Center for Open Access in Science

Open Journal for

Information Technology

2022 ● Volume 5 ● Number 1

https://doi.org/10.32591/coas.ojit.0501

ISSN (Online) 2620-0627

OPEN JOURNAL FOR INFORMATION TECHNOLOGY (OJIT)
ISSN (Online) 2620-0627
www.centerprode.com/ojit.html
ojit@centerprode.com

Publisher:

Center for Open Access in Science (COAS)
Belgrade, SERBIA
www.centerprode.com
office@centerprode.com

Open Journal for Information Technology, 2022, *5*(1), 1-40.
ISSN (Online) 2620-0627

CONTENTS

Center for Open Access in Science ▪ https://www.centerprode.com/ojit.html
Open Journal for Information Technology, 2022, *5*(1), 1-16.
ISSN (Online) 2620-0627 ▪ https://doi.org/10.32591/coas.ojit.0501.01001o

Machine Learning Techniques, Features, Datasets, and Algorithm Performance Parameters for Sentiment Analysis: A Systematic Review

Bernard Ondara, Stephen Waithaka & John Kandiri

Kenyatta University, Nairobi, KENYA
School of Engineering and Technology

Lawrence Muchemi

University of Nairobi, KENYA
Department of Computing and Informatics

Received: 29 November 2021 ▪ Revised: 20 February 2022 ▪ Accepted: 28 March 2022

Abstract

The purpose of this paper is to review various studies on current machine learning techniques used in sentiment analysis with the primary focus on finding the most suitable combinations of the techniques, datasets, data features, and algorithm performance parameters used in most applications. To accomplish this, we performed a systematic review of 24 articles published between 2013 and 2020 covering machine learning techniques for sentiment analysis. The review shows that Support Vector Machine as well as Naïve Bayes techniques are the most popular machine learning techniques; word stem and n-grams are the most extensively applied features, and the Twitter dataset is the most predominant. This review further revealed that machine learning algorithms' performance depends on many factors, including the dataset, extracted features, and size of data used. Accuracy is the most commonly used algorithm performance metric. These findings offer important information for researchers and businesses to use when selecting suitable techniques, features, and datasets for sentiment analysis for various business applications such as brand reputation monitoring.

Keywords: sentiment analysis; machine learning technique; machine learning algorithm; sentiment classification technique; sentiment classification algorithm.

1. Introduction

The proliferation of Internet and Mobile technologies has led to the rapid adoption of social networks and micro-blogging sites. Equally, there is a rising trend in sharing user views and experiences about products and services daily on the Internet. For example, when customers want to purchase a new product or service, they often check on the reviews left behind by previous customers. By the same token, companies can access product and service reviews made by their clients through the different micro-blogging and social media networks like Twitter, Instagram, Facebook, YouTube, and LinkedIn to help them discover and eventually make improvements on the targeted products and/or services and/or make better business decisions (Ahmad et al., 2018). It is not feasible for humans to read all the posts (tweets, comments, and reviews) made by clients.

Correspondence Bernard Ondara, Kenyatta University, School of Engineering and Technology, Nairobi, KENYA. E-mail: ondara.Bernard@ku.ac.ke.

Consequently, this has become a fertile area for research, particularly in the design of automated computer algorithmic techniques for carrying out two popular and important tasks: text classification and sentiment analysis.

- Naive Bayes and Support Vector Machines are the two most popularly used machine learning techniques in sentiment analysis.
- Twitter data and reviews data are the most popular datasets used in sentiment analysis with machine learning techniques.
- The most popular features in sentiment analysis using machine learning techniques are Word stem, n-grams, and Bag of Word.
- Accuracy is the most commonly used algorithm performance metric in sentiment classification where machine learning techniques are used.

Sentiment analysis involves classifying a particular text into positive, neutral, or negative classes. From published literature, there exist three approaches to sentiment analysis. These include Lexicon-based, machine learning, and the blended technique called ensemble/hybrid approach, which bears some aspects from the other two approaches. Various Lexicon-based sentiment analysis techniques and tools have been explored (Ahmad et al., 2017b). Diverse machine learning tools and techniques for performing sentiment analysis have been explored and discussed in depth (Ahmad et al., 2017). In an attempt to improve sentiment classification and general sentiment analysis quality and accuracy, research has advanced to involve the combination of lexicon-based and machine-learning-based techniques in what is classically known as ensemble or hybrid techniques (Ahmad et al., 2017a). In the machine learning-based approach, several techniques have been used, including the following: Support Vector Machine (SVM), Naïve Bayes (NB), Random Forests (RF), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), Decision Trees (DT), Logistic Regression, and Maximum Entropy (ME).

These algorithms generally belong to a class of algorithms called classification algorithms. These algorithms also fall in the group of supervised machine learning algorithms. The algorithms must first be trained using a pre-identified classes of outputs called training data and thereafter gain the capability of classifying real input data called test data. For purposes of sentiment analysis and text classification, many annotated datasets exist for different application domains. These datasets include Twitter data, Amazon product review dataset (Jindal et al., 2008), gender classification dataset (Mukherjee & Liu, 2010), and customer review dataset (Hu & Liu, 2004), among others. This study examined 235 papers on sentiment analysis utilizing numerous machine learning techniques published from 2013 to 2020. We especially used three online libraries: Science Direct, Academia.edu, and Research-gate. Based on the specific query strings used, we identified 235 articles. Upon applying the systematic review framework outlined in sections 3.3 and 3.4 of this study, 24 papers that met our inclusion criteria were selected for a thorough and comprehensive review.

This systematic review paper is structured in the following fashion. Section 2 defines the related works in this research sphere. Section 3 outlines the research methodology employed in this study. Section 4 gives a detailed review of the carefully selected papers. Section 5 is the discussion of the key results of this comprehensive review. Lastly, section 6 concludes this review paper.

2. Related work

Sentiment analysis is a research area that involves analyzing the sentiments, opinions, attitudes, emotions, appraisals, and evaluations of people towards entities like products, services, people, topics, brands, and their attributes (Liu, 2015). An opinion has the following four attributes:

- Object – the target of the opinion;
- Aspect – an object's targeted attribute;
- Sentiment orientation – what indicates if the opinion is positive, neutral, or negative;
- Opinion holder – this is the individual or party that articulates an opinion.

Considering the above attributes of an opinion (quintuple), sentiment analysis presents a very thought-provoking research area with multifaceted tasks (Kharde & Sonawane, 2016). Sentiment classification, subjectivity classification, aspect extraction, spam detection, and lexicon creation are some of the most frequently studied sentiment analysis tasks.

2.1 *Granularity levels in sentiment analysis*

Three sentiment analysis levels exist. They include aspect-level together with sentence-level and document-level (Saranya et al., 2016). At the aspect level, the primary objective is to discover the four attributes of the opinion. Aspect extraction and the subsequent aspect sentiment classification are the two primary tasks defined at this finer-grained level of sentiment analysis. A positive, neutral, or negative opinion on a given object is referenced to an object's attribute and not the entire object.

At the sentence level, the focus is to identify if a sentence carries a sentiment or not, in addition to assessing the sentimentality of particular independent sentences. This granularity is a bit more puzzling because the sentiment orientation is exceedingly reliant on the context in which the word is used. Common challenges in this level include handling sarcasm and comparisons at the sentence level.

The last level is the document-level granularity. In this case, the text is treated as one unit and consequently assigned a positive, neutral, or negative sentiment class. This involves assuming the whole document presents a single opinion about the opinion holder and cannot, therefore, be used where a document compares or evaluates multiple entities.

2.2 *Approaches to sentiment analysis*

Typically, sentiment analysis through machine learning techniques is performed using any of the following approaches or their combinations: unsupervised, supervised, and hybrid approaches. According to Boudad et al. (2018), the supervised machine learning approach uses algorithms like Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbor (KNN) and Artificial Neural Networks (ANN). A huge dataset of categorized data is used to train a sentiment classifier or a set of multiple sentiment classifiers in this approach. The purpose of this is to develop a sentiment classification model that can predict the opinion or sentiment in new pieces of text. This is because oftentimes, labeled data is not available, or the process of annotating data is cumbersome.

On the other hand, unsupervised machine learning approaches use words together with their respective sentiments. Every word found in the lexicon bears sentiment polarity scores that signify if the sentiment is positive, neutral, or negative. Lexicons can be created from a current corpus or dictionary. Contrasting this with the supervised approach, the unsupervised machine learning approach does not necessitate using labeled datasets. In its place, it requires an extensive lexicon that includes the largest possible number of sentiment words.

Instead of utilizing either the supervised or unsupervised approach, a few researchers have opted to integrate both of these approaches. The resulting new model, usually known as the semi-supervised approach, employs a huge quantity of unlabeled data and a small quantity of partially labeled data in creating improved sentiment classifiers. Then, the model classifies the unlabeled data employing supervised classifiers that have been trained using labeled data.

2.3 *Common sentiment classifiers / classification techniques*

In this section, we briefly describe some of the most popular machine learning techniques used in sentiment analysis. This list is therefore not exhaustive.

2.3.1 *Naïve Bayes (NB)*

This sentiment classification technique is probabilistic and often performs well when large datasets are used. The classifier must first compute a posterior probability and then assume that the features involved are conditionally independent. Even so, to do away with unwanted effects, smoothing techniques are employed (Kaur, 2016).

2.3.2 *Support Vector Machine (SVM)*

Like the NB model, this model is also probabilistic and must have training data for model training. SVM employs nonlinear mapping to locate a big margin between various classes. The classifier tries to discover a decision margin that makes the most of the separation gap between two classes. SVM classifier is highly accurate even though it requires additional time in model training. Contrasting with the NB classifier, SVM does not make any assumption regarding the conditional independence of classes (Ankit & Saleena, 2018).

2.3.3 *Logistic Regression (LR)*

This model is employed in classification tasks. It is normally used to associate one definite independent variable with at least one independent variable. The LR classifier tries to identify a hyper-plane that makes the most of the separation gap amid classes (Ankit & Saleena, 2018).

2.3.4 *Random Forests (RF)*

The RF classifier is an ensemble of decision-tree based classifiers. As such, the RF classifier creates a set of decision trees from the existing dataset for training. Upon getting votes from the various decision trees, the RF classifier determines the ultimate class or label of an object in the test dataset (Ankit & Saleena, 2018).

2.4 *Datasets*

The following are some of the common Twitter datasets.

2.4.1 *Twitter sentiment analysis dataset*

This particular dataset contains 99,989 tweets for model training. Each of the tweets is labeled as positive or negative. Of this set, 56,457 tweets are labeled as positive, while the remaining 43532 tweets are labeled as negative, according to Ankit and Saleena (2018).

2.4.2 *Stanford-sentiment140 corpus*

This dataset has 1.6 million tweets for model training. This set is evenly divided into positive and negative tweets (Go et al., 2009).

2.4.3 *First GOP debate Twitter sentiment dataset*

Crowdflower hosts this dataset, and it contains tweets about the first GOP debate held for the presidential nomination in 2016. The dataset has 13,871 tweets, with 2,236 being positive, 8,493 being negative, and the remaining 3,142 tweets bearing a neutral sentiment label (Ankit & Saleena, 2018).

2.4.4 *Health care reform (HCR)*

This dataset was collected from Twitter using the #hrc search tag (Go et al., 2009). It has 888 tweets, of which 365 are positive, and 523 are negative.

