LONDON
METROPOLITAN
UNIVERSITY

islington college
(इस्लिङ्टन कलेज)

Module Code & Module Title

CU6051NI Artificial Intelligence

Assessment Weightage & Type

25% Individual Coursework

2023-24 Autumn

Student Name: Amit Baniya

London Met ID: 21039840

College ID: NP01CP4A210096

Assignment Due Date: Wednesday, December 20, 2023

Assignment Submission Date: Wednesday, December 20, 2023

I confirm that I understand my coursework needs to be submitted online via MySecondTeacher under the relevant module page before the deadline in order for my assignment to be accepted and marked. I am fully aware that late submissions will be treated as non-submission and a marks of zero will be awarded.

Abstract

Spam has been one of the biggest problems in our lives, in this document we will be seeing, how AI and machine learning have helped us in different sectors. The AI industry is growing rapidly each year, and it is the future. Therefore, there are ways that we can implement it in our lives for a better living style. We will be discussing the small part of machine learning usage which is spam filtering. There are many ways to filter spam whether it is a traditional way by not using machine learning or by using machine learning. Nowadays, all the major companies use machine learning for spam filtering. However, the impact of spam is still high as spam covers half of our mail. Therefore, we will be researching about what are the possible machine learning algorithms that can be used to identify spam. After researching about the algorithms from different research papers done by others. We will be selecting three algorithms, learning about them, and seeing how we will be able to implement them to detect spam. We will be using a dataset found from the website called Kaggle which is one of the best places to find datasets. We will be laying a foundation in this document to implement the algorithms to detect spam and in the future, we will be doing another project where we will test the algorithms by modeling dataset and comparing those algorithms to each other side by side in order to understand the effectiveness of those algorithms.

Table of Contents

Table of Figures

List of Tables

Table of equations

1. Introduction

Artificial intelligence (AI) is the abilities of a computer or robot to carry out operations typically performed by intelligent entities by computer or robot themselves. The phrase is commonly used to describe the attempt of creating computer systems that include human-like cognitive functions, like reasoning, meaning-finding, generalization, and experience-based learning. It has been shown that computers can perform extremely complicated jobs, including finding proofs for mathematical theorems or playing chess, with remarkable proficiency ever since the digital computer was developed in the 1940s. Nevertheless, despite ongoing improvements in computer memory and processing power, no software can yet fully simulate human adaptability over a larger range of areas or in jobs requiring a great deal of common knowledge. However, in certain limited applications, such as medical diagnosis, computer search engines, voice, or handwriting recognition, and chatbots, artificial intelligence has evolved to the point where some programs can perform at the same levels as humans. (Copeland, 2023)

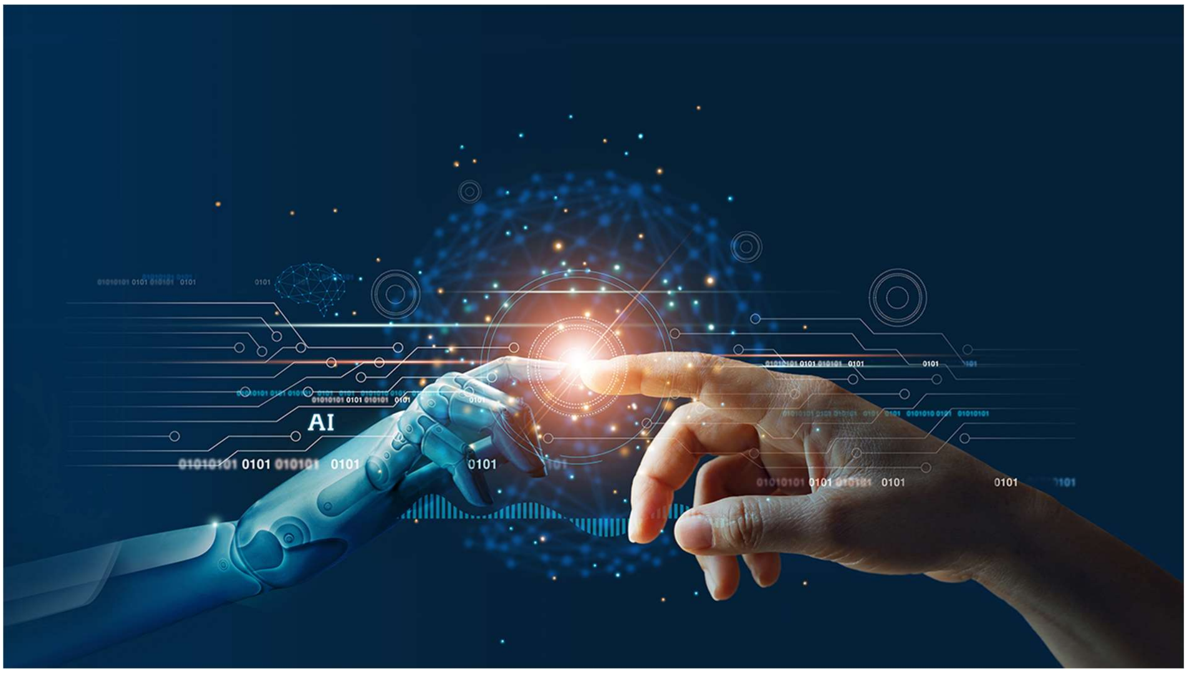


Figure 1 - AI and human (poptika/shutterstock.com)

Since defining intelligence can be difficult, AI specialists usually distinguish between strong and weak AI.

Strong AI

Strong artificial intelligence, sometimes referred to as artificial general intelligence, is the ability of a machine to solve problems like those on which it has never been trained. The robots from different movies are examples of this type of artificial intelligence (AI). There isn't really any AI like this yet. Though the complexity of gaining this level of achievement hasn't decreased over time, strong AI, in contrast to weak AI, indicates a computer with a full set of cognitive abilities and an equally large selection of use cases. (Schroer, 2023)

Weak AI

A form of artificial intelligence that is limited to a particular or narrow field is known as weak artificial intelligence (AI), also known as narrow AI. Human cognition is simulated by weak AI. It can help society by automating laborious tasks and doing data analysis in ways that people aren't always able to. We can compare weak AI to strong AI as the weak AI are theoretical kinds of computer intelligence comparable to human intelligence. (Frankenfield, 2022)

Some examples of Weak AI are as follows:

- Self-driving cars
- Google search
- Email Spam filters
- Recommendation systems
- Smart assistants like google home, Alexa etc.

Deep Learning vs Machine Learning

While the words "deep learning" and "machine learning" are often used in discussions about artificial intelligence, they should not be used synonymously. Machine learning, a branch of artificial intelligence, includes deep learning as one of its forms.

Deep Learning

A kind of machine learning called deep learning processes inputs via a neural network architecture that takes reference from biology. The data is processed through a number of hidden layers in neural networks, which enable the machine to learn "deeply," form connections, and weight input for optimal outcomes. (Schroer, 2023)

Machine Learning

Artificial intelligence (AI)'s machine learning (ML) field gives computers the capacity to autonomously learn from data and past experiences, finding patterns to generate predictions with little to no human input. Without explicit programming, computers can function independently thanks to machine learning techniques. Applications for machine learning are fed fresh data and have the ability to learn, grow, evolve, and adapt on their own. (Kanade, 2022)

Types of Machine Learning Techniques

There are three types of machine learning techniques, and they are:

Supervised Learning

The target value is chosen beforehand by supervised learning algorithms, which map labelled inputs to known outputs. Machine learning models are trained using supervised learning techniques, which require outside supervision. Thus, the term "supervised." To achieve the intended result, they require instruction and further details. Supervised learning can be used in filtering spams, prediction weather conditions, share market etc. (Menon, 2023)

Unsupervised Learning

A machine learning that trains algorithms using unlabelled data. Data without labels lacks a specific output variable. After learning from the data and identifying its patterns and features, the model produces an output. In order to figure out the output, unsupervised learning analyzes trends and patterns in the data. In considering this, the model attempts to label the data using the attributes of the input data. Unsupervised learning approaches do not require supervision during the training stage in order to construct models. They make predictions and pick up knowledge on their own. They can be used to segment customers in businesses, identification of accident-prone areas, classification of heavenly bodies. (Menon, 2023)

Reinforcement Learning

A machine can be trained through learning to maximize its rewards and take the appropriate actions in a given scenario in this kind of learning technique. It generates actions and rewards by utilizing an agent and an environment. There are start and finish states for the agent. However, there may be multiple routes, much to a maze, to arrive at the destination. This learning method does not have a predetermined target variable. It can be used to make AI in games, or make robots do human tasks. (Menon, 2023)

Types of AI

AI are divided in four types, and they are as follows:

