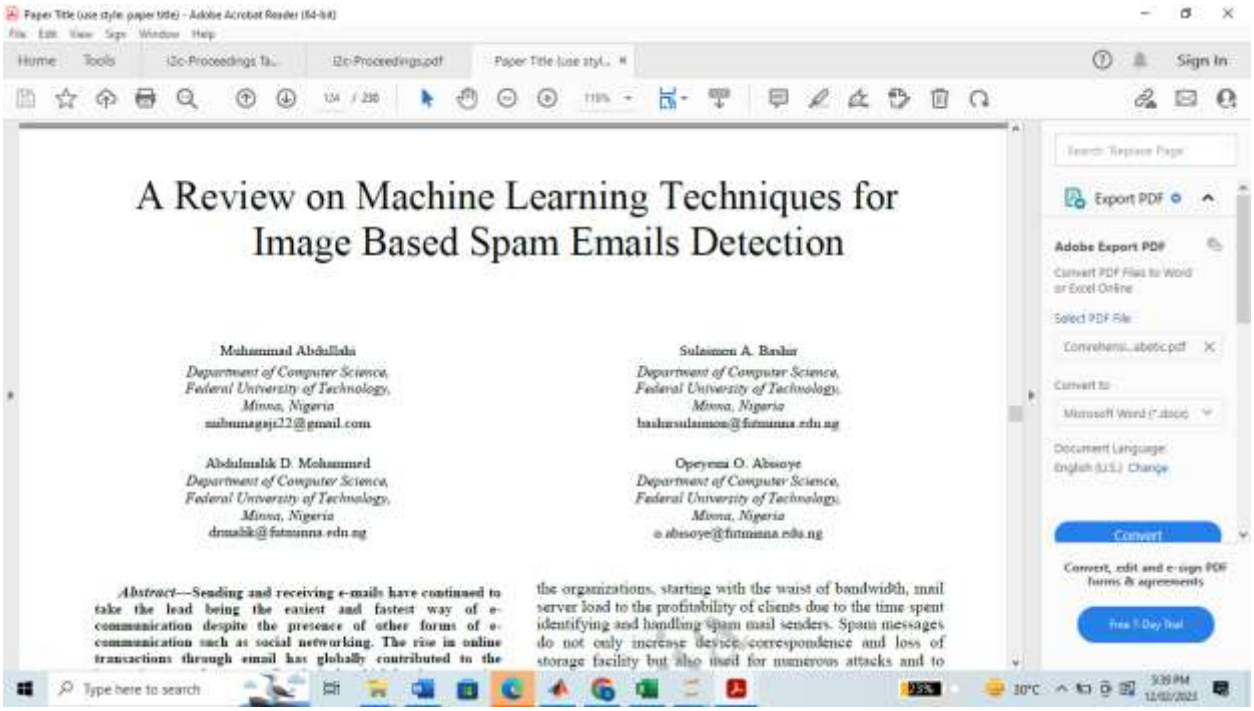


A Review on Machine Learning Techniques for Image Based Spam Emails Detection

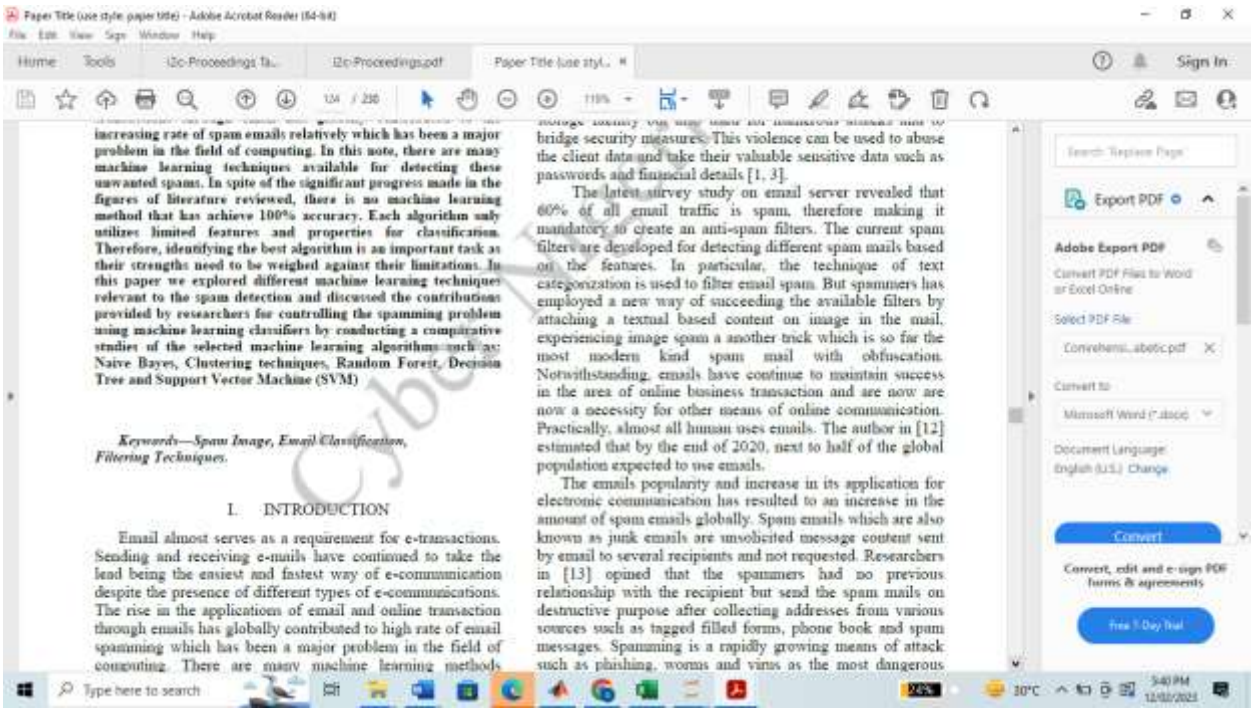
Muhammad Abdullahi
Department of Computer Science, Federal University of Technology, Minna, Nigeria
muhmmagaji22@gmail.com