2.5 *Algorithm performance parameters*

Based on the review conducted, the performance of different sentiment classifiers is based on the following algorithm performance evaluation metrics: Accuracy, Recall, Precision, and F-score, whose formulae are given below (Ahuja et al., 2019). However, it was observed that most studies focus on accuracy alone while it is our understanding that the other three parameters could be used to give a better estimation of the performance of sentiment classification techniques.

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$\text{F1} = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

Accuracy measures the percentage of correctly classified cases. Precision measures that ratio positive cases to the sum of all positive cases. Recall measures the number of correctly cases classified into a class in relation to all cases present. F-score is a measure of the weighted average of recall and precision. Where classes are unevenly distributed, this parameter is the best to use in gauging algorithm performance (Ahuja et al., 2019).

3. Research methodology

This study was carried out to obtain key information from the most current research publications on sentiment analysis with particular emphasis on machine learning techniques, features, datasets and algorithm performance metrics used. The articles are published in the last 6-7 years.

Systematic reviews scrutinize gaps amid various researches over some time (Kitchenham et al., 2009). A research methodology outlines the step-wise structure followed in carrying out the study. According to Brereton et al. (2007), an in-depth procedure, specific structure, and boundary lines should be used to guide the process of selecting the most pertinent study articles that bear the utmost quality. The methodology adopted in preparing this systematic review is provided in this section. The exact steps followed consist of searching for relevant publications, publication inclusion criteria, publication exclusion criteria, and the findings in the publications. A Systematic Literature Review Guidelines document in the domain of software engineering was equally important in this regard (Ashraf & Aftab, 2017; Ashraf, 2017; Anwer & Aftab, 2017).

3.1 *Research questions*

To reflect the primary objectives of this study, especially while carrying out the systematic review of the selected articles, five research questions required to be answered:

RQ1: Which current machine learning techniques are the most popular in sentiment analysis?

RQ2: What machine learning techniques are particularly applied in Twitter sentiment classification?

RQ3: What kinds of datasets are commonly used for performance evaluation?

RQ4: What are the main features of the datasets for Twitter sentiment analysis?

RQ5: What are the key algorithm performance metrics for sentiment classifiers?

3.2 *Query string and search space*

A query string refers to the blend of certain keywords applied to extract search publications or articles from the chosen online libraries. The specific keywords that were extracted from the six research questions include: Machine learning techniques, Sentiment analysis, Twitter sentiment analysis, machine learning techniques. The following search query was completed using the above keywords:

(("Machine learning techniques" OR "Machine learning algorithms") AND ("Twitter sentiment analysis"))

The following popular online search libraries were mainly used: Science Direct, Research Gate, and Academia.edu. Since these libraries differ in search characteristics, we adjusted the query strings slightly to ensure that more appropriate articles are extracted. This means that the same query was applied at different times, with each iteration having some slight modifications to the keywords' arrangement.

3.3 *Selection criteria*

This section involves articles Inclusion Criteria (IC) rules:

IC1: Articles published from 2013 to 2020;

IC2: Articles that used one machine learning technique in sentiment analysis;

IC3: Articles that used hybrid techniques that involved various machine learning algorithms;

IC4: Articles with experiments and results;

IC5: Articles written in English;

Besides the inclusion above, only those articles bore greater relevance to the research questions considered for this study.

3.4 *Quality assessment*

For purposes of achieving high-quality results, the following specific quality assessment parameters were considered:

- Credible electronic scientific libraries were used to extract the appropriate research materials
- To ascertain the highest quality of results, only the most recent publications were considered
- The process of selecting the articles was unbiased.
- The systematic literature review steps discussed above were strictly followed.

3.5 *Data extraction and synthesis*

Upon completing the search process, 24 most relevant publications were identified.

4. Literature analysis

The selected papers were analyzed and the results presented in the tables below. The results are presented based on the machine learning technique(s) the studies used, features extracted, highest accuracy attained, and the article's reference. The reason for this breakdown is to help identify what machine learning techniques, features, and algorithm performance parameters are used for different datasets and social media platforms. It makes it easier to analyze, for instance, the techniques, features and accuracy levels of the techniques used for different social media platforms like Twitter and Facebook.

4.1 *Synthesis of the experimental results*

Tables 1-6 represent sentiment classification techniques using machine learning approach.

Table 1. Twitter dataset

Machine Learning Technique	Features Extracted	Accuracy	Reference
SVM, Back-Propagation Neural Networks (BPNN), NB, Decision Tree	Word stem	96.06%	Hammad & Mouhammd, (2016)
SVM, NB	Word stem	95% F-score	Duwairi (2015)
NB, DT	Word stem	64.85%	Al-Horaibi & Khan (2016)
SVM, NB, KNN	Word stem, n-grams	67%	Duwairi & El-Orfali, 2014)
SVM, NB, KNN	Word stem, n-grams	69.97%	Duwairi & Qarqaz (2014)
NB, SVM, MaxEnt, ANN	Unigram, Bigram, Hybrid of Unigram and Bigram	92% for SVM with PCA	Anjaria & Guddeti (2014)
MNB, SVM, BNB, Passive Aggression KNN, LR, SGD	n-grams	69.1%	Nabil et al. (2015)
Prind, KNN, NB, SVM, RF, NBMN (NB Multinomial)	Text Blob, SentiWordNet, and WSD	79% for NB with WSD; 76% for NB with TextBlob	Hasan et al. (2018)
SVM	Sentence-level features, Standard features, linguistic	95%	Ibrahim et al. (2015)
Complement NB	Sentiment Lexicon features, text-related features, emoticon-based features	79.4%	El-Beltagy & Ali (2013)
NN, Hierarchical and DBSCAN clustering			Stojanovski et al. (2016)
RF	User and network features	*Improved*	Novalita et al. (2019)

Table 2. Facebook dataset

Machine Learning Technique	Features Extracted	Accuracy	Reference
SVM, KNN, BN	Word stem + n-grams	69.97%	Duwairi & Qarqaz (2016)
SVM, Back-Propagation Neural Networks (BPNN), NB, Decision Tree	Word stem	96.06%	Hammad & Mouhammd (2016)

Table 3. YouTube dataset

Machine Learning Technique	Features Extracted	Accuracy	Reference
NB, SVM, DT, based lexicon	Unigram, bigram, word stem, Sentiment orientation weight	95%	Elawady et al. (2015)
SVM, Back-Propagation Neural Networks (BPNN), NB, Decision Tree	Word stem	96.06%	Hammad & Mouhammd (2016)

Table 4. Reviews datasets: Amazon reviews, hotel reviews, book reviews, multi-domain reviews

Machine Learning Technique	**Features Extracted**	**Accuracy**	**Reference**
NB	Word features	75%	Jain et al. (2016)
SVM, ANN, and MaxEnt	Opinion features, stylistic features, discourse markers, morpholexical features, and domain-dependent features	85.06%	Bayoudhi et al. (2015)
Linear SVM, LR, Bernoulli NB, KNN, Stochastic Gradient Descent.	n-grams, Lexicon entries	88%	ElSahar & El-Beltagy (2015)
SVM	Low-level stem	83%	Cherif et al. (2015)
Hierarchical Classification using: SVM, KNN, DT, NB	Bag-Of-Word	57.8%	Al Shboul et al. (2015)
SVM	Standard features, sentence-level features, linguistic features,	95%	Ibrahim et al. (2015)
SVM, NB, Back-Propagation Neural Networks (BPNN), Decision Tree	Word stem	96.06%	(Hammad & Mouhammd (2016)

Table 5. Aljazeera dataset

Machine Learning Technique	**Features Extracted**	**Accuracy**	**Reference**
SVM, KNN, NB	Word stem + n-grams	96%	Duwairi & El-Orfali (2014)
SVM, MaxEnt, and ANN	Opinion features, domain-dependent features, discourse markers, morpholexical Features and stylistic features	85.06%	Bayoudhi et al. (2015)

Table 6. Other datasets: Blogs, Goodreads.com, Subjectivity, NetvizzApp

Machine Learning Technique	**Features Extracted**	**Accuracy**	**Reference**
SVM, NB	Word stem + n-grams	75%	Akaichi et al. (2013)
J48, Decision Table, KNN, SVM, NB, MNB	Word stem	58%	Al Shboul et al. (2015)
NB, Variational Expectation-Maximization (VEM)		84.6%	Adeborna & Siau (2014)
KNN	Link, photo, status update, and video.	82.3%	Poecze et al. (2018)

4.2 *Accuracy of the techniques/algorithms*

In their study, Ibrahim et al. (2015) used SVM alone on Twitter reviews and comments, attaining a 95% accuracy with an extensive feature set. A different study by Cherif et al. (2015) applied SVM to hotel reviews and attained 83% accuracy with feature weights generated using a new mathematical approach. This significant variation in the accuracy shows that SVM accuracy depends on the dataset and features used. In other cases, SVM was used in combination with other machine learning methods. For instance, Duwairi and Al-Rifai (2015) used an SVM and NB ensemble on the Twitter dataset and attained 87% F-measure accuracy.

The lowest SVM performance was recorded by Al Shboul et al. (2015), who applied it to the goodreads.com dataset and attained a 58% accuracy but with using multi-way classification where other algorithms such as KNN, MB, MNB, Decision Table, and J48 were also studied. This is a classic example of where ensemble methods perform lower than independent methods. The highest accuracy was obtained where SVM was used in the study by Hammad and Mouhammd (2016), who used SVM, BPNN, NB, and Decision Trees on various datasets, and this resulted in 96.06% accuracy with POS tagging.

In the Naive Bayes (NB) method, the highest accuracy recorded in the papers reviewed is also in the same study where SVM performed the best according to Hammad and Mouhammd (2016). NB was used alone, such as in the study by Duwairi and Qarqaz (2014) using Amazon reviews dataset and word features, the accuracy was 75%. A variation of NB, Complement Naïve Bayes, was applied in a study by ElSahar and El-Beltagy (2015) with an extensive set of Twitter features that resulted in 79.4% accuracy. The lowest recorded performance of NB in this study by Bayoudhi et al. (2015). Just as was the case with SVM, this is one reason that many studies argue that SVM and NB have comparatively similar accuracy levels.

It is worth noting that other machine learning methods have good accuracy levels even though they are not popular. For instance, the Random Forest (RF) method achieved 84.0% accuracy, while the KNN method achieved 82.3% accuracy in a study by Poecze et al. (2018). Unfortunately, other techniques such as Decision Trees, Logistic Regression, and Maximum Entropy were not used alone in any of the 24 studies; hence difficult to get their independent accuracies.

4.3 *Datasets used*

13 out of the 24 papers reviewed used Twitter as a source of data. Other studies used different data sources, including Aljazeera (Duwairi & El-Orfali, 2014), Facebook and Blogs (Akaichi et al., 2013), Amazon Reviews (Jain et al., 2016), goodreads.com (Al Shboul et al., 2015), Aljazeera movie reviews (Bayoudhi et al., 2015), YouTube comments (Elawady et al., 2015; Baccouche et al., 2019), Multi-domain reviews (ElSahar & El-Beltagy, 2015), hotel reviews (Cherif et al., 2015), book reviews (Al Shboul et al., 2015), subjectivity dataset (Hammad & Mouhammd, 2016; Poecze et al., 2018), and a combination of Hotels reviews, Facebook, Twitter, and YouTube (Hammad & Mouhammd, 2016).