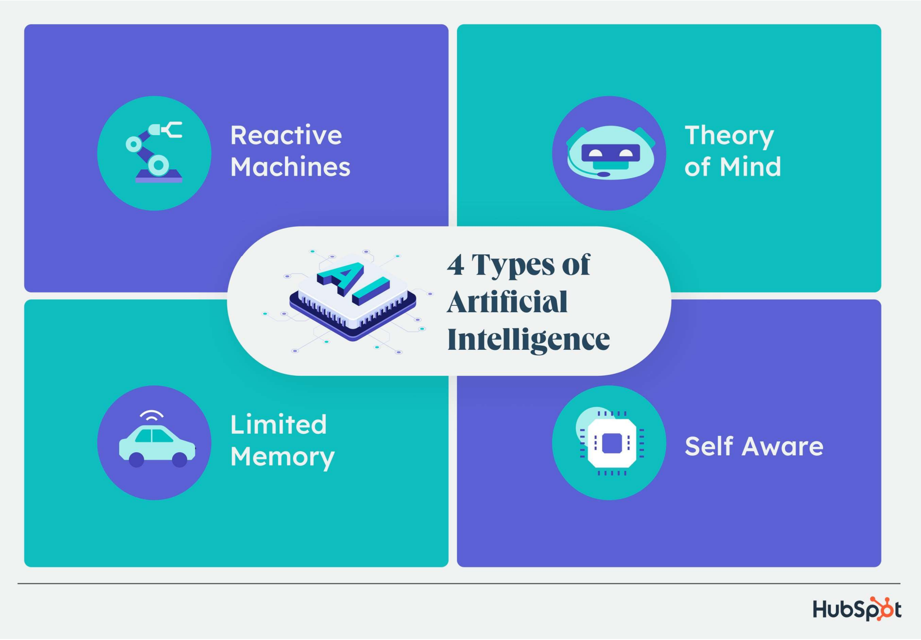


Figure 2 - Types of AI - (Erin Rodrigue,Flori Needle, 2023)

Reactive Machines

Reactive machines, as the name indicates react and respond to various inputs. It achieves this without the aid of recollection or a deeper awareness of the situation. In game design, this kind of AI is frequently utilized to create opponents. The opponent is not aware of the main goal of the game, but they will react to your moves, attacks, and activities in real time. Furthermore, because it lacks memory, it is unable to draw lessons from the past or modify its gameplay. Numerous marketing solutions are powered by reactive AI. Chatbots are one such instance. Reactive AI is used by these algorithms to provide the appropriate information in response to messages or inputs. (Erin Rodrigue,Flori Needle, 2023)

Limited Memory

In order to process information and weigh possible actions, limited memory AI can store historical facts and forecasts. This is effectively looking into the past for hints about what might happen next. Reactive machines are less complex than limited memory AI, which offers more opportunities. Limited memory AI is developed when a group of people consistently trains a model to evaluate and make use of fresh data, or when an environment is developed specifically for AI to enable automatic model renewal and training. Self-Driving cars can be said to be the example of Limited Memory AI. (Schroer, 2023)

Theory of Mind

This type of AI as the name suggests itself, it only exits as a concept. It stands for a specialized category of technology capable of interpreting human brain states. For example, Google Maps does not become angry or provide emotional support if you yell at it because you missed a turn. As a response, it chooses another path.

The goal of theory of mind is to build machines that, by understanding human wants, objectives, and motivations, will be better able to interact with humans. An AI system, for instance, can react more delicately if it realizes the annoyances of a dissatisfied consumer. Theory of mind AI has the potential to significantly impact marketing in the long run. However, it's still in the early phases, so it's hard to say when it will happen. (Erin Rodrigue,Flori Needle, 2023)

Self-Awareness

The last stage of AI development will be self-awareness, which will happen long after theory of mind has been created. This type of artificial intelligence is conscious on the same level with humans, recognizing both its own presence and the presence and emotional states of others. It would have the ability to comprehend what other people could require based on both what they say to them and how they say it.

In order to include self-awareness into AI, human researchers must first grasp the fundamentals of consciousness before figuring out how to replicate it in machines. (Schroer, 2023)

1.1 Explanation of topic and concepts

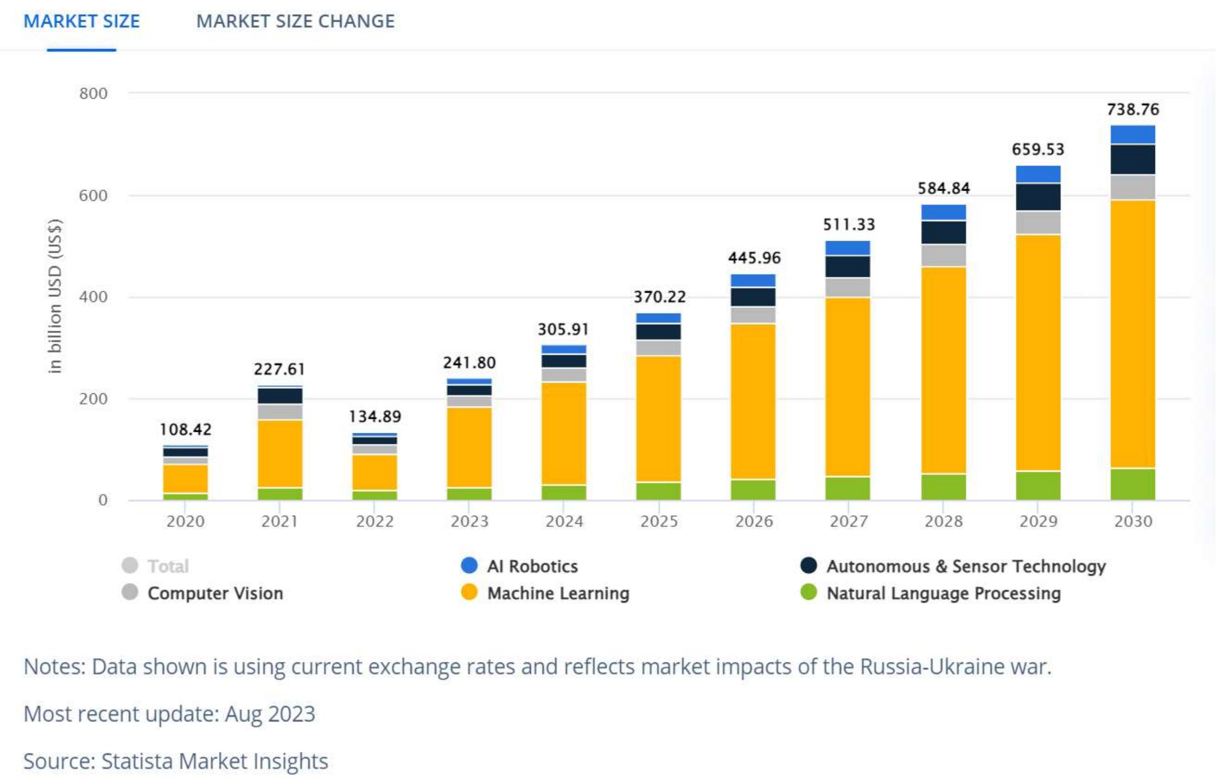
1.1.1 Artificial Intelligence and Machine Learning

Artificial intelligence, essentially, is the ability of a machine to replicate intelligent human behaviour and machine learning is the subset of the artificial intelligence, meaning machine learning is AI. Systems using artificial intelligence are utilized to carry out difficult jobs in a manner comparable to how people solve issues. AI aims to build computer models which demonstrate "intelligent behaviours" identical to those of humans, as stated by CSAIL main research scientist and leader of the InfoLab Group Boris Katz. This includes devices that can identify an image, read a document written in natural language, or carry out a task in the real world. (Brown, 2021)

Artificial intelligence is a concept that has been widely used in computer science and related fields. Recent advancements in machine learning and artificial intelligence have led to a significant increase in the popularity of this term. Machine learning is the branch of artificial intelligence where machines are thought to be smarter than people and are in charge of automating everyday tasks. Robotics and the Internet of Things have elevated machines to a new degree of intelligence and thought, to the point that they can now outsmart people. They are known to learn, adapt, and function far more quickly than humans are expected or programmed to. (Pedamkar, 2023)

The AI industry has grown significantly these past few years. As per below image from, the Statista Market Insights, the AI market reached about 108 billion USD dollars in

2020 and increased by double in 2021 which was about 227.61 billion USD dollars. However, due to the Russia-Ukraine war the market of AI declined to 134.64 billion dollars. However, it is projected that by the end of 2023 the market of AI will reach whopping 241.90 billion USD Dollars. And with this rate it will reach about 738 billion USD dollars by the end of 2030. This shows us how AI is growing rapidly and how important it is going to be in our lives. (Insights, 2023)



Notes: Data shown is using current exchange rates and reflects market impacts of the Russia-Ukraine war.

Most recent update: Aug 2023

Source: Statista Market Insights

Figure 3 - Market Size of AI (Insights, 2023)

Artificial Intelligence has made a big impact in our lives for past few years now. Machine learning is being used by analysts in the financial services industry to identify fraud, automate trading processes, and offer clients financial advice. There are several ways that machine learning can help businesses increase their efficacy, productivity, and services. For example, chatbots enable companies to offer more responsive, expedient customer care without requiring them to use call centers or put clients on hold until a representative becomes available. (Engineering, n.d.)In numerous airplanes, artificial intelligence has been used to do tasks like as navigation maps, taxing routes, and a fast inspection of the complete cockpit panel to guarantee that every component is operating correctly. This has been supported a lot because it has been producing really encouraging results. (Pedamkar, 2023)Machine learning can be used in the healthcare industry to assist hospitals and personnel detect and identify infectious diseases more effectively, customize medical treatments, and accelerate administrative procedures. (Engineering, n.d.)

