification for Planetary Gearboxes' condition monitoring with the use of vibration data in multidimensional symptom space. *Sensors*, 20(21), 5979.
[21]. Luo, Y., & Xu, X. (2021). Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *International Journal of Hospitality Management*, 94, 102849.
[22]. Pacheco, A. D. P., Junior, J. A. D. S., Ruiz-Armenteros, A. M., & Henriques, R. F. F. (2021). Assessment of k-nearest neighbor and random forest classifiers for mapping forest fire areas in central portugal using landsat-8, sentinel-2, and terra imagery. *Remote Sensing*, 13(7), 1345.
[23]. Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).
[24]. Rao, S., Verma, A. K., & Bhatia, T. (2021). A review on social spam detection: challenges, open issues, and future directions. *Expert Systems with Applications*, 186, 115742.
[25]. Saini, D. K., & Yousif, J. H. (2021). Vulnerability and Attack Detection Techniques: Intrusion Detection System. In Cybersecurity (pp. 17-26). CRC Press.
[26]. Sjarif, N. N. A., Azmi, N. F. M., Chuprat, S., Sarkan, H. M., Yahya, Y., & Sam, S. M. (2019). SMS spam message detection using term frequency-inverse document frequency and random forest algorithm. *Procedia Computer Science*, 161, 509-515.
[27]. Tagg, C. A Corpus Linguistic Study of SMS Texting. Ph.D. Thesis, University of Birmingham, Birmingham, UK, 2009.
[28]. Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, 18(1), 18.
[29]. Venkatesan, E., & Velmurugan, T. (2015). Performance analysis of decision tree algorithms for breast cancer classification. *Indian Journal of Science and Technology*, 8(29), 1-8.
[30]. Xu, C., Chiew, K., Chen, Y., & Liu, J. (2011). Fusion of text and image features: A new approach to image spam filtering. *In Practical Applications of Intelligent Systems* (pp. 129-140). Springer, Berlin, Heidelberg.
[31]. Yousif, J. H., Khan, F. R., Al Jaradi, S. N., & Alshibli, A. S. (2021). Exploring the Influence of Social Media Usage for Academic Purposes Using a Partial Least Squares Approach. Computation, 9(6), 64.
[32]. Yousif, J. H., & Saini, D. K. (2020). Big Data Analysis on Smart Tools and Techniques. In Cyber Defense Mechanisms (pp. 111-130). CRC Press.