4.4 *Features extracted*

Different studies used different features from their datasets. The popular feature, Word stem, was used in the studies by Duwairi and Al-Rifai (2015), Duwairi and El-Orfali (2014), Duwairi and Qarqaz (2016), Akaichi et al. (2013), Duwairi et al. (2015) and Al Shboul et al. (2015). Another common feature, n-gram, was used in the following studies: Duwairi and El-Orfali (2014), Duwairi and Qarqaz (2016), Akaichi et al. (2013), Nabil et al. (2015), ElSahar and El-Beltagy (2015), Duwairi et al. (2015), and Aldayel and Azmi (2015).

Some of the reviewed papers also used features that were not common or popular among the reviewed papers. These include Bag of Words (BoW), Bigrams, unigrams, word features, opinion features, Sentiment Lexicon features, sentence-level features, emoticon-based features, discourse markers, user and network features, stylistic features, text-related features, domain-dependent features, and morpholexical features, Text Blob, SentiWordNet, and WSD, Sentiment orientation weight, Lexicon entries, Low-level stem, Standard features, linguistic features, and others.

5. Results and discussions

The discussion of our findings below is aligned with the research questions of this study.

RQ1: Which current machine techniques are the most popular in sentiment analysis?

The papers reviewed machine learning techniques used in sentiment analysis. Out of the 24 papers reviewed, 20 articles used Naïve Bayes (NB) technique in its simple form or other enhanced forms like Complement Naïve Bayes (CNB), Bernoulli NB (BNB), or Multinomial Naïve Bayes (MNB). This qualifies the Naïve Bayes technique as the most popular (83.3%) machine technique for sentiment analysis based on the papers reviewed. This was closely followed by Support Vector Machine (SVM) that was used in 16 of the 24 papers. This is the popularity of 66.67%. Interestingly, SVM was not used in other forms apart from a particular use of Linear SVM in one study.

The other machine methods attracted a usage popularity of less than 50%. For instance, K-Nearest Neighbor (KNN) method was used in 8 studies, which presents a 33.33% popularity among the considered studies. This was followed by Decision Trees that was used in 5 studies (20.83% popularity), Artificial Neural Networks (NN) was used in 4 studies (16.67% popularity), and Random Forests (RF) was used in 3 studies representing a 12.5% popularity. Both Maximum Entropy (ME) and Logistic Regression (LR) methods were used in 2 studies, representing a popularity of 8.33%.

RQ2: What machine learning techniques are popularly applied to sentiment classification on Twitter?

Of the 24 articles reviewed, 13 of them used Twitter datasets, which confirms the popularity of Twitter as a data source for sentiment analysis among popular studies. Interestingly, this review observed a positive correlation between machine learning techniques generally used for sentiment analysis and the those specifically used for Twitter sentiment analysis. In particular, this review observed that nine of the reviewed papers used SVM. Nine of the reviewed papers equally used the NB technique. Both algorithms represent a 69.23% popularity among researchers in Twitter sentiment analysis. This confirms the popularity of these two machine techniques among researchers in sentiment analysis. The other techniques in their decreasing popularity are KNN, ANN, RF, DT, MaxEnt, LR. The RF and DT techniques tied at 2 articles each while MaxEnt and LR tied at 1 article. This still confirms that SVM and NB are the most popularly or commonly used techniques for Twitter sentiment analysis.

RQ3: What kinds of datasets are commonly used for performance evaluation?

The 24 articles reviewed in this paper featured the application of different datasets from various platforms. This included Twitter, Facebook, YouTube, Amazon Reviews, Hotel Reviews, and Movie Reviews. Twitter social media network is the most popular as it offered the most popular dataset used in 12 of the studies, representing 50% popularity. This was followed by reviews (books, movies, hotels), representing 25% popularity of this study. Facebook was used in 3 studies, which is 12.5% in popularity, while YouTube had 8.33% popularity. This shows that Twitter is the most popular platform for studies on sentiment analysis. This is because Twitter data is easily accessible, available, and the quantity of tweets is good for model development, training, evaluation, and implementation.

RQ4: What are the main features of the datasets for Twitter sentiment analysis?

The papers reviewed contained machine learning techniques that exploited various features. The list of features used in the studies includes word stem, n-grams, unigrams and bigrams, hybrid unigrams and bigrams, word features, opinion features, Bag of Words (BoW), discourse markers, stylistic markers, morpho-lexical features, sentiment orientation weights, lexicon entries, low-level stem, domain-dependent features, and standard features among others. The most popular features in their order of decreasing popularity are Word stem, n-grams, and Bag of Word tying with unigrams/bigrams or a combination of unigrams with bigrams. The other features appeared only in one study each. The popularity of word stem was 33.33%, featuring in 8 of the 24 reviewed papers. In comparison, n-gram followed closely with a popularity of 20.83% as both n-grams, and unigram/bigram each scored 8.3% popularity.

RQ5: What are the key algorithm performance metrics for sentiment classifiers?

With reference to section 2.5, this study identified accuracy as the main and most popularly used measure of the performance of machine learning techniques in sentiment analysis. However, it should be noted that other parameters like Recall, Precision and F-score could be used to select better machine learning algorithms for different application areas in sentiment analysis. For example, if data classes are unevenly distributed, then it is best to use F-score parameter.

Limitations of this study

The following are the limitations of this study:

1. The performance of the machine learning algorithms reviewed was reported by various researchers who tested them. These researchers' actual performance or accuracy levels may be inaccurate, and this may generally affect the analysis of this research findings.

2. While a stringent approach was used in selecting the 24 papers reviewed in this study, there may be a few research works relevant to this study that was still left out. This may arise from a few factors, including but not limited to the inclusion criteria that were used in this study.

3. The reviewed articles showed varying performances of the same machine learning technique, making it challenging to pinpoint the best performing technique.

4. This review was limited to conventional machine learning techniques. There was no coverage of deep learning techniques.

This paper came up with five key findings that could contribute greatly to the advancement of sentiment analysis in the future. First, whereas different methods have been used to carry out a sentiment analysis on data from different social media platforms, this review establishes that the commonest machine learning techniques in the decreasing order of their popularity are: Naïve Bayes (NB), Support Vector Machines (SVM), and K-nearest neighbor (KNN), Decision Trees, Artificial Neural Networks (ANN) and Random Forests. Of these, the most popular three are NB, SVM, and KNN. Other techniques such as were used in one article and therefore considered very rare. It was noted that the choice of the technique was dependent on the data. NB and SVM have similar accuracy levels. Researchers' concern is to factor in text structure, data volume, and duration taken for model training and running. Machine learning techniques work well even with huge quantities of data, which requires more time to train the model than lexicon-based approaches. However, to improve the accuracy and quality of sentiment classification and analysis results, we suggest the application of ensemble approaches that involve combining machine learning methods with lexicon-based methods.

Second, this review shows that the commonest machine learning techniques used in carrying out sentiment analysis on different datasets are equally the commonest techniques used on Twitter datasets. They include NB, SVM, KNN, and ANN. This finding provides a ready recipe for those interested in studying these popular algorithms or even combining them to achieve better results by creating hybrid models that benefit from the advantages of the techniques combined while diminishing their disadvantages or limitations when used independently. Moreover, these findings give researchers room to investigate unpopular machine learning techniques for sentiment analysis.

Third, we also identified Twitter as the commonest social media platform for extracting datasets suitable for sentiment analysis throughout this review. This is evident because most of the papers we reviewed used Twitter as their preferred data source. This is because Twitter is available, has rich content in the form of tweets, and is accessible. Checking on any common topic on Twitter could reveal that daily, there are millions of tweets. Conversely, we observed that there is little focus on other social networks like YouTube, Facebook, and WordPress blogs. Although the content and structure of data on these other social media platforms may differ from that on Twitter, this area is worth researching. It may yield new interesting findings and knowledge.

Fourth, this review shows that Word stem and n-grams are the commonest features used in sentiment analysis both on Twitter and other platforms or datasets. It would be interesting to investigate further if other features would yield better results than these common ones hence becoming a research area of interest.

Finally, accuracy is the single-most commonly referenced and used sentiment classifier performance parameter. In all the 24 papers reviewed, accuracy was considered and reported. One paper however simply indicated that the accuracy improved without giving the percentage.

6. Conclusion and future work

Sentiment analysis is currently a hot area of research within the larger knowledge discovery domain. Considering the huge volumes of data generated daily on the different social networking and micro-blogging sites in the form of tweets, posts, comments, and reviews, sentiment analysis techniques are often applied to get useful insights to help with brand reputation monitoring, getting the sentiments of the public regarding a given product or service just before it is launched, predicting election results, and a plethora of other applications. Three approaches to sentiment analysis are available. This systematic review presents research on sentiment analysis on data from social media networks and microblogging websites.

In conclusion, sentiment analysis has seen wide-ranging applications in various areas, including brand reputation monitoring, forecasting political election results, disaster location and response, data security awareness creation, business strategy and quality improvement, disease outbreak monitoring, and perceptions of people towards certain sports. These vast application areas show that sentiment analysis is useful in improving decision-making through gathering and analyzing people's perceptions towards a phenomenon, concept, person, or thing. We recommend that in the future, further studies be carried out to create a universal sentiment analysis model that could be applied to various data types and other social networking sites for purposes of obtaining user sentiments in a bid to expand the application of sentiment analysis in real life.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public commercial, or not-for-profit sectors.

The authors declare no competing interests.

References

Adeborna, E., & Siau, K. (2014). An approach to sentiment analysis – The case of airline quality rating. *Proceedings - Pacific Asia Conference on Information Systems, PACIS 2014*, *363*.

Ahmad, M., Aftab, S., Ali, I., & Hameed, N. (2017a). Hybrid tools and techniques for sentiment analysis: A review. *International Journal of Multidisciplinary Sciences and Engineering*, *8*(4), 28-33.

Ahmad, M., Aftab, S., Ali, I., & Hameed, N. (2017b). Tools and techniques for lexicon driven sentiment analysis. *International Journal of Multidisciplinary Sciences and Engineering*, *8*(1), 17-23.

Ahmad, M., Aftab, S., Bashir, M. S., & Hameed, N. (2018). Sentiment analysis using SVM: A systematic literature review. *International Journal of Advanced Computer Science and Applications*, *9*(2), 182-188. https://doi.org/10.14569/IJACSA.2018.090226

Ahmad, M., Aftab, S., Muhammad, S., & Ahmad, S. (2017). Machine Learning Techniques for Sentiment Analysis: A Review. *Int. J. Multidiscip. Sci. Eng*, *8*(3), 27–32.

Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, *152*, 341-348. https://doi.org/10.1016/j.procs.2019.05.008

Akaichi, J., Dhouioui, Z., & López, M. (2013). Social networks' text mining for sentiment classification: the case of Facebook'statuses updates. *17th International Conference*, 640-645.