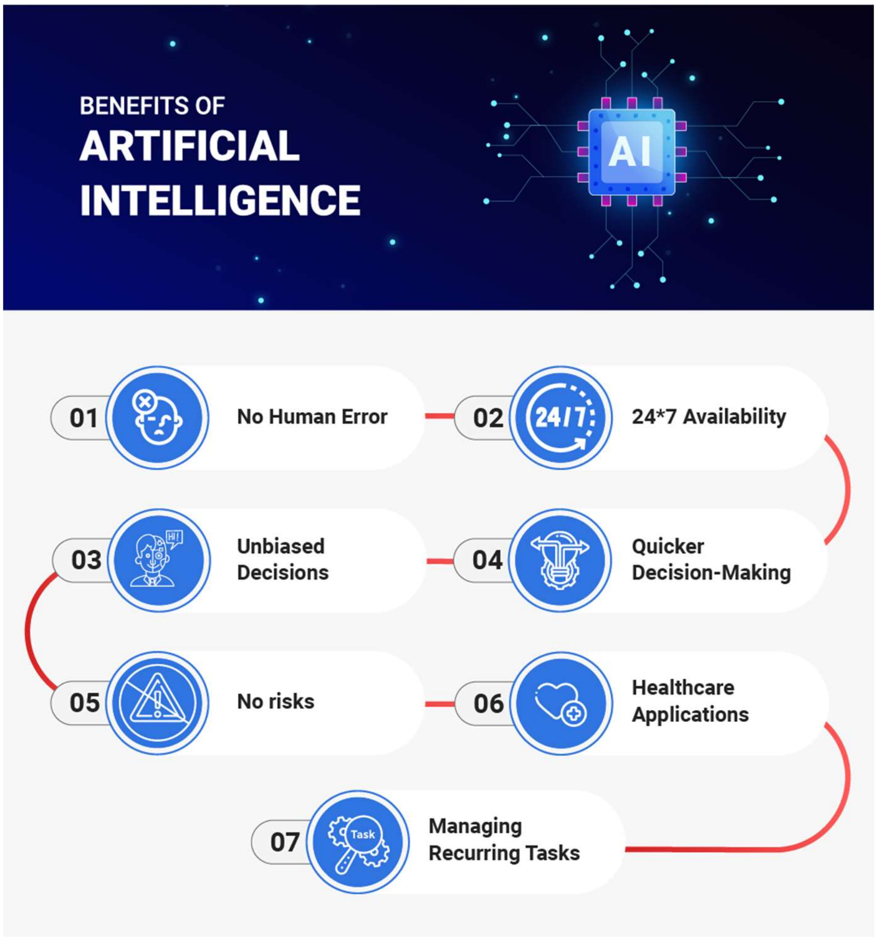


Figure 4 - Importance of AI (Oza, 2021)

1.2 Explanation of chosen topic/domain

For any unwanted message, it is known as spam. Although spam is primarily junk email, it can also be texts, calls, or messages on social media or even can be phone messages. Spam's the end goal is to convince an individual to open a message and purchase a good or service—which may or may not be fraudulent. Alternatively, some spammers try to trick customers into downloading a virus onto their systems in order to either steal important data or demand money from the victim. Spammers use bulk emailing to carry out phishing frauds. (Pickle, 2022)

Figure 5 - Spam (Gatefy, 2021)

The process of identifying and preventing unsolicited, undesired, and potentially virus-filled emails from getting into email inboxes is known as spam filtering. Spam filtering technology is used by Internet service providers (ISPs) and small- to medium-sized businesses (SMBs) to protect their networks, consumers, and staff against potential dangers. (MailChannels, n.d.)

Importance of Spam filtering

As per the data found in Statista published by Ani Petrosyan and shown below in the diagram, it was found that 48.63% of email sent was a spam mail in 2022. But when compared to the 80.26% spam email of 2011, it is a way better number. However, the email spam percentages are almost half, meaning you are getting only 50% email that are not spam. Therefore, it is still a very large number of spam mails. And this doesn't even count towards other spam sources. (Petrosyan, 2023)

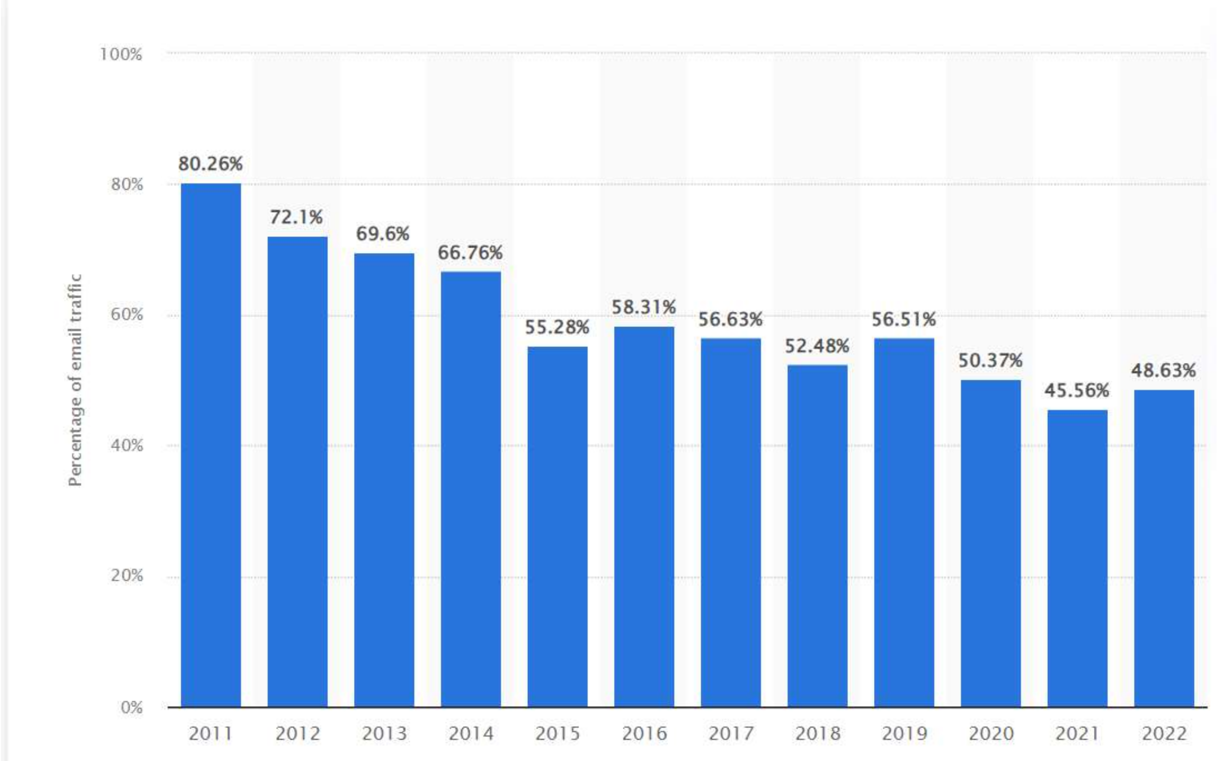


Figure 6 - Email Spam percentage (Petrosyan, 2023)

Since, there is this large percentage of spam mails in our lives every year, without spam filtering, we would potentially be scammed or tricked into buying something that we don't want to. For these, reasons spam filtering is required in our day to day lives.

2. Background

2.1 Research work done on the chosen topic/problem domain.

2.1.1. Spam Filtering

A great protection against unwanted messages (or, in the case of robocalls, phone calls) is known as spam filtering. Spam filtering helps detect and stop unwanted, unsolicited, and virus-filled emails (often referred to as spam) from entering email accounts. (Slavin, 2022)

Figure 7 - Spam Filtering (Slavin, 2022)

Traditional spam filtering without using machine learning.

To determine whether an incoming email is spam or not, non-machine learning techniques are utilized. These techniques primarily consist of a list of email addresses and a list of terms that are used to determine whether a given email is spam or not. List- and content-based spam filters are examples of non-machine techniques. While content-based filters identify based on the email's content, list-based filters do so based on a specified list. (Sandhi Kranthi Reddy, T Maruthi Padmaja, 2019)

Limitations of spam filtering without using machine learning

This way of spam filtering has significant drawbacks despite their overall effectiveness. These systems are less able to deal with newer and complex types of spam because they are unable to adjust to changing spamming patterns and methods. In addition, as spam increased in quantity and variety, it became more difficult to manage and update the regulations. (Md Rafiqul Islam, Morshed U. Chowdhury, 2015)

2.1.2. Machine Learning in Spam Filtering

The field of spam filtering saw a change in direction with the introduction of machine learning. Machine learning algorithms could be trained to identify patterns and features in data rather than depending on pre-established rules, which would allow them to adapt dynamically to evolving spam strategies. The ability to adapt greatly increased spam detection's effectiveness and accuracy. We can see in the figure below that the spam has decreases significantly after the year 2016 as this is the time when spam filtering software using machine learning has been started.