Sulaimon A. Bashir
Department of Computer Science, Federal University of Technology, Minna, Nigeria
bashirsulaimon@futminna.edu.ng

Abdulmalik D. Mohammed
Department of Computer Science, Federal University of Technology, Minna, Nigeria
dmalik@futminna.edu.ng

Opeyemi O. Abisoye
Department of Computer Science, Federal University of Technology, Minna, Nigeria
o.abisoye@futminna.edu.ng

***Abstract*—Sending and receiving e-mails have continued to take the lead being the easiest and fastest way of e-communication despite the presence of other forms of e-communication such as social networking. The rise in online transactions through email has globally contributed to the** the organizations, starting with the waist of bandwidth, mail server load to the profitability of clients due to the time spent identifying and handling spam mail senders. Spam messages do not only increase device, correspondence and loss of storage facility but also used for numerous attacks and to



increasing rate of spam emails relatively which has been a major problem in the field of computing. In this note, there are many machine learning techniques available for detecting these unwanted spams. In spite of the significant progress made in the figures of literature reviewed, there is no machine learning method that has achieve 100% accuracy. Each algorithm only utilizes limited features and properties for classification. Therefore, identifying the best algorithm is an important task as their strengths need to be weighed against their limitations. In this paper we explored different machine learning techniques relevant to the spam detection and discussed the contributions provided by researchers for controlling the spamming problem using machine learning classifiers by conducting a comparative studies of the selected machine learning algorithms such as: Naive Bayes, Clustering techniques, Random Forest, Decision Tree and Support Vector Machine (SVM)

***Keywords*—Spam Image, Email Classification, Filtering Techniques.**

I. INTRODUCTION

Email almost serves as a requirement for e-transactions. Sending and receiving e-mails have continued to take the lead being the easiest and fastest way of e-communication despite the presence of different types of e-communications. The rise in the applications of email and online transaction through emails has globally contributed to high rate of email spamming which has been a major problem in the field of computing. There are many machine learning methods

bridge security measures. This violence can be used to abuse the client data and take their valuable sensitive data such as passwords and financial details [1, 3].

The latest survey study on email server revealed that 60% of all email traffic is spam, therefore making it mandatory to create an anti-spam filters. The current spam filters are developed for detecting different spam mails based on the features. In particular, the technique of text categorization is used to filter email spam. But spammers has employed a new way of succeeding the available filters by attaching a textual based content on image in the mail, experiencing image spam a another trick which is so far the most modern kind spam mail with obfuscation. Notwithstanding, emails have continue to maintain success in the area of online business transaction and are now are now a necessity for other means of online communication. Practically, almost all human uses emails. The author in [12] estimated that by the end of 2020, next to half of the global population expected to use emails.

The emails popularity and increase in its application for electronic communication has resulted to an increase in the amount of spam emails globally. Spam emails which are also known as junk emails are unsolicited message content sent by email to several recipients and not requested. Researchers in [13] opined that the spammers had no previous relationship with the recipient but send the spam mails on destructive purpose after collecting addresses from various sources such as tagged filled forms, phone book and spam messages. Spamming is a rapidly growing means of attack such as phishing, worms and virus as the most dangerous

Cyber Kings