Al-Horaibi, L., & Khan, M. B. (2016). Sentiment analysis of Arabic tweets using text mining techniques. *First International Workshop on Pattern Recognition*, *10011*(July 2016), 100111F. https://doi.org/10.1117/12.2242187

Al Shboul, B., Al-Ayyouby, M., & Jararwehy, Y. (2015). Multi-way sentiment classification of Arabic reviews. In *2015 6th International Conference on Information and Communication Systems, ICICS 2015* (pp. 206-211). https://doi.org/10.1109/IACS.2015.7103228

Aldayel, H. K., & Azmi, A. M. (2015). Arabic tweets sentiment analysis – A hybrid scheme. *Journal of Information Science*, *42*(6), 782-797. https://doi.org/10.1177/0165551515610513

Anjaria, M., & Guddeti, R. M. R. (2014). Influence factor based opinion mining of Twitter data using supervised learning. *2014 6th International Conference on Communication Systems and Networks, COMSNETS 2014*. https://doi.org/10.1109/COMSNETS.2014.6734907

Ankit, & Saleena, N. (2018). An ensemble classification system for Twitter sentiment analysis. *Procedia Computer Science*, *132*(Iccids), 937-946. https://doi.org/10.1016/j.procs.2018.05.109

Anwer, F., & Aftab, S. (2017). Latest customizations of XP: A systematic literature review. *International Journal of Modern Education and Computer Science*, *9*(12), 26-37. https://doi.org/10.5815/ijmecs.2017.12.04

Ashraf, S. (2017). Scrum with the spices of agile family: A systematic mapping. *International Journal of Modern Education and Computer Science*, *9*(11), 58-72. https://doi.org/10.5815/ijmecs.2017.11.07

Ashraf, S., & Aftab, S. (2017). Latest transformations in scrum: A state of the art review. *International Journal of Modern Education and Computer Science*, *9*(7), 12-22. https://doi.org/10.5815/ijmecs.2017.07.02

Baccouche, A., Garcia-Zapirain, B., & Elmaghraby, A. (2019). Annotation technique for health-related tweets sentiment analysis. *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*, 382-387. https://doi.org/10.1109/ISSPIT.2018.8642685

Bayoudhi, A., Belguith, L. H., & Ghorbel, H. (2015). Sentiment classification of Arabic documents:

Experiments with multi-type features and ensemble algorithms. *29th Pacific Asia Conference on Language, Information and Computation, PACLIC 2015,* 196-205.

Boudad, N., Faizi, R., Oulad Haj Thami, R., & Chiheb, R. (2018). Sentiment analysis in Arabic: A review of the literature. *Ain Shams Engineering Journal*, *9*(4), 2479-2490. https://doi.org/10.1016/j.asej.2017.04.007

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571-583. https://doi.org/10.1016/j.jss.2006.07.009

Cherif, W., Madani, A., & Kissi, M. (2015). A new modeling approach for Arabic opinion mining recognition. In *2015 Intelligent Systems and Computer Vision, ISCV 2015*. https://doi.org/10.1109/ISACV.2015.7105541

Duwairi, R., & El-Orfali, M. (2014). A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. *Journal of Information Science*, *40*(4), 501-513. https://doi.org/10.1177/0165551514534143

Duwairi, R. M., Ahmed, N. A., & Al-Rifai, S. Y. (2015). Detecting sentiment embedded in Arabic social media - A lexicon-based approach. *Journal of Intelligent and Fuzzy Systems*, *29*(1), 107-117. https://doi.org/10.3233/IFS-151574

Duwairi, Rehab M. (2015). Sentiment analysis for dialectical Arabic. *2015 6th International Conference on Information and Communication Systems, ICICS 2015*, 166-170. https://doi.org/10.1109/IACS.2015.7103221

Duwairi, Rehab M., & Qarqaz, I. (2014). Arabic sentiment analysis using supervised classification. *Proceedings – 2014 International Conference on Future Internet of Things and Cloud, FiCloud 2014, August*, 579-583. https://doi.org/10.1109/FiCloud.2014.100

Duwairi, Rehab M., & Qarqaz, I. (2016). A framework for Arabic sentiment analysis using supervised classification. *International Journal of Data Mining, Modelling and Management*, *8*(4), 369. https://doi.org/10.1504/ijdmmm.2016.10002311

El-Beltagy, S. R., & Ali, A. (2013). Open issues in the sentiment analysis of Arabic social media: A case study. *2013 9th International Conference on Innovations in Information Technology, IIT 2013*, 215-220. https://doi.org/10.1109/Innovations.2013.6544421

Elawady, R., Barakat, S., & Elrashidy, N. (2015). Sentimentanalysis for Arabic and English datasets. *International Journal of Intelligent Computing and Information Sciences*, *15*(1), 55-70. https://doi.org/10.21608/ijicis.2015.10911

ElSahar, H., & El-Beltagy, S. R. (2015). Building large arabic multi-domain resources for sentiment analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9042*, 23-34. https://doi.org/10.1007/978-3-319-18117-2_2

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *Processing*, 1-6.

Hammad, M., & Mouhammd, A. (2016). *Sentiment analysis for arabic reviews in social networks using machine learning*. *Apri*, 131-139. https://doi.org/10.1007/978-3-319-32467-8_13

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for Twitter accounts. *Mathematical and Computational Applications*, *23*(1), 11. https://doi.org/10.3390/mca23010011

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *KDD-2004 – Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *November*, 168-177. https://doi.org/10.1145/1014052.1014073

Ibrahim, H. S., Abdou, S. M., & Gheith, M. (2015). Sentiment analysis for modern standard Arabic and colloquial, ArXiv Prepr. *International Journal on Natural Language Computing (IJNLC)*, *4*(2), 95-105. https://doi.org/10.1109/ReTIS.2015.7232904

Jain, J., Panchal, P., Suryawanshi, N., & Shinde, A. A. (2016). Sentiment analysis using supervised machine learning. *Imperial Journal of Interdisciplinary Research*, *2*(6), 2454-1362.

Jindal, N., Liu, B., & Street, S. M. (2008). Opinion spam and analysis. *Proceedings of First ACM International Conference on Web Search and Data Mining (WSDM-2008).* https://doi.org/10.1145/1341531.1341560

Kaur, J. (2016). A review paper on Twitter sentiment analysis techniques. *International Journal for Research in Applied Science & Engineering Technology*, *4*(X), 61-70.

Kharde, V., & Sonawane, S. S. (2016). Sentiment analysis of Twitter data: A survey of techniques. *International Journal of Computer Applications*, *139*(11), 5-15. https://doi.org/10.5120/ijca2016908625

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7-15. https://doi.org/10.1016/j.infsof.2008.09.009

Liu. (2015). Sentiment analysis: Mining opinions, sentiments, and emotions. In *Cambridge University Press*. https://doi.org/10.14569/ijacsa.2018.090981

Mukherjee, A., & Liu, B. (2010). Improving gender classification of blog authors. *EMNLP 2010 – Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 207-217.

Nabil, M., Aly, M., & Atiya, A. F. (2015). ASTD: Arabic sentiment tweets dataset. *Conference Proceedings – EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, *September*, 2515-2519. https://doi.org/10.18653/v1/d15-1299

Novalita, N., Herdiani, A., Lukmana, I., & Puspandari, D. (2019). Cyberbullying identification on Twitter using random forest classifier. *Journal of Physics: Conference Series*, *1192*(1). https://doi.org/10.1088/1742-6596/1192/1/012029

Poecze, F., Ebster, C., & Strauss, C. (2018). Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts. *Procedia Computer Science*, *130*, 660-666. https://doi.org/10.1016/j.procs.2018.04.117

Saranya, N., Phil, M., & Gunavathi, R. (2016). A study on various classification techniques for sentiment analysis on social networks. *International Research Journal of Engineering and Technology*, *3*(8), 1332-1337.

Stojanovski, D., Strezoski, G., Madjarov, G., & Dimitrovski, I. (2016). Finki at SemEval-2016 task 4: Deep learning architecture for Twitter sentiment analysis. *SemEval 2016 – 10th International Workshop on Semantic Evaluation, Proceedings*, 149-154. https://doi.org/10.18653/v1/s16-1022

Center for Open Access in Science ▪ https://www.centerprode.com/ojit.html
Open Journal for Information Technology, 2022, *5*(1), 17-32.
ISSN (Online) 2620-0627 ▪ https://doi.org/10.32591/coas.ojit.0501.02017d

Effect of Supplier Relationship Management on Organizational Performance: A Case Study of the Plastic Manufacturing Industry in Harare Between 2015-2019

Emmanuel Taengwa Denhere

Chinhoyi University of Technology, Harare, ZIMBABWE
School of Entrepreneurship and Business Sciences

Farai Choga

Zimbabwe Open University, Faculty of Commerce, Harare, ZIMBABWE

Received: 1 February 2022 ▪ Revised: 18 April 2022 ▪ Accepted: 20 May 2022

Abstract

Supplier relationship management (SRM) is the overall coordination, collaboration and information sharing between an organization and its suppliers. The study focused on the effect of supplier relationship management on organizational performance for firms in the plastic manufacturing industry in Harare. This research adopted an interpretivism philosophy and data was collected using open-ended questionnaires and telephonic interviews. The population was derived from plastic manufacturing companies operating in Harare. A purposive sampling technique was used to select twenty participants. The study findings revealed that organizations in the plastic industry enjoyed several supplier relationship management benefits that included information sharing and involvement of suppliers in new product development as these contributed positively to their overall organizational performance. However, firms in the plastic manufacturing industry also encountered supplier relationship management challenges that affected their organization's performance. Challenges such as organizations failing to meet their obligations to the buyer–supplier relationship resulting in negative reactions by suppliers. Supplier relationship management implementation was also constrained by challenges that included the unavailability of supplier relationship management (SRM) team and lack of resources to support the SRM system. It was however noted that organizations in the plastic of supplier relationship management to the organizational performance. These strategies included open communication with suppliers for the purposes of sharing critical information, the involvement of suppliers in the new product manufacturing industry implemented SRM strategies that ensured the positive contribution development and supplier certification as a pre-requisite to supplier engagement. The study, therefore, recommends that organizations in the plastic manufacturing industry continue to develop and maintain clear lines of communication with their suppliers. These would enable suppliers to share information that is critical in their strategic decision-making process. Further recommendations were also made to the effect that organizations in the plastic manufacturing industry should always endeavor to honor their obligation to the buyer-supplier relationship.

Keywords: plastic manufacturing industry, supplier relationship management, organizational performance.

Correspondence: Emmanuel Taengwa Denhere, Chinhoyi University of Technology, School of Entrepreneurship and Business Sciences, Harare, ZIMBABWE. E-mail: etdenhere@gmail.com.

1. Background of the study

1.1 *SRM Global trends*

Unilever an international manufacturing organization with production site doted across the globe demonstrated that SRM was an important aspect of the organization by partnering with Greenpeace Indonesia in a campaign to stop further deforestation of Indonesia forest by palm oil companies in Indonesia that supplied palm oil to Unilever International McDonald (2015). Not only did Unilever protect its brand image but also provided support resources for its suppliers to operate within acceptable environmental standards (McDonald 2015). As a leading automobile manufacturer, Toyota has developed long-term relationships with its key suppliers and as such very few international organizations have a strong reputation in supplier relationship management that matches Toyota (Sutton, 2018).