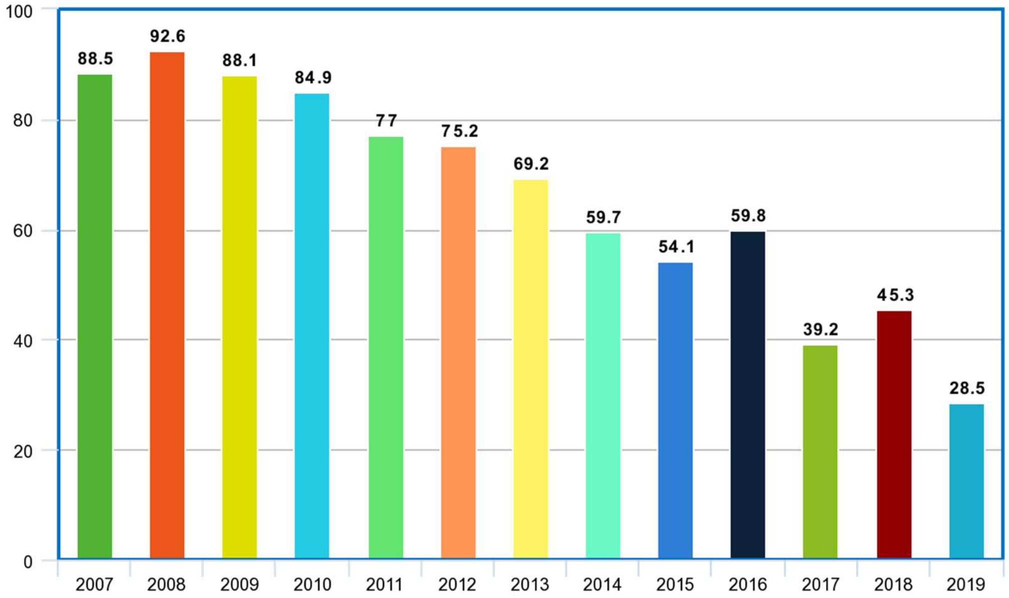


Figure 8 - Machine Learning spam protection (Raj, n.d.)

2.1.3. Advantages of Problem Domain

2.1.3.1. Accuracy

Spam filters can obtain extra information about the sender and their intentions by considering the website name. This aids in separating reputable emails from spam emails that have been cleverly disguised and may have faked well-known sender identities.

2.1.3.2. Reduced clutter

Users can keep their inbox free of unnecessary messages and make it simpler to locate critical emails by filtering spam emails.

2.1.3.3. Security

Spam emails are frequently used by malicious individuals to spread malware and phishing scams. Blocking spam might help in preventing users from becoming victims of these frauds.

2.1.3.4. Reduced Server Load

Server resources might be severely strained by spam emails. Reducing server load and enhancing overall system performance can be achieved by blocking spam.

2.1.3.5. Protects against scams.

Spam filters are essential for protecting people against phishing attempts and other internet scams that try to obtain personal information.

2.1.4. Disadvantages of Problem Domain

2.1.4.1. False Positives and Negatives

Sometimes spam filters can wrongly classify emails that are authentic as spam. Also, sometimes, Spam emails may manage to get past your security measures when spam filters are unable to recognize them.

2.1.4.2. Privacy

Privacy concerns may arise from the fact that spam filters frequently gather and examine your personal data.

2.1.4.3. Technical Challenges

It's challenging to stay on top of the most recent threats because spammers are always coming up with new ways to get around spam filters.

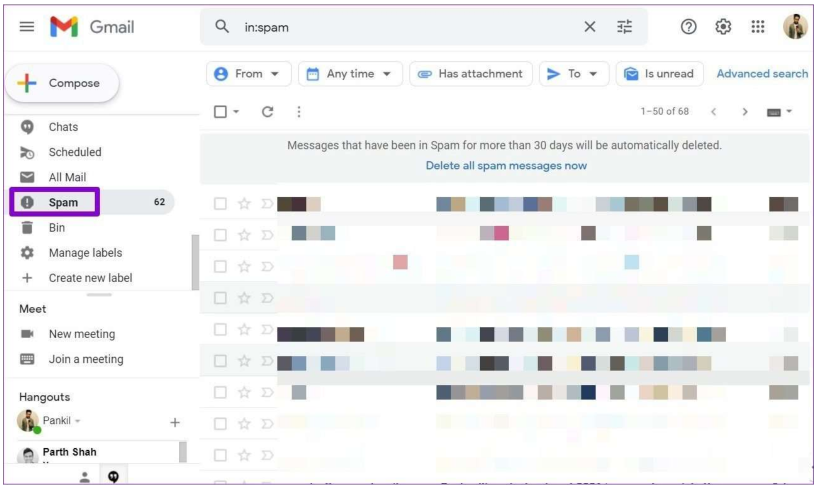
2.1.4.4. Over blocking

Over blocking may occur when spam filters are overly strict, preventing both spam and legitimate emails from getting through, which may affect communication.

2.1.5. Examples of spam filtering used by email service providers.

2.1.5.1. Gmail

Gmail is a popular email service provider by Google. It is one of the email service providers which uses machine learning to detect and block spam.



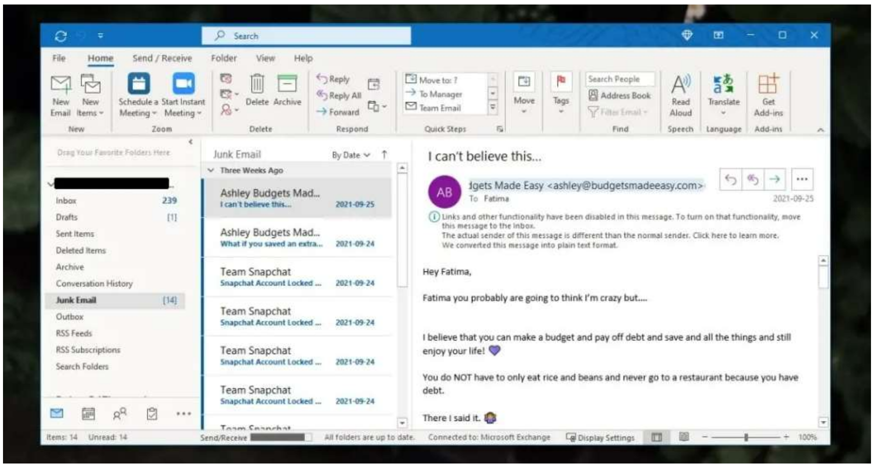
Messages that have been in Spam for more than 30 days will be automatically deleted.

Delete all spam messages now

Figure 9 - Gmail Spam (Shah, 2022)

2.1.5.2. Outlook

Outlook is another popular email service provider by Microsoft which also uses machine learning algorithms to detect spams, or they call it junk.



Ashley Budgets Mad...	I can't believe this...	2021-09-25
Ashley Budgets Mad...	What if you saved an extra...	2021-09-24
Team Snapchat	Snapchat Account Locked ...	2021-09-24
Team Snapchat	Snapchat Account Locked ...	2021-09-24
Team Snapchat	Snapchat Account Locked ...	2021-09-24
Team Snapchat	Snapchat Account Locked ...	2021-09-24

I can't believe this...

AB ɟgets Made Easy <ashley@budgetsmadeeasy.com>
To Fatima
2021-09-25

Links and other functionality have been disabled in this message. To turn on that functionality, move this message to the Inbox.
The actual sender of this message is different than the normal sender. Click here to learn more.
We converted this message into plain text format.

Hey Fatima,

Fatima you probably are going to think I'm crazy but....