1.2 *SRM trends in Africa*

In Kenya, Ondieki and Biraori (2015) concluded that the effectiveness of the supply chain management in the public sector was as a result of an effective supplier relationship management. Tangus (2015) studied manufacturing firms in Kenya's Kisumu County and concluded that SRM positively contributed to the organizational performance. Badenhorst-Weiss (2016) in their study on the South African Automotive industry concluded that the automotive industry valued its supplier relationships. In Zambia, Milambo and Phiri (2019) in their study revealed that the non-adherence to good supplier relationship management principles in the Zambian Aviation Industry has caused delays in the supply of aircraft spares giving testimony of the importance of an effective SRM to organizational performance.

1.3 *SRM trends in Zimbabwe*

Although the Zimbabwe government has repeatedly emphasized the importance of the manufacturing sector to the revival of the economy sadly the plastic manufacturing industry is not among the critical sectors that are prioritized for foreign currency allocation by the Reserve Bank of Zimbabwe. This has resulted in the plastic industry firms failing to meet their financial obligations which left the suppliers with no option but to demand payment in advance (http://www.rbz.co.zw, 10.04.20). According to the CZI (2018) survey report capacity utilization of the plastic and packaging manufacturing sector continued on the downward trend from 53% in 2016 to 51% in 2018 with a projected further decline of below 34% beyond 2019 as a result of the current economic challenges being faced in Zimbabwe.

2. Statement of the problem

None of the Zimbabwean studies looked at the effects of supplier relationship management on organizational performance for firms in the plastic manufacturing industry in Harare. This study set out to address this yawning knowledge gap on supplier relationship management on organizational performance of firms in the plastic manufacturing industry in Harare.

3. Research objectives

The objectives were to:

i. determine the benefits of supplier relationship management on organizational performance of firms in plastics manufacturing industry in Harare.

ii. determine the challenges of supplier relationship management on organizational performance of firms in the plastics manufacturing industry in Harare.

iii. establish SRM strategies that are being used by firms in the plastics manufacturing industry.

4. Research questions

i. What are the benefits of SRM to organizational performance of firms in the plastics manufacturing industry?

ii. What are the challenges of SRM to organizational performance of firms in the plastics manufacturing industry?

iii. What are the SRM strategies that are being used by firms in the plastics manufacturing industry?

5. Literature review

5.1 *Social exchange theory*

Social exchange theory (SET) is a broad conceptual paradigm that was developed by Homans in 1958. It is a theory that focuses on the study of social behavior in the interaction of two parties that implement a cost benefit analysis to determine risks and benefits (Cropanzano et al., 2017). Social exchange theory cuts across several social science disciplines such as management, social psychology, and anthropology (Cropanzano et al., 2017). This also includes economic relationships such as the buyer-supplier relationship. Through the process of reciprocity, resources are exchanged as one party tends to repay for the good deeds done by the other party. Muema (2016) highlighted that SET tends to look at the inter-organizational association from an interaction viewpoint concentrating on the social structure of the association rather than the transaction. One of the conclusions of SET in that social interactions are molded through the use of the cost benefit analysis. Parties will only remain in the relationship as long as they are accruing benefits from it (Crapaznano et al., 2017).

Social exchange theory, therefore, becomes a critical element in the development and selection of supplier relationship management strategies (Kingshott, 2006) and is relevant for this research.

5.2 *Benefits of SRM to organizational performance*

By entering, a long-term buyer-supplier relationships organizations would be expected to yield benefits that would have a positive effect on their organizational performance. As stated by the social exchange theory parties will enter relationships that are beneficial to both parties. Some of the expected buyer-supplier relationship benefits are discussed below.

5.2.1 *Information sharing*

Through information sharing, Awan, Kraslawski and Huiskonen (2018), in their study based on 239 companies in Pakistan concluded that SRM enabled the buyer and the supplier to appreciate each other's cultural values. Through SRM organizations and their suppliers have come to realize that their organizations can be much more profitable if they adopt a closer co-operation and implement comprehensive communication in areas such as new product development, quality, engineering, and logistic (Roushdy et al., 2015).

As such SRM creates a platform for clearer communication between the buying firm and its suppliers and enables conflicts to be resolved at their infant stage thereby reducing their impact on the relationship and the organizational performance.

5.2.2 *Joint decision making*

Another benefit that was highlighted by Onyango et al. (2015) was that of joint decision making by both parties to the relationship which they defined as "the degree to which each party penetrates into the other party's organizational boundaries with an aim of improving their organizational performance." Joint decision making is one of the most sophisticated forms of information sharing as it requires a very high level of trust and transparency. This process generally involves the maintenance of information flow, assignment of resources, problem solving, and preparation of detailed activity reports, inter-organizational strategic decisions and plans, (Onyango et al., 2015). Joint decision making would come in the form of deciding on the direction in which a new product development should take Sjoerdsma and Weele (2015). In times of economic crisis buyers and the suppliers would make joint decisions that would ensure minimum impact on both organizations' performance.

5.2.3 *Supplier segmentation*

Manufacturing firm deals with a wide range of suppliers that includes suppliers of raw material, repairs and maintenance spares as well as general goods including service providers. Tangus (2015) while concurring with Onyango (2015) added on one of the benefits of SRM as being able to segment the suppliers according to their level of importance. This enabled the organizations to handle the variety, complexity, and heterogeneity of their supplier base. The organizations are then able to treat each supplier with a uniqueness that results in a positive contribution to the overall organizational performance. The commitment of adequate resources to the relationship would also be made possible. In agreeing with Tangus (2015), Diirr and Cappelli (2018) based on their presentation at the Hawaii International Conference on Systems Science, stated that organizations are able to select a specific type of relationship, considering the identified opportunity for joint work, faced reality and established agreements. This will map the way the supplier relationship will proceed to achieve a common goal for both parties thereby enhancing organizational performance.

5.2.4 *Supplier development programs*

Bai and Sarkis (2016) in their study defined supplier development as a deliberate effort by the buying organization to allocate resources to its suppliers in order to improve the supplier's performance and capabilities. Supplier development according to Bai and Sarkis (2016) can be classified into tangible and intangible. Tangible supplier development included amongst other practical activities such as investment in the supplier's capital equipment as well as the human investment that might involve transferring technical staff to the supplier. While intangible supplier development might include supplier staff training. This will in turn assist the buying organization to achieve its set objectives of delivering a quality end product to the consumer as the supplier would be delivering quality raw materials as a result of the development program. Bai & Sarkis (2016) referred to this as the supplier development investment. Tangus (2015) went on to highlight the fact that SRM enabled organizations to implement supplier development programs that are aimed at improving the supplier's performance and capability. This ultimately resulted is the buying organization being able to meet their short-term and long term needs thereby improving their organizational performance that included delivering customer requirements on time. Due to significant benefits such as improved new product development time, improved capacity utilization, product quality, and reduction in manufacturing cost supply development have continued gained ample popularity with the manufacturing organizations in India (Pradhan & Routroy, 2017). Bai and Sarkis (2016) in their study, however, indicated that while both parties to the development program stood to benefit the buyers and suppliers where reluctant to commit resources to these development programs unless they had a clear vision of gaining profitably from the investments.

5.2.5 *Reduction in inventory levels and costs*

Munyimi and Chari (2018) in their study based on the telecommunication sector in Zimbabwe, agreed with Kumar and Rahman (2016) on the fact that SRM enabled organizations to reduce inventory levels by introducing systems such as JIT (just in time) and reduced time to the market. This resulted in organizations in the private telecommunication sector achieving huge cost savings. Zenir, Findikli and Celtekligil (2018) weighed in by stating that not only did the supplier relationship management through strategic partnership significantly reduced cost but also improved time new products were introduced to the market, increased productivity, enhanced product quality.

5.2.6 *New product development*

Nguyen et al (2018) defined new product development as the transformation of market opportunities into products that are made available on to the market. Sjoerdsma and Weele (2015) in their study stated that supplier relationship management systems allow organizations to involve their suppliers in their new product development as suppliers are able to contribute by providing technical advice that includes technological developments on the supply side as well as the availability of the intended raw materials. The successful launch and acceptability of the new product by the market not only would it benefit the organization but also the suppliers. As the production volume increased so did the demand on raw materials from the suppliers.

5.2.7 *Management of business risk*

Business risk is the exposure an organization has to factors that will affect its bottom financial position (profit), this may be as a result of internal factors or external sources such as a change in regulatory requirements or supplier relations (www.investopedia.com, 2019). Dubey et al. (2018) in their study on supplier relationship management for the economy again concluded that firms with a strong supplier relationship management are able to better manage their business risks through collaborative learning and organizational sustainability and are likely to enjoy a superior organizational performance (Neumuller et al., 2016). Williams and Hausman, (2017) in their study pointed out that the first stage in the business risk management process was to carry out a risk analysis that included risk identification, risk categorization, and risk assessment. The effectiveness of the business risk assessment was largely dependent upon the completeness of the initial business risk identification and risk categorization (Williams & Hausman, 2017).

5.2.8 *Identification of supplier development needs*

Shahzad and Sillanpaa (2015) weighed in with the fact that SRM enabled organizations to identify supplier's developmental needs which they would provide assistance to the supplier. The resultant effect would be an improved delivery of quality goods and services by the supplier which would ultimately contribute positively to the organizational performance. The improved quality delivery by the supplier would also contribute towards the organization's cost reduction (Shahzad & Sillanpaa, 2015).

5.3 *Challenges of SRM on organizational performance*

Although past studies have proved that SRM provides a lot of positive benefits to the organization's overall performance it also has met with its challenges that have affected the organizational performance. According to Kaufmann, Carter and Esslinger (2018) when a buying firm fails to manage its relationship with the supplier by failing to meet its obligation in the contract the supplier will reciprocate in order to protect their interest as the relationship cost would outweigh the benefits incurred. The social exchange theory also states that parties will only remain in a relationship as long as they are enjoying some economic benefits (Crapaznano et al., 2017).

Sjoerdsma and Weele (2015) stated that one of the challenges of SRM on organizational performance is the failure by the buying firm to carry out a comprehensive supplier selection process. In one of their case studies, the buying firm Alpha wrongly assumed that their selected supplier Delta had the development capability that would assist Alpha in their new product development. They later discovered that Delta did not even allocate an engineer to the project as they also assumed it was a routine project. This affected Alpha on the progression of their project.

Butt (2019) in his study on absence of a personal relationship in a buyer-supplier relationship highlighted that SRM can have a negative effect on organizational performance if there is no personal relationship developed between the procurement manager of the buying firm and the sales manager of the supplying firm. Both parties will be reluctant to share information as the level of trust between the two managers will be very low.

Diirr and Cappelli (2018) also pointed out that besides the numerous benefits that SRM brings along to the organization trading on the global market, challenges might also be encountered if the buying firm and its suppliers fail to manage their different cultural values especially for organizations whose suppliers are beyond their country's borders. Misunderstandings may arise which will affect the organizational performance. Awan (2018) weighed in by stating that one of the biggest challenges of SRM is the management of cultural differences across a geographically dispersed location for international organizations. He thus noted that employees should understand and have the capability to handle and react appropriately to situations that arose as a result of cultural differences.

5.4 *Supplier Relationship Management (SRM) strategies*

These SRM strategies would be part of the organization's oval corporate strategies.

5.4.1 *Information sharing as an SRM strategy*

In their study Kurmar and Rahman (2015) highlighted that information sharing and information gathering were among the SRM strategies that organizations would use to strengthen their relationship with their suppliers. Awan, Kraslawski and Huiskonen (2018) supported this strategy as it helped the buying firm and the supplier appreciate each other's cultural values especially when dealing with cross border partners.