I believe that you can make a budget and pay off debt and save and all the things and still enjoy your life! 💜

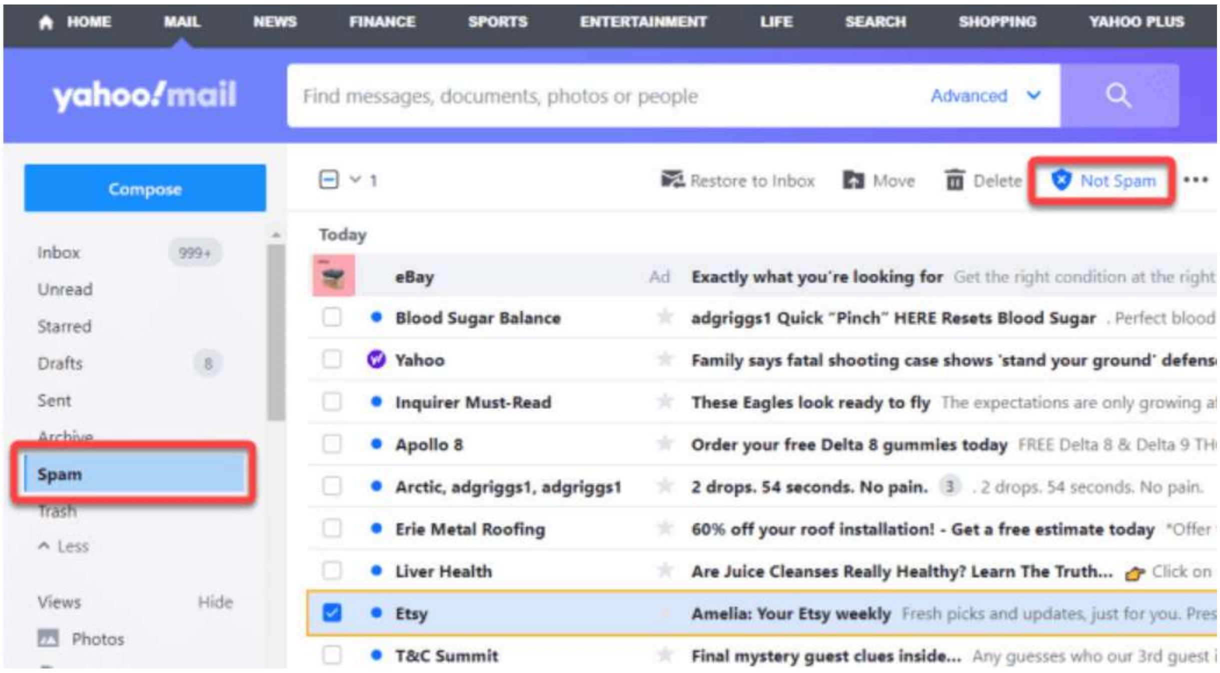
You do NOT have to only eat rice and beans and never go to a restaurant because you have debt.

There I said it.

Figure 10 - Outlook Spam (Wahab, 2021)

2.1.5.3. Yahoo Mail

This is another example of email service providers which also uses machine learning to detect and block spams.



eBay	Ad	**Exactly what you're looking for** Get the right condition at the right
Blood Sugar Balance		**adgriggs1 Quick "Pinch" HERE Resets Blood Sugar** . Perfect blood
Yahoo		**Family says fatal shooting case shows 'stand your ground' defens**
Inquirer Must-Read		**These Eagles look ready to fly** The expectations are only growing a
Apollo 8		**Order your free Delta 8 gummies today** FREE Delta 8 & Delta 9 TH
Arctic, adgriggs1, adgriggs1		**2 drops. 54 seconds. No pain.** 3 . 2 drops. 54 seconds. No pain.
Erie Metal Roofing		**60% off your roof installation! - Get a free estimate today** "Offer
Liver Health		**Are Juice Cleanses Really Healthy? Learn The Truth...** Click on
☑ Etsy		**Amelia: Your Etsy weekly** Fresh picks and updates, just for you. Pres
T&C Summit		**Final mystery guest clues inside...** Any guesses who our 3rd guest i

Figure 11 - Yahoo Spam (GRIGGS, 2022)

Despite all the email service provider use machine learning to spam filtering, still they haven't perfected the filtering, as we see that we must sometime, go to spam to find the mails that are not spam, or some mails that are spam can be found in the main email section. There, there is need of constant research to get more effective system that will help detect spam or ham mails.

2.1.6. Dataset

The dataset we are going to use to train the model is taken from Kaggle. This dataset contains three columns which is numbering(unnamed in the dataset), body, and labels. The dataset is not clean dataset which means before doing anything with it, we need to clean or prepare it. There are empty rows and numerous things to be removed before extracting the bag of words from it in order to train models. The dataset is extracted from: https://www.kaggle.com/datasets/nitishabharathi/email-spam-dataset

The main two columns that we will be using are:

- **Body:** All the email contents are found in this column which needs to be converted into bag of words.
- **Label:** The label has two numbers 0 and 1 in which 0 means not a spam mail and 1 means a spam mail.

Name	**Description**	**Datatype**
Body	All the contents of emails that are spam and are not spam.	Object
Label	Identification of body whether it is spam or not in 0 and 1.	Integer

Table 1 - Dataset Description

2.2 Review and analysis of work in the problem domain

2.2.1. Research Paper 1

Title: SMS Spam Detection using Machine Learning and Deep Learning Techniques

Authors: Sridevi Gadde, A. Lakshmanarao, S. Satyanarayana

Website Extracted from: https://ieeexplore.ieee.org/document/9441783

Citation: (Sridevi Gadde, A.Lakshmanarao, S.Satyanarayana, 2021)

Summary

This research paper basically explains how in this modern age where the number of phones is increasing day by day in fact it has reached 3.8 billion users of mobile users in just five years from 1 billion users. As per the research, the aim is to see how different machine learning can be used to detect spam and how they compare to each other.

In their research paper, they did their own research where they found out they that deep neural network and Hidden Markov Models (HMM) is the most accurate in finding spams with accuracy of 98%. However, as their aim is to find the effectiveness of ML algorithms themselves, they took a dataset from UCI repository and cleaned it using NLP and trained models for different algorithms and compared them to each other. They found out that Logistic Regression and Decision Tree is the most accurate algorithm to detect spam with the accuracy of 94%. And the SVM was second in their list with 93%. But they saw that the dataset was unbalanced. After balancing the dataset using a sampling technique called SMOTE (Synthetic Minority Oversampling), they achieved that Logistic regression is still at the top with accuracy of 95% whereas the SVM came second with 94%.

2.2.2. Research Paper 2

Title: Detection of Email Spam using Machine Learning Algorithms: A Comparative Study

Authors: Prazwal Thakur, Kartik Joshi, Prateek Thakral, Shruti Jain

Website Extracted from: https://ieeexplore.ieee.org/document/10009149

Citation: (Prazwal Thakur, Kartik Joshi, Prateek Thakral, Shruti Jain, 2022)

Summary

Here, in this paper, they are talking about how internet is a very integral part of your lives in this era and how there are so much internet spam. However, they want to focus on the email spam part of the internet spam. This research paper wants to detect spam using machine learning techniques rather than the knowledge engineering method of spam detection and filtration. This document is more through on the process of training a model with the flowchart and description of the algorithms as well as how algorithms compared to each other. At last, they chose, SVC (Support Vector Classifier), KNN (K-Nearest Neighbours), Logistic Regression, Naïve Bayes, and Decision Tree. In the test, they came to conclusion that SVC or SVM came at the very top with the accuracy of 98.09%. However, we wouldn't know which one of these algorithms came in second as they are describing KNN to be the second with accuracy of 95.3% however, in the table it is shown that Logistic Regression is second with the accuracy of 95.59%. However, we know that the SVM is the most effective in finding the spam in this research paper.

2.2.3. Research Paper 3

Title: A Comprehensive Review on Email Spam Classification using Machine Learning Algorithms

Authors: Mansoor RAZA, Nathali Dilshani Jayasinghe, Muhana Magboul Ali Muslam

Website Extracted from: https://ieeexplore.ieee.org/document/9334020

Citation: (Mansoor RAZA, Nathali Dilshani Jayasinghe, Muhana Magboul Ali Muslam, 2021)

Summary

The research paper is basically about how different algorithms or sets of algorithms can be used to detect spam and how they compare to each other. Here, it is described that nearly 50 percent of emails or unwanted emails among the 40 -50 emails a person gets in a day. They have also highlighted that almost, 3.5 USD million dollars are earned by spammers every year. As per the FBI, in the year 2019, the businesses financial loss due to phishing and spam is over 12.5 USD billion dollars. Here, in the document, they want to see how, and which machine learning are popular to detect spams. Here, unlike the other research papers, they have also done research on unsupervised machine learning approach for spam detection.

In their test, they have checked multiple groups of algorithms for their accuracy. Only two out of twelve are the algorithms that are not in groups to find the accuracy. However, at last they came to conclusion that the two algorithms are highly favoured algorithms for spam detection. They also found out that, multi algorithms are used for better outcome that using a single algorithm.

2.2.4. Research Paper 4

Title: A Study of Machine Learning Classifiers for Spam Detection

Authors: Shrawan Kumar Trivedi

Website Extracted from: https://ieeexplore.ieee.org/document/7743279

Citation: (Trivedi, 2016)

Summary

Here, in this paper different types of machine learning classified have been described thoroughly with tall the formulas and how it works. The classifiers are SVM, Naïve Bayes, Bayesian, J48(Decision Tree), along with the same classifier but using it with adaboost to re-assess the accuracy. The comparison showed that the accuracy of SVM is the highest of them all coming at 93.3%, after that Naive Bayes with boosting came at second with 93.2% of accuracy. It seems that SVM is the best classifier to be used and the research paper also concludes that SVM is the algorithm to be used for spam detection. And the lowest accuracy was of Bayesian with the score of 92.0%. Therefore, the Bayesian should not be used comparatively to others as the research paper shows.

2.2.5. Research Paper 5

Title: Evaluating the Effectiveness of Machine Learning Methods for Spam Detection

Authors: Kingshuk Debnath, Nirmalya Kar

Website Extracted from: https://ieeexplore.ieee.org/document/9850588

Citation: (Kingshuk Debnath, Nirmalya Kar, 2022)

Summary

This research like other research is also about seeing what technique the best is in finding spam, however, here they also focused on deep learning along with machine learning unlike other research papers where the machine learning is only implemented. While researching they found out that, there are 45.1% spam email traffic in March of 2021. They have described that their goal is to make a model using machine learning as well as deep learning to have more precision than the previous models. They found from other research that SVM and Naïve Bayes are the algorithms with a high accuracy. However, to meet their goal they have to tests of their own. Therefore, they took a dataset from Enron, and pre-processed the data and used it to create models of different machine learning algorithms as well deep learning algorithms.

In their test, they found out that the MNB (Multinominal Naïve Bayes) algorithm has the most accuracy with 98.13% while SVM was second on the list with 98.06%, which when they compared to another research that they did was higher. Also, among the deep-learning algorithms, the BERT (Bidirectional Encoder Representations from Transformers) deep-learning technique was the most effective with 99.14% accuracy. As per the result, they did succeed in their goal. However, they didn’t compare the precision score which was the main aim.

2.2.6. Summarized Review and analysis

Together, the five studies research machine learning methods for spam recognition. Naïve Bayes, SVM, KNN, Neural Networks, Decision Trees, Random Forests, TF-IDF, BART, Logistic Regression and many others are the algorithms that have been examined among this five research. Every research analyzes the effectiveness of various algorithms and highlights the importance of filtering to reduce spam.

Research 1 shows Logistic Regression is the most accurate in their own tests (95%) and found that SVM was the second most accurate. They determined that SVM and Logistic Regression performed better than other methods.

In Research 2, The best algorithms are found to be SVM in the top and Logistic Regression and KNN are in the second.

The third study compares different multi-algorithms techniques as well as single algorithms techniques for email spam detection. They concluded that SVM and Naïve Bayes are the most popular and wanted algorithms even in the multi algorithms techniques.

Research 4 provides analysis of various classifiers, giving SVM the highest ranking (93.3%) and indicating that it is the recommended algorithm for spam detection.

Research 5 compares both machine learning and deep learning algorithms. The most accurate machine learning algorithms is Multinomial Naïve Bayes with 93.13%, closely followed by SVM (98.06%). In the deep-learning technique, the BERT came in top with 99.14%.

To sum up, studies frequently show that Naïve Bayes and SVM are efficient algorithms for spam identification, beating other approaches and showing high accuracy and precision.

3. Solution

3.1 Proposed Approach to solving the problem.

The proposed approach in solving this problem of spam is to use machine learning algorithms to find the spam and filter them. As per the research we have done above, we can see that Support Vector Machine (SVM) and Naive Bayes are the two of the algorithms that were on top in almost all the research. We will also be these two machine learning algorithms along with k-Nearest Neighbours (KNN) to detect the spams. We will be doing our own test checking which one of the three algorithms are the most accurate in order to find the right algorithm as a solution for spam detection.

3.1.1. Elaboration of the proposed Algorithms

3.1.1.1. Support Vector Machine (SVM)

The goal of SVM supervised machine learning algorithm is to identify the hyperplane that divides the two classes the best. Regression and SVM may seem similar, but they are not the same. whereas both techniques aim to identify the optimal hyperplane, the primary distinction between them is that support vector machines rely on statistical methods, whereas logistic regression takes a probabilistic approach.

Since the margin in SVM is calculated using the points that are closest to the hyperplane (support vectors), we don't need to worry about additional observations; in logistic regression, on the other hand, the classifier is defined over all of the points. As a result, SVM naturally speeds faster. Let's use an example to better understand how SVM functions. Assume we have two classes in our dataset (green and blue). The new data point needs to be categorized as either blue or green. The primary goal of SVM is to find the best hyperplane, which is the plane with the maximum distance from both classes. This is accomplished by identifying many hyperplanes that best classify the labels, after which it selects the hyperplane that is the furthest from the data points or that has the largest margin. We can see a hyper plane that best classify the labels in the below diagram. (Saini, 2023)

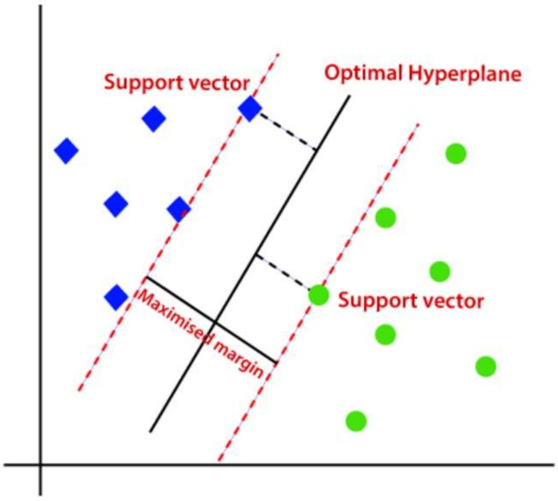


Figure 12 - SVM example (Saini, 2023)

3.1.1.2. Naïve Bayes

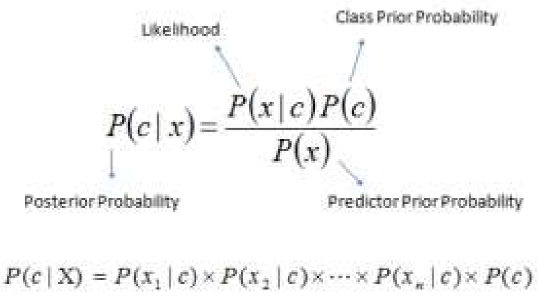
This is a popular supervised machine learning approach for classification applications like text classification. It is a member of the generative learning algorithm family, which implies that it simulates the input distribution for a certain class or category. This method relies on the assumption that the input data's attributes are conditionally independent given the class, which enables the algorithm to produce precise and fast predictions.

Naive Bayes classifiers are regarded as straightforward probabilistic classifiers in statistics that make use of the Bayes theorem. The probability of a hypothesis, given the facts and some prior information, is the basis of this theorem. The naive Bayes classifier assumes that every characteristic in the input data is unrelated to every other feature, which is frequently false in practical applications. The naive Bayes classifier is nevertheless frequently used because to its effectiveness and strong performance in numerous real-world applications. (Ray, 2023)

Using P(c), P(x), and P(x|c), one can calculate the probability P(c|x) using the Bayes theorem. and the equation is as follows:

$$P(c|x) = \frac{P(x|c).P(c)}{P(x)}$$

Equation 1 - Bayes equation



$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 13 - Bayes equation (Ray, 2023)

where,

- P(c|x) represents the class (c, target) posterior probability given a predictor (x, characteristics).
- The prior probability of the class is P(c).
- The likelihood, or probability of the predictor given class, is expressed as P(x|c).
- The predictor's prior probability is denoted by P(x).

3.1.1.3. K-nearest Neighbours (KNN)

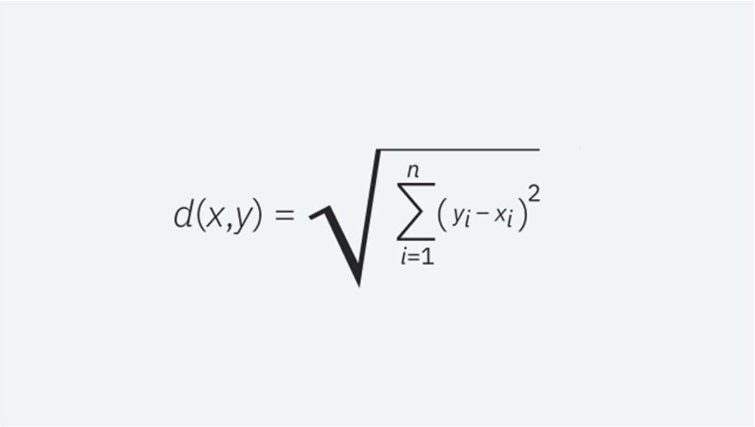
The k-nearest neighbours’ algorithm is a non-parametric supervised learning classifier that groups individual data points based on closeness in order to classify or predict data. Although it can be applied to classification or regression issues, it is usually employed as a classification algorithm, based on the idea that comparable points can be located next to each other. The k-nearest neighbour algorithm seeks to locate a query point's closest neighbours so that a class label can be applied to it.

The distance between a query point and the other data points must be computed in order to figure out which data points are closest to a particular query point. The decision boundaries that divide query points into various areas are formed in part by these distance measurements. While there are several distances measures, we will be seeing the equation of Euclidean distance.

Euclidean distance is the most widely used distance metric, although it can only be applied to vectors with real values. It measures a straight line between the query location and the other point being measured using the formula below. (IBM, n.d.)

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^{\,2}}$$

Equation 2 - Euclidean distance equation



$$d(x,y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

Figure 14 - Euclidean distance equation (IBM, n.d.)

where,

- d represents the distance
- x represents the first point
- y represents the second point
- i represent the index to iterate over coordinates.
- n is number of dimensions

3.1.2. Implementation methodology of used Algorithm in the system

3.1.2.1. Data Preparation

- Dataset cleaning and preparation by taking care of missing values and eliminating unnecessary data.
- Feature extraction to make it machine-learning-friendly.