Pei and Yan (2019) in their study on cooperative behaviour and information sharing in the e-commerce age' stated that as an effective mechanism to improve information accuracy information sharing has become one of the strategic thrust for senior managers to improving organizational performance. The improved decision-making process is made possible by the use of accurate information from both parties of the commercial relationship thus reducing costs of trial-and-error process which ultimately improves the profitability of the organization (Pei & Yan, 2019). Jermsittipasert and Rungsrisawat (2019) stated that by encouraging information sharing organisations in a relationship would improve their performance as such it is critical for the buying firm and its supplier to develop a suitable communication pattern.

5.4.2 *Supplier development programs as an SRM strategy*

Tukimin et al. (2019) defined supplier development (SD) as a collaboration process between the buying organization and its selected suppliers in order to improve their operations to meet their short-term and long-term supply needs. These supplier development programs (SDP) include training and education of supplier's personal, personnel exchange between the buying firm, and the supplier, raising performance expectations and direct investment in the supplier by the buying firm (Jagtap & Teli, 2017). By engaging in supplier development programs manufacturing organizations would ensure that their suppliers are able to provide them with the best quality material, on time at the right place and right service level. Not only does supplier development have a direct positive effect on supplier quality and organizational performance but it is also important for the organization in achieving world class performance levels which are being demanded by the global market (Tukimin et al., 2019).

5.4.3 *Supplier involvement in new product development*

The involvement of suppliers in the transformation of market opportunities into new products reduced the lead time of new product development Nguyen et al. (2018). Nyamasege and Biraori (2015) and Sjoerdsma and Weele (2015) concurred that involvement of suppliers in joint development program such as new product development was an SRM strategy that would contribute to organizational performance as suppliers would be better prepared to support the new product by also improving their capabilities. Suppliers would also share their knowledge in terms of market developments on the supply side (Kumar & Rahman, 2015). Shirkanh, Keramati and Rezaie (2015) in their study 'Investigating the effects of customer relationship management and supplier relationship management on new product development' highlighted that one of the strategic objectives of SRM in the fast-changing customer demands was to collaborate with suppliers in new product development.

By involving suppliers in new product development, the buying organisation will accrue several benefits such as cost reduction, and reduction in new product development cycle time from the supplier innovativeness (Jermsittipasert & Rungsrisawat, 2019). The involvement of suppliers in new product development would range from simple technical advice to a more complex design of a component, sub-assembly, or system design (Lawson et al., 2015). According to Nguyen et al. (2018), global studies have revealed that new product development has become an important element in an organization as such collaborative relationships have positively impacted the new product development processes.

5.4.4 *Supplier selection process as an SRM strategy*

In recent years supplier selection process has gone beyond selecting suppliers with the lowest acquisition cost only as this has proved that most of these suppliers offering a lower price end up failing to deliver on time resulting in customer dissatisfaction (Fonseca & Lima, 2015). In order for organization to implement an effective supplier relationship management system Bouhnik, Giat and Zarruk (2017) stated that organizations should implement a robust supplier selection, evaluation, and assessment strategy that ensures a long-term relationship with committed suppliers. This is in concurrence with Kumar and Rahman (2015) who listed supplier evaluation and assessment as one of the strategies that organization should implement in order to have an effective SRM that contributes positively to the organization's performance.

5.4.5 *Supplier certification as an SRM strategy*

According to Fonseca & Lima (2015) supplier certification process is now part of the organizations quality management (QM) requirement. In order to put suppliers under pressure to get certification by these public standards organizations, buying firms are now putting it as part of the contractual requirement in any transaction for suppliers to be certified (Jajja et al., 2019). The benefit to the buyer for entering into a relationship with suppliers that are certified by these public standards is that they harness maximum value from the relationship at minimum cost (Jajja et al., 2019).

These certification processes have assisted organizations in their supplier selection process as certification is an indication that the supplier is committed to supplying quality products that are produced under acceptable environmental conditions (Wiengarten et al., 2018).

5.4.6 *Information technology (IT) and SRM*

Mukamutembe and Mulyungi (2018) in their study based on the Skol Breweries Rwanda limited, concluded that ICT greatly improved the buyer-supplier relationship and had a positive effect on the organizational performance of the buying firm and its key suppliers were able to share critical information in real time. This technological integration of SRM applications has allowed cooperating partners to the buyer-supplier relationship to conceptualize best practice in their interaction (Mukamutembe & Mulyungi, 2018). Information technology has also allowed partners in a commercial relationship to access information that would have been otherwise kept omitted by the other party. Collaborative buyer-supplier relationships have benefited from the use of ICT tools such as business intelligence and demand planning tools (Enrique et al., 2018). These tools have allowed organizations in buyer-supplier relationships to collect and analyze data used in developing new products for the market. The ICT tools such as those mentioned above have eliminated the potential negative impact of physical distance between business partners in new product development (Enrique et al., 2018).

5.4.7 *Relationship satisfaction*

In agreeing with Murphy and Sashi (2018), Fehr and Rocha (2018) confirmed that satisfaction is a key factor to any relationship as such collaborative relationship has provided greater benefits as opposed to the transactional oriented relationship. One way buying firms can increase supplier satisfaction is by initiating supplier development programs as these are a sign of commitment by the buying firm to a long-term relationship (Dastyar & Pannek, 2019). The supplier satisfaction will be both economic and non- economic.

The social exchange theory (SET) argues that relationship satisfaction is influenced by relationship elements such as partner's commitment, relationship bonds, and trust in the exchange (Shanka & Buvik, 2019).

6. Research methodology

Interpretivism philosophy was used in this research and a case study design was adopted as the research strategy. The target population was composed of an estimated 122 employees involved in procurement process in firms in the plastic manufacturing industry operating in Harare. The population was derived from plastic manufacturing companies operating in Harare that are listed in Zimbabwe Business Directory (www.thedirectory.co.zw, accessed 10.06.2020). Purposive sampling was employed. Open-ended questionnaires and unstructured interviews were used for data collection.

7. Findings and discussion

7.1 *Response rate*

Most of the interviews scheduled were done and 60% of the questionnaires were returned.

7.2 *Benefits of supplier relationship management to organization's performance*

The participants agreed that joint decision making with their suppliers contributed positively to their organization's performance. This is in line with Onyango et al. (2015) who stated that joint decision making with suppliers improves organizational performance.

7.2.1 *Supplier segmentation*

Most of the participants agreed that supplier segmentation as part of their supplier relationship management has allowed them to effectively manage their suppliers thereby improving their organization's performance. This supports Onyango et al. (2015) study where he stated that as one of the benefits of supplier relationship management, supplier segmentation enabled organizations to classify their suppliers according to their level of importance to the organization. This enabled the organization be able to put more focus on those suppliers that are critical to the organization. This also allowed organization to select a specific type of relationship with the supplier. One of the interviewees stated that:

> "Supplier segmentation enabled them to put more focus on their critical suppliers as well as manage their level of information sharing. This also enabled them to target small suppliers who showed potential for growth."

7.2.2 *Supplier development programs*

Most of the participants agreed to the fact that SRM enabled them to implement supplier development programs. This concurred with Tangus (2015) who highlighted in their study that supplier relationship management enabled organization to implement supplier development programs which were aimed at improving the supplier's performance and capability. This in turn ensured an uninterrupted supply of critical raw materials to the buying organisation there by positively contributing to the organization's performance.

7.2.3 *Reduction in inventory levels*

Participants strongly agreed that supplier relationship management had contributed to the reduction of inventory holding levels through the implementation of inventory management systems. This is in concurrence with Kumar and Rahman (2016) who concluded in their study that supplier relationship management had enabled buying organization to reduce inventory levels as suppliers came up with planned delivery schedules that were based on the buying organization's planned production requirements.

7.2.4 *New product development*

All participants acknowledged that supplier relationship management enabled them to involve suppliers in their new product development. This is in agreement with Sjoerdsma and Weele (2015) who in their study concluded that supplier relationship management allowed organizations to involve their suppliers in their new product development as supplier would weigh in with their technical advice as regard among other things the availability and quality of the intended raw materials.

7.2.5 *Reduction in the organization's business risk*

In agreeing with Dubey et al. (2018) in their study who concluded that firms with a strong supplier relationship management are able to better manage their business risks and are likely to enjoy superior organizational performance, the participants indicated the same.

8. Challenges of SRM to organization's performance

While supplier relationship management presents a lot of benefit to organizations in the plastic industry failure to manage the relation can have a negative effect on the organization's performance as highlighted in the findings below.

8.1 *Failure to meet buyer obligations*

Most participants agreed that they failed to meet their buyer obligations in their buyer-supplier relationship. This as stated by Kaufmann, Carter and Esslinger (2018) has resulted in suppliers reacting by suspending their services or deliveries in order to force the buyer to adhere to their contractual obligations as well as protecting their interest.

8.2 *Supplier's response*

After failing to meet their obligation to the buyer – supplier relationship many participants saw their suppliers' suspending deliveries while few had their credit facilities suspended and suppliers permanently stopping trading with the buyer organization. This is in agreement with Kaufmann, Carter and Esslinger (2018) finding in their study where they stated that when buying firms fail to manage their supplier relationship by failing to meet their obligation the supplier will reciprocate by suspending trading with the buying firm in order to protect their interest. This will not only affect the relationship but also negatively affect the organization performance.

8.3 *Effect of supplier's reaction*

Participants confirmed that the supplier's reaction to the buying organization failure to meet their obligation affected their overall organizational performance and others were not affected by the supplier's reaction. Two interviewees whose organization were not affected by the supplier's reaction stated had this to say:

> "They were able to look for alternative suppliers within a very short space of time."

8.4 *Personal relationship between buyer and supplier*

With regards to a personal relationship between the buyer and the supplier's employees, most participants agreed that it had an effect on organizational performance while very few did not agree. Those who agreed indicated that this had a positive effect on the organizational performance yet some indicated that the relationship had a negative effect on organizational performance. These findings are in concurrence with Butt (2019) who in his study highlighted that the supplier relationship management can have a negative effect on the organizational performance if there is no personal relationship developed between buying team of the organization and the sales team of their suppliers. One of the interviewees stated that:

> "The personal relationship that they had developed with sales manager of their supplier has resulted in the building of trust between the two organizations as well as increased information sharing."

8.5 *Cultural differences and organizational performance*

On cultural differences between the organization and its suppliers' participants agreed that cultural difference has an effect on organizational performance if not managed properly while others confirmed that cultural differences between the buying organization and its supplier does not have an effect on their organization performance. According to Diirr and Cappelli (2018), organizations that are trading extensively on the global market face the challenge of cultural differences affecting their organizational performance if not managed properly.

9. SRM Strategies

The researchers found out some SRM strategies as indicated by the participants.

9.1 *Supplier development programs*

Supplier development program was used as a strategy for managing their relationships with their suppliers. This is in line with Kumar and Routroy (2018) findings on their study they carried out on the Indian manufacturing firms. Funds was a limiting factor.

One of the interviewees stated that "their organization could not use supplier development as a supplier relationship strategy because they did not have enough resources to channel towards such programs due to the current economic challenges."

9.2 *Supplier involvement in new product development*

Most of the participants indicated that they involved their suppliers in their new product development as an SRM strategy. This is in line with the research findings by Nyamasege and Biraori (2015) and Sjoerdsma and Weele (2015) who both concurred that the involvement of supplier in the organization's new product development was an SRM strategy the contributed positively to the organization's performance. The above response is a clear indication of the importance firms in the plastic industry place on the involvement of suppliers in their new product development. This guarantees the commitment by suppliers involved in the new product development to supply quality raw materials that will ensure that the end product will meet the customer's expectations. This will intern increase volume sales that will ultimately contribute to the organization's overall performance.

9.3 *Establishing a systematic supplier selection process*

Most participants used a systematic supplier selection process as part of their supplier relationship management strategy. This is in concurrence with Bouhnik, Giat and Zarruk (2017) who in their research finding concluded that organizations should implement a robust supplier selection, evaluation and assessment strategy that ensured the development of a long-term relationship with the suppliers. Five interviewees confirmed that the use of a systematic supplier selection process assisted their organization in engaging suppliers who had a potential of establishing a long-term relationship as well as supplying quality products consistently.

9.4 *Supplier certification as a requirement for supplier engagement*

Half of the responses confirmed that their organizations used the supplier certification as a requirement for supplier engagement. The findings are in line with Fonseca & Lima (2015) who in their study stated that supplier certification was a realization by organizations that suppliers were an important partner whose collaboration had resulted in improved organization's performance. However, three of the interviewees stated that although supplier certification requirement was a very noble SRM strategy majority of the local suppliers did not have the resource that are required to carry out the certification process later alone maintain the certification which required an annual renewal fees. As clearly shown by the responses supplier certification requirement is a very good strategy however lack of resources for local supplier in the plastic manufacturing industry restricted them from using it as an SRM strategy. This is mainly due to the current economic challenges that the country is facing.

9.5 *Use of latest technology in supplier relationship management*

More than half of the organizations adopted the use of latest technology as part of their supplier relationship management strategy. This is in agreement with Makumutemebe and Mulyungi (2018) who in their study based on Skol Breweries Rwanda limited concluded that information technology greatly improved their buyer-supplier relationship and had a positive effect on the organization's performance. Its key suppliers were able to share information on real time.

Two of the interviewees confirmed that they used internet communication facilities such as conference calls, Skype and zoom to conduct meetings with some of their suppliers.

10. Conclusion and recommendations

10.1 *Conclusion*

The following conclusions were made from above finding:

Supplier relationship management enabled organizations in the plastic manufacturing industry to share information with their suppliers, this positively contributed to their organizational performance as they were able to plan their operations in line with any new developments on the critical raw material supply side.

The lack of resource allocation has affected the effective implementation of the supplier relationship management system within organizations in the plastic manufacturing industry.

The current economic challenges in the Zimbabwean economy have contributed to some firms in the plastic industry to default in honoring their obligations in their contracts with suppliers (failed to settle their accounts on time). However, the continual engagement with the suppliers has enabled the organizations to reduce their business risk as suppliers continued to supply critical raw material on upfront payment basis.

Although organizations in the plastics manufacturing industry are trading in a very difficult economic environment the implementation of various supplier relationship management strategies that include, information sharing with suppliers, involvement of suppliers in new production development, use of latest technology in supplier relationship management, a systematic supplier selection process and supplier segmentation has enabled organizations to improve their organizational performance.

10.2 *Recommendations*

Based on the conclusions made above the following recommendations are made to organization in the plastic industry in Harare.

Organization in the plastics manufacturing industry should continue to implement supplier relationship management systems as supplier relationship has an effect on organization performance despite the economic challenges. They should seek to be on the preferred customer list with their suppliers. For them to be able to maximize on the benefit of SRM they should maintain open channels of communication that will allow their suppliers to continue to share information that is vital for them to maintaining a competitive advantage in the packaging industry market.

The procurement personal of the buying organization should seek to establish a closer personal relationship with the supplier's sales team in order to build trust between them that would enable the sharing of critical information among them. This critical information would also assist the organization to gain competitive advantage in the market. It is only natural for a person to share sensitive information with people they trust to be able to handle that information with the confidentiality it deserves.

In order to build a strong relationship with their suppliers, firms in the plastic industries should strive to always honor their obligation to the buyer – supplier relationship (e.g., paying their accounts on time). One way of building trust with your suppliers is to honor your obligation.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public commercial, or not-for-profit sectors.

The authors declare no competing interests.

References

Africa Development Bank (2019). *Zimbabwe Economic Brief,* February 2019. http://www.afdb.org.

Awan, U. (2019). Effects of buyer-supplier relationship on social performance improvement and innovation performance improvement. *International Journal of Applied Management Science*, *11*(1), 21-35.

Awan, U., Kraslawski, A., & Huiskonen, J. (2018). Buyer-supplier relationship on social sustainability: Moderation analysis of cultural intelligence. *Cogent Business Management*. https://doi.org/10.1080/23311975.2018.1429346

Badenhorst-Weiss, J. A., & Tolmay, A. S. (2016). Relationship value, trust and supplier retention in South African automotive supplier chain. *The Journal for Applied Business Research*, *32*(5). RSA.

Bai, X., Sheng, S., & Li, J. J. (2016). Contract governance and buyer-supplier conflict: The moderating role of institutions. *Journal of Operations Management*, *41*, 12-24.

Bai, C., & Sarkis, J. (2016). Supplier development investment strategies: A game theoretic evaluation. *Annals of Operations Research*, *240*, 583-615.

Bouhnik, D., Giat, Y., & Zarruk, I. (2017). Supplier selection and assessment by university officers. *International Journal of Information System and Supply Chain Management*, *10*(1).

Butt, A. S. (2019). Absence of personal relationship in a buyer- supplier relationship: A case study of buyers and suppliers of Logistics services providers in Australia. *Haliyon*, *5*(6), e01799.

Butt, A. S. (2019). Mitigating knowledge hiding in buyer-supplier relationship: An exploratory study. *Wiley Online Library*. https://doi.org/10.1002/kpm.1626

Chari, F. (2016). The significance of supplier relationship management in industrial competitiveness: A case of bakeries in Harare, Zimbabwe. *Journal of Business Management Science*, *2*(1), 29-42.

Charted Institute of Procurement and Supply (CIPS). www.scm-portal.net. Arm's length relationship: Supply Chain Glossary.

Charted Institute of Procurement and Supply (CIPS). www.cips.org.com.

Confederation of Zimbabwe Industry (CZI), manufacturing sector survey 2018. http://www.czi.co.zw.

Cropanzano, R., Anthony, E. I., Daniel, S. R., & Hall, A. V. (2017). Social exchange theory: A critical review with theoretical remedies. *The Academy of Management Annals*, *11*, 1-38. http://dx.doi.org/10.5465/annals.2015.0099

Dastyar, H., & Paanek, J. (2019). Numeric evaluation of game-theoretic collaboration mode in supplier development. *Applied Science*, *9*(20).

Dirri, B., & Cappelli, C., (2018). A systematic literature review to understand cross-organisational relationship management and collaboration. *Proceeding of the 51st Hawaii International Conference on Science Systems*. http://hdl.handle.net/10125/49907

Dubey, R., Gunesekaran, A., Childe, S., Papadopoulos T., & Helo, P. (2018). *Supplier relationship management for circular economy influence of external pressures and top management commitment*. University of Kent. ISSN 0025-1747 UK.

Enrique, D. V., Ayana, F. N., Lima, M. J. R. F., Marodin, G. A., Gzara, L., & Frank, A. G. (2018). The use of ICT to support collaborative product development activities: Evidence from Brazilian Industry. *Production*, Vol. 28, Sao Paulo 2018. https://doi.org/10.1590/0103-6513.20170099

Fehr, L. C. F., & Rocha, W. (2018). Open- book accounting and trust: Influence on buyer-supplier Relationship. *RAUSP Management Journal*. ISSN. 2531-0448. Published by Emerald Publishing Limited.

Fonseca, L. M., & Lima, V. M. (2015). Impact of supplier relationship strategies on organisational performance of ISO 9001 certified organisations. *Quality Innovation Prosperity*, *19*(2). http://hdl.handle.net/10400.22/7308

Jagtap, M. M., Teli, S. M., Raikar, V. A., & Deshmukh, P. M. (2017). Effects of supplier quality management on performance measures of manufacturing organisation. Conference paper, December 2017. www.researchgate.net/publication/322498772.

Jajja, M. S. S., Asif, M., Montanan, F., & Chatha, A. K. (2019). Buyer-supplier relationships and organisational value in supplier social compliance. *Journal of Cleaner Production*, *214*, 331-344.

Jajja, M. S. S., Kannan, V. R., Brah, S. A., & Hassan, S. Z. (2017). Product innovation and business performance: Incites from resource dependence theory. *International Journal of Operations and Production Management*, *37*(8), 1054-1075.

Jermsittiparsert, K., & Rungsrisawat, S. (2019). Impact of strategic sourcing, supplier innovativeness, and information sharing on supply chain agility. *International Journal of Innovation, Creativity and Change*, 5(2), Special edition.

Kaufmann, L., Carter, C. R., & Esslinger, J. (2018). Towards relationship resilience: Managing buyer induced breaches of psychological contract during buyer-supplier project. *Journal of Supply Chain Management*, *54*.

Kmetec, A., Rosi, B., & Kac, S. M. (2019). The importance of partnership in supply chains. *Journal of Contemporary Management Issues*, *24*(2), 95-106.

Kumar, C. V. S., & Routroy, S. (2017). Modelling supplier development barriers in Indian manufacturing firms. *Asia Pacific Management Review*, *23*, 235-250.

Kumar, D., & Rahman, Z. (2015). Sustainability adoption through buyer supplier relationships across supply chain: A literature review and conceptual framework. *International Strategic Management Review*, *3*(1-2), 110-127.

Lawson, B., Krause, D., & Potter, A. (2015). Improving supplier performance in new product development: The role of supplier development. *Journal of Production Innovation Management*, *32*(5).

McDonald, S. (2015). Managing issues through cross-section collaboration: Unilever and Greenpeace. *Crisis Communication in a Digital World 2015*. http://dro.deakin.edu.au/eserv/DU:30072866/mcdonald-managingissues-post-2015.pdf

Milambo, D., & Phiri, J. (2019). Aircraft spares supply chain management for the aviation industry in Zambia based on the Supply Chain Operations reference (SCOR) Model. *Open Journal for Business and Management*, *7*(3).

Mukamutembe, B., & Mulyungi, P. (2018). Role of supplier relationship management on procurement performance in manufacturing sector in Rwanda: As case of Skol Breweries Rwanda Limited. *International Journal of Science and Research*, *8*(5).

Munyimi, T. F., & Chari, F. (2018). The role of buyer-supplier relationship in achieving economic sustainability in the private telecommunication sector in Zimbabwe. *Cogent Business Management*. https://doi.org/10.1080/23311975.2018.1540917

Murphy, M., & Sashi, C. M. (2018). Communication, interactivity, and satisfaction in B2B relationship. *Industrial Marketing Management*, *68*, 1-12.

Muelensteen, T. M., Vermeulen, W. J. V., & Meeman, S. (2016). Creating shared value in buyer-supplier relationship through the implementation of sustainability requirements. *Global Economics and Business review*, *18*(6).