3.1.2.2. Implementation of Algorithm

- Implementing SVM, Naive Bayes, and KNN algorithms to use with a machine learning library (such as Python's scikit-learn).

3.1.2.3. Model Evaluation

- dividing the dataset into sets for testing and training.
- Modelling the training datasets and assessing it using testing datasets

3.1.2.4. Comparison and Analysis

- Analysing the accuracy, precision, recall, and F1-score performance of SVM, Naive Bayes, and KNN.

3.2 Pseudocode

```
START
IMPORT libraries
IMPORT datasets
ANALYZE the dataset
REMOVE unwanted columns
IF missing values
    REMOVE missing values
END IF
REMOVE http links
REMOVE unnecessary other characters expect numerical and alphabetical characters
LOWER all the words
SPLIT the words
CREATE a list of the split words
CREATE bag of words
SPLIT the dataset into train and test sets
TRAIN models using train datasets using SVM (Support Vector Machine), Naïve Bayes and KNN (K-Nearest Neighbours)
TEST the model using test datasets
CHECK accuracy, precision, recall, and F1-score
STOP
```

3.3 Flowchart

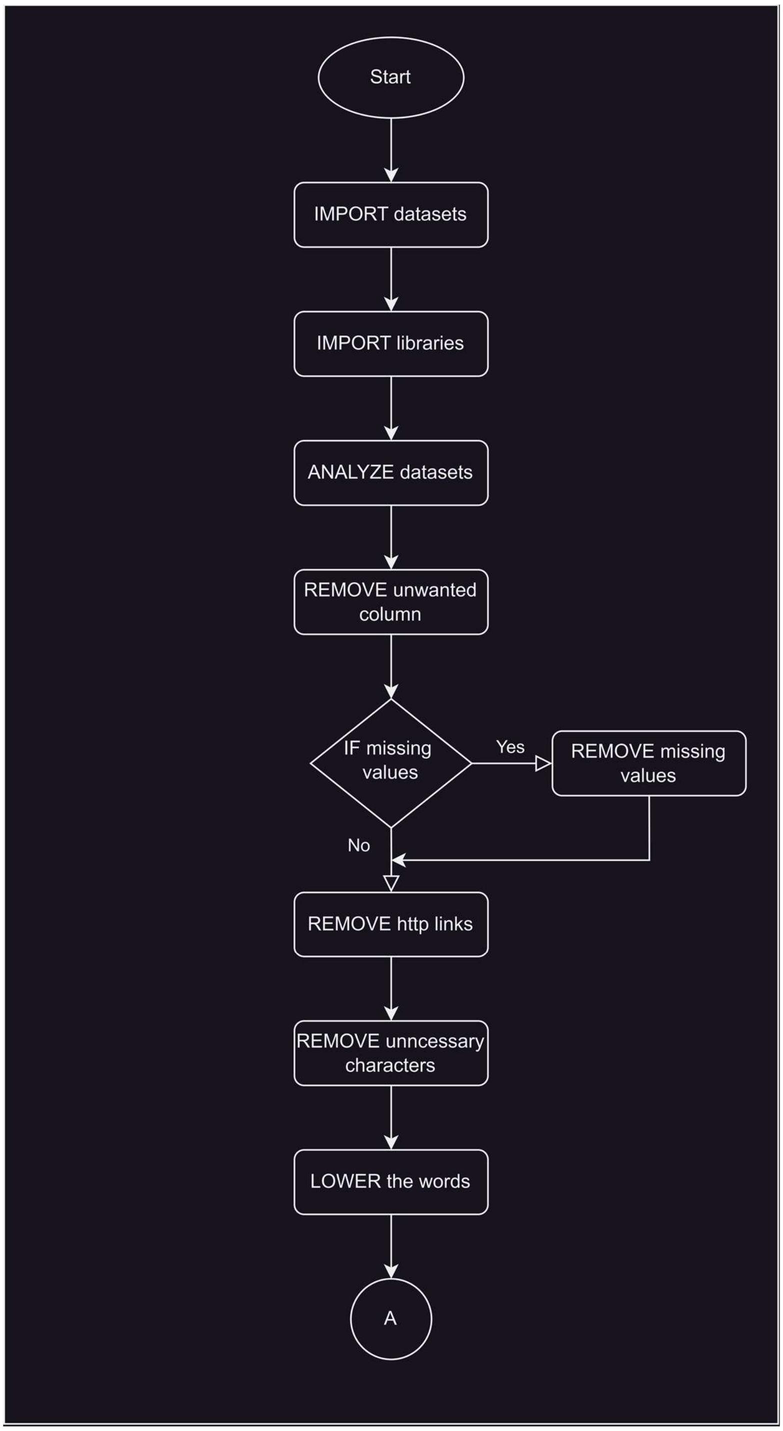


Figure 15 - Flowchart Part 1

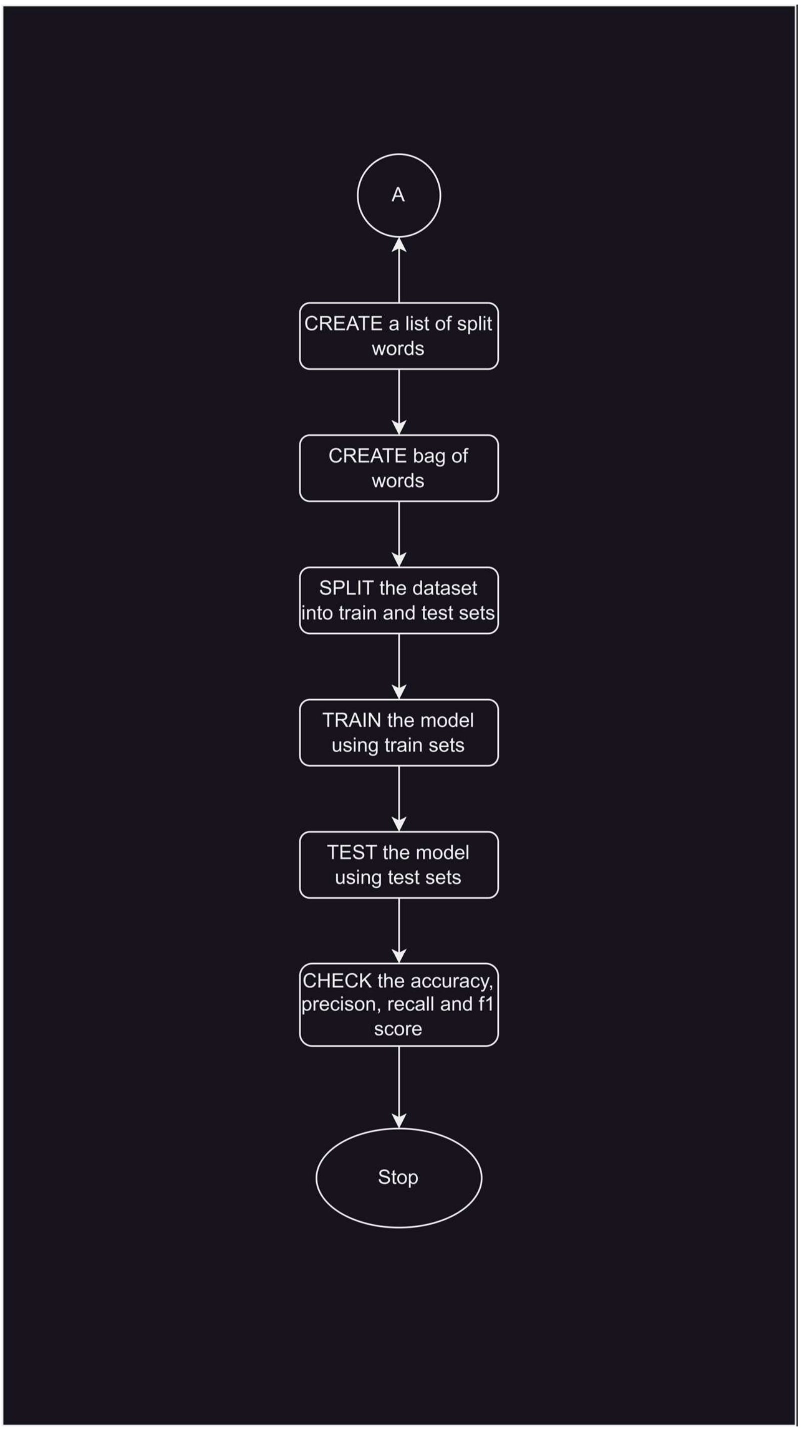

A
CREATE a list of split words
CREATE bag of words
SPLIT the dataset into train and test sets
TRAIN the model using train sets
TEST the model using test sets
CHECK the accuracy, precison, recall and f1 score
Stop

Figure 16 - Flowchart Part 2

4. Conclusion

4.1 Analysis of the work done

Research on spam filtering using AI and machine learning highlights the present and future possibilities of this technology. Analysing various AI types—from simple reactive robots to sophisticated self-awareness—offers a framework for comprehending the evolution of AI. Applications of machine learning, especially deep learning via neural networks, are significant in a number of domains, including spam filters and self-driving automobiles. Studies on spam detection show how ML algorithms, such as Naive Bayes and SVM, are replacing traditional methods because they are more accurate and flexible. These algorithms consistently perform well, as demonstrated by an analysis of five research articles. This makes them promising options in the field spam filtering. In conclusion we will be doing our own test of effectiveness of different algorithms for spam detection. Therefore, here in this documentation we have laid the foundation to create, analyse and select the most effective algorithms among the popular algorithms for spam detection.

4.2 Solution that addresses real world problems

The machine learning techniques SVM, Naive Bayes, and KNN are chosen in order to tackle the common problem of spam. These algorithms were chosen because to their demonstrated effectiveness in previous research articles. Improving the accuracy of spam filtering and reducing the problems caused by spams are the main goals of the solution.

Application in Real-world

Cybersecurity

Cybersecurity measures are strengthened by the application of machine learning algorithms, which offer preventive protection against any risks hidden in spam emails.

Protection against phishing attacks

The application of spam filtering using machine learning will ensure a safer online experience by detecting and blocking phishing emails, which will stop unwanted access to any information that is not supposed to be in the hands of unwanted person.

Business and Individual Benefits

When workers spend less time sorting through spam and more time on important duties, businesses benefit from increased operational efficiency. Furthermore, by protecting companies from email-based dangers and avoiding possible breaches of security and phishing schemes, an efficient spam detection system maintains brand reputation. Also, people feel more confident when interacting with others online because they are aware that their lines of communication are protected from spam, which enhances the online experience.

4.3 Further work

In the future, we will be implementing all the things that we learned throughout this paper and will use it to make a model. The further works that we will be doing are as follows:

Dataset Preparation

Thorough dataset preparation is essential before putting the spam detection model into practice. This is a multi-step procedure that begins with data cleaning by removing missing values. Techniques for feature extraction, in our case, bag-of-words, convert unprocessed email data into a format that is appropriate for machine learning algorithms. After that, the dataset is divided into training and testing sets, and any necessary scaling or normalization is implemented.

Model Implementation

After the dataset has been prepared and divided into training and testing sets, we will be using the chosen machine learning algorithms from this paper which are SVM, Naïve Bayes, and KNN to train the model using training sets for spam detection.

Evaluation

When the model has been trained for the spam detection, it should be evaluated using the test datasets. After that we will be using this test dataset to evaluate the accuracy, precision, recall and F1-Score.

Algorithm Comparison

After the accuracy, precision, recall and F1-Score have been found of all the algorithm at the very last we will be using plots to see which among the three algorithm is more effective than the other two.

5. References

Brown, S., 2021. *Machine learning, explained.* [Online]
Available at: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained
[Accessed 9 December 2023].

Copeland, B., 2023. *artificial intelligence.* [Online]
Available at: https://www.britannica.com/technology/artificial-intelligence
[Accessed 6 December 2023].

Engineering, C., n.d. *Top Machine Learning Applications by Industry: 6 Machine Learning Examples.* [Online]
Available at: https://bootcamp.cvn.columbia.edu/blog/machine-learning-applications/
[Accessed 9 December 2023].

Erin Rodrigue,Flori Needle, 2023. *4 Types of Artificial Intelligence & What Marketers Are Using Most (Research).* [Online]
Available at: https://blog.hubspot.com/marketing/types-of-ai
[Accessed 6 December 2023].

Frankenfield, J., 2022. *Weak AI (Artificial Intelligence): Examples and Limitations.* [Online]
Available at: https://www.investopedia.com/terms/w/weak-ai.asp
[Accessed 6 December 2023].

Gatefy, 2021. *7 most common types of email spam.* [Online]
Available at: https://gatefy.com/blog/most-common-types-email-spam/
[Accessed 10 December 2023].

GRIGGS, A., 2022. *How to Set up Spam Settings in Yahoo Email.* [Online]
Available at: https://turbofuture.com/internet/How-to-Setup-Spam-Settings-in-Yahoo-Email-Including-Marking-and-Unmarking-Email-Messages-Plus-Understanding-Spam
[Accessed 18 December 2023].

IBM, n.d. *What is the k-nearest neighbors algorithm?.* [Online]
Available at: https://www.ibm.com/topics/knn
[Accessed 16 December 2023].

Insights, S. M., 2023. *Artificial Intelligence - Worldwide.* [Online]
Available at: https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide#market-size
[Accessed 9 December 2023].

Kanade, V., 2022. *What Is Machine Learning? Definition, Types, Applications, and Trends for 2022.* [Online]
Available at: https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/
[Accessed 6 December 2023].

Kingshuk Debnath, Nirmalya Kar, 2022. Email Spam Detection using Deep Learning Approach. In: *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON).* Faridabad: IEEE, pp. 37-41.

MailChannels, n.d. *What is Spam Filtering.* [Online]
Available at: https://blog.mailchannels.com/what-is-spam-filtering/
[Accessed 10 December 2023].

Mansoor RAZA, Nathali Dilshani Jayasinghe, Muhana Magboul Ali Muslam, 2021. A Comprehensive Review on Email Spam Classification using Machine Learning Algorithms. In: *2021 International Conference on Information Networking (ICOIN).* Jeju Island: IEEE, pp. 327-332.

Md Rafiqul Islam, Morshed U. Chowdhury, 2015. Spam filtering using Ml algorithms. *Spam filtering using Ml algorithms*, 12 February, p. 425.

Menon, K., 2023. *Different Types of Machine Learning: Exploring AI's Core.* [Online]
Available at: https://www.simplilearn.com/tutorials/machine-learning-tutorial/types-of-machine-learning
[Accessed 18 December 2023].

Oza, H., 2021. *Importance And Benefits Of Artificial Intelligence.* [Online]
Available at: https://www.hdatasystems.com/blog/importance-and-benefits-of-artificial-intelligence
[Accessed 9 December 2023].

Pedamkar, P., 2023. *Importance of Artificial Intelligence.* [Online]
Available at: https://www.educba.com/importance-of-artificial-intelligence/
[Accessed 9 December 2023].

Petrosyan, A., 2023. *Global spam volume as percentage of total e-mail traffic from 2011 to 2022.* [Online]
Available at: https://www.statista.com/statistics/420400/spam-email-traffic-share-annual/
[Accessed 10 December 2022].

Pickle, B., 2022. *Spam.* [Online]
Available at: https://techterms.com/definition/spam
[Accessed 10 December 2023].

Prazwal Thakur, Kartik Joshi, Prateek Thakral, Shruti Jain, 2022. Detection of Email Spam using Machine Learning Algorithms: A Comparative Study. In: *2022 8th International Conference on Signal Processing and Communication (ICSC).* Noida: IEEE, pp. 349-352.

Raj, R., n.d. *Email Spam and Non-spam filtering using machine Learning.* [Online]
Available at: https://www.enjoyalgorithms.com/blog/email-spam-and-non-spam-filtering-using-machine-learning
[Accessed 12 December 2023].

Ray, S., 2023. *Naive Bayes Classifier Explained: Applications and Practice Problems of Naive Bayes Classifier.* [Online]
Available at: https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/
[Accessed 16 December 2023].

Saini, A., 2023. *Guide on Support Vector Machine (SVM) Algorithm.* [Online]
Available at: https://www.analyticsvidhya.com/blog/2021/10/support-vector-

machinessvm-a-complete-guide-for-beginners/#:~:text=optimize%20in%20SVM.-,Margin%20in%20Support%20Vector%20Machine,and%20b%20is%20an%20offset.&text=If%20the%20value%20of%20w,it%20is%20a%20negativ
[Accessed 16 December 2023].

Sandhi Kranthi Reddy, T Maruthi Padmaja, 2019. Non Machine and Machine Learning Spam. *International Journal of Recent Technology and Engineering,* 7(5S4), p. 5.

Schroer, A., 2023. *Artificial Intelligence.What Is Artificial Intelligence (AI)? How Does AI Work?.* [Online]
Available at: https://builtin.com/artificial-intelligence
[Accessed 6 December 2023].

Shah, P., 2022. *Top 3 Ways to Stop Emails from Known Senders Going in Gmail's Spam Folder.* [Online]
Available at: https://www.guidingtech.com/ways-to-stop-emails-from-known-senders-going-in-gmails-spam-folder/
[Accessed 18 December 2023].

Slavin, B., 2022. *What Is Spam Filtering And How Does It Work?.* [Online]
Available at: https://www.duocircle.com/email-hosting/what-is-spam-filtering-and-how-does-it-work
[Accessed 12 December 2023].

Sridevi Gadde, A.Lakshmanarao, S.Satyanarayana, 2021. SMS Spam Detection using Machine Learning. In: *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS).* Coimbatore: IEEE, pp. 358-362.

Trivedi, S. K., 2016. A study of machine learning classifiers for spam detection. In: *2016 4th International Symposium on Computational and Business Intelligence (ISCBI).* Olten: IEEE, pp. 176-180.

Wahab, F., 2021. *How to fix Outlook keeps sending emails to Junk or Spam Folder.* [Online]
Available at: https://www.addictivetips.com/windows-tips/outlook-sending-emails-to-

junk-folder/
[Accessed 18 December 2023].