Nyamasenge, O. J., & Biraori, O. E. (2015). Effect of supplier relationship management on the effectiveness of supply chain management in the Kenya public sector. *International Journal of Managing Value and Supply Chains*, *6*(1).

Nguyen, H. A., Nguyen, H., Nguyen, H. T., Phan, A. C., & Matsui, Y. (2018). Empirical study on role of collaboration in new product development in manufacturing companies. *International Journal for Quality Research*, *12*(2), 363-384.

Onyango, O. J., Onyango, M. B., Kiruri, S. N., & Karanja S. N. (2015). Effect of strategic supplier relationship management on internal operation performance of manufacturing firms: A case study of East African Breweries Limited, Kenya. *International journal for Economics, Finance and Management Sciences*, *3*(2), 115-124.

Pei, Z., & Yan, R. (2019). Cooperative behaviour and information sharing in the e-commerce age. *Industrial Marketing Management*, *73*, 12-22.

Pradhan, S. K., & Routroy, S. (2017). Improving supply chain performance by supplier development programme through enhanced visibility. *Materials today: Proceedings,* Vol. 5 Issue 2. Part 1. Pp. 3629-3638.

Reserve Bank of Zimbabwe. www.rbz.co.zw.

Roushdy, M., Mahomed, M., Hesham, S., Elzark, S., & Hafaz, L. (2015). Investigating the impact of supplier relationship management on firm's performance: A multiple case study approach on manufacturing companies in Egypt. *Proceeding of the 2015 International Conference Operations Excellence and Service Engineering*. Orlando Florida. USA.

Shanka, M. S., & Buvik, A. (2019). When does relationship exchange matters? Social bond, trust and satisfaction. *Journal of Business-to-Business Marketing*, *26*(1).

Shahzad, K., Ali, T., Makala, J., Helo, P., & Zaefarain, G. (2019). The varying roles of Governance mechanism on ex post transaction cost and relationship commitment in buyer-supplier relationship. *Industrial Marketing Management*, 131-142.

Sillanpaa, I., Shahzad, K., & Sillanpaa, E. (2015). Supplier development and buyer-supplier relationship strategies – A literature review. *International Journal of Procurement Management*, *8*(1/2).

Sjoerdsma, M., & Weele A. J. (2015). Managing supplier relationship in new product development context. *Journal of Purchasing and Supply*, *21*, 192-203.

Tangus, C. C., Oyugi, L. A., & Rambo, C. (2015). Effects of supplier relationship management practices of performance of manufacturing firms in Kisumu County*, International Journal of Economics, Commerce and Management*. UK, volume III.

Tukimin, R., Hasrulnizzam, W., Mahmood, W., Mohamed, N., Rosdi, M. N. H. M., & Nordin, M. M. (2019). Prioritisation of supplier development practices: A fuzzy method. *International Journal for Recent Technology and Engineering*, *8*(4).

Unilever. www.unilever.com.

Wiengarten, F., Onofrei, G., Humphrey, P., & Fynes, B. (2018). A supply chain view on certification standards: does supply chain certification improve performance outcomes? In *Measuring Operations Performance* (pp. 193 -214). Springer International Publication.

William, S. P., & Hausman, V. (2017). Categorising of business risk of social media. *Procedia Computer Science*, 121, 266-273. Published by Elsevier B.V.

Center for Open Access in Science ▪ https://www.centerprode.com/ojit.html
Open Journal for Information Technology, 2022, *5*(1), 33-40.
ISSN (Online) 2620-0627 ▪ https://doi.org/10.32591/coas.ojit.0501.03033m

Time-Series Prediction of Gamma-Ray Counts Using XGB Algorithm

Vincent Mutuku

South Eastern Kenya University, School of Science and Computing, Kitui, KENYA

Joshua Mwema

South Eastern Kenya University, Kitui, KENYA
Directorate of Information and Communication Technology

Mutwiri Joseph

Tharaka University College, Science Engineering & Technology, Chuka, KENYA

Received: 11 March 2022 ▪ Revised: 13 June 2022 ▪ Accepted: 25 July 2022

Abstract

Radioactivity is spontaneous and thus not easy to predict when it will occur. The average number of decay events in a given interval can lead to accurate projection of the activity of a sample. The possibility of predicting the number of events that will occur in a given time using machine learning has been investigated. The prediction performance of the Extreme gradient boosted (XGB) regression algorithm was tested on gamma-ray counts for K-40, Pb-212 and Pb-214 photo peaks. The accuracy of the prediction over a six-minute duration was observed to improve at higher peak energies. The best performance was obtained at 1460keV photopeak energy of K-40 while the least is at 239keV peak energy of Pb-212. This could be attributed to higher number of data points at higher peak energies which are broad for NaITi detector hence the model had more features to learn from. High R-squared values in the order of 0.99 and 0.97 for K-40 and Pb-212 peaks respectively suggest model overfitting which is attributed to the small number of detector channels. Although radioactive events are spontaneous in nature and not easy to predict when they will occur, it has been established that the average number of counts during a given period of time can be modelled using the XGB algorithm. A similar study with a NaITi gamma detector of high channel numbers and modelling with other machine learning algorithms would be important to compare the findings of the current study.

Keywords: radioactivity, extreme gradient boost, regression, Gamma-rays, photo-peaks, NaITi.

1. Introduction

Radioactivity is the spontaneous emission of energy and particles from unstable atoms of radioactive material. Naturally occurring radionuclides; Potassium-40, Uranium-238 and Thorium-232 comprise terrestrial radiation (NRCC, 1999) and are used to quantify the radiological safety of a given material. Measurement of gamma rays from these radionuclides is mainly done with NaITi or HpGe gamma-ray spectrometry systems which differ in terms of energy resolution, detection efficiency and mechanism of detection (Hossain, Sharip & Viswanathan,

Correspondence: Vincent Mulwa, South Eastern Kenya University, School of Science and Computing, Kitui, KENYA. E-mail: mulwamutukuu@gmail.com.

2012). For a given range of channels in the detector representing a photopeak, the integral sum of the counts of each energy is a very important parameter for quantifying the radionuclide. Counts per second (cps), i.e., the intensity of the radionuclide is obtained by the ratio of the background-corrected integral sum of counts normally referred to as net area to the live time of the gamma-ray counting. The formation of a peak resulting from γ-ray emissions of a certain radionuclide in the sample is primarily a result of Compton scattering and photoelectric absorption from the incident and scattered photons (James & Christine, 2015). Highly-resolved photo peaks are obtained by longer measurement times by accumulating the radiation absorption and scattering events. However, this is determined by the activity of the sample wherein the intensity of the radionuclide is an indicator of the former, i.e., high-intensity samples imply high activity and take a shorter duration to form peaks. A review of some radiation surveys shows that different researchers use different sample run times, e.g., 27.7hrs (Sharma, Singh, Esakki & Tripath, 2016), 23.8hrs (Asaduzzaman, Mannan, Khandaker, Farook, Elkezza & Amin, 2015), 6.1hrs (Aslam, Gul, Ara & Hussain, 2012), and 5.5hrs (Viruthagiri, Rajamannan & S., 2013). While longer measurement times are recommended, treating all samples as low intensity may unnecessarily lengthen the data collection leading to delayed research output especially in developing countries where research equipment are few compared to a large number of researchers. Given that radioactivity is spontaneous, it's impossible to predict when the next unstable atom would decay and emit a gamma-ray. However, when γ-ray counting starts from the time

$$t = 0 \; to \; t1, t2 \; and \; t3 \; all \; the \; way \; to \; t_n$$

an average of n_t counts are registered by the detector at the end of each duration t_n. Since the half-lives of the radionuclides are long enough (Ebbing & Wentworth, 1995; Connell & Pike, 2005), the activity of each radionuclide remains the same within the measurement time. Thus, the intensity (cps) of the radionuclides within a sample material is characteristic and probabilistic. Based on the scattering angle, the energy counts are registered at one of the three regions, Compton continuum, Compton edge or full energy peak. The capability of XGB to deduce the hidden patterns in the interaction events leading to a full energy peak were examined to predict the number of counts in a given time. Accurate prediction of the number of counts in a range of channels could result in a predicted spectrum which implies that shorter measurement times can be adopted to accurately predict counts over a longer time for rapid research and development.

2. Literature review

Studies related to the current study have been reviewed here to understand the current scope of applications of machine learning in nuclear studies.

Klaus and John (1995) described application of neural networks (NN) in predicting probabilities of nuclear stability and relaxation to ground state. In the study, a feedforward network was implemented where the inputs were nuclide parameters which include the proton and neutron numbers. The dynamics of NN weights were managed by a stochastic back-propagation algorithm coupled by entropy function. Whilst NNs were retrained severally using different architectures leading to different models which performed well, it is difficult to obtain a high-quality performance with a global model in regard to the existing nuclear theory.

Niu, Liang, Sun, Long and Niu (2019) investigated the prediction of nuclear β decay using neural networks. Although some physics theories underlying nuclear β such as Fermi theory of β-decay and dependencies of half-lives which include pairing correlations and decay energies were embedded into a Bayesian NN (BNN), other unclear physics were left for the BNN to learn. To the researchers, the high prediction accuracy achieved is very instrumental in simulations involving the r-process.

Empirical formulas in nuclear decays are normally used with less modification as they are conventionally established for computations (Saxena, Sharma & Prafulla, 2021). A study by Saxena, Sharma and Prafulla (2021) shows that inclusion of machine learning in understanding certain phenomena can help modify the existing formulars thereby improving the precision. The researchers showed that adding asymmetry components predicted the half-lives in α-decay with more precision than the empirical formulas. Machine learning methods used include; XGBoost, rando forest, decision trees and multilayer perceptron NN whose results excellently agreed with experimental decay modes. At the same time, S., Freitas and John (2019) predicted the systematics of α-decay of heavy and superheavy nuclei using artificial neural networks (ANN) by backpropagation algorithm with regularization. The investigation highlighted the strengths and limitations of applying machine learning in studying nuclear events beyond stability.

The two body-bound state of deuteron was studied with a single layer feed-forward NN (Keeble & Rios, 2020). The NN successfully represented the S and D state wave functions. Compared with solutions of diagonalization tools, the study's results show that a 6 hidden node NN can seamlessly represent the ground state wavefunction with binding energy that is 0.1% of the theoretical dimensions. It is postulated that this method can pave way for variational ANN to solve nuclear many body problems.

Most of the studies have investigated half-lives and nuclear stability landscape in α-decay using different machine learning methods. The neural networks are the most used in learning the complex concepts in nuclear physics. While α-decay are important, γrays are also very critical in terms of the health effects they cause as they are ionizing radiation. It is important to apply machine learning techniques to have a deeper understanding of the decay behaviour. The number of gamma decays in a given interval is predicted in this study using XGBoost algorithm.

3. Materials and methods

The investigation was implemented in two phases; experimental data collection and machine learning implementation on the spectrum samples as outline in figure 1. A γ-ray spectrometer system comprising; NaITi γ-ray detector, lead shield and a multichannel analyser software was used in gamma-ray counting and acquisition of the sample spectra. The soil samples were prepared according to Sharma, Singh, Esakki, and Tripath (2016), packed in airtight containers with an Aluminum foil reinforced lid, and stored for 30 days to achieve secular equilibrium (Aslam, Gul, Ara & Hussain, 2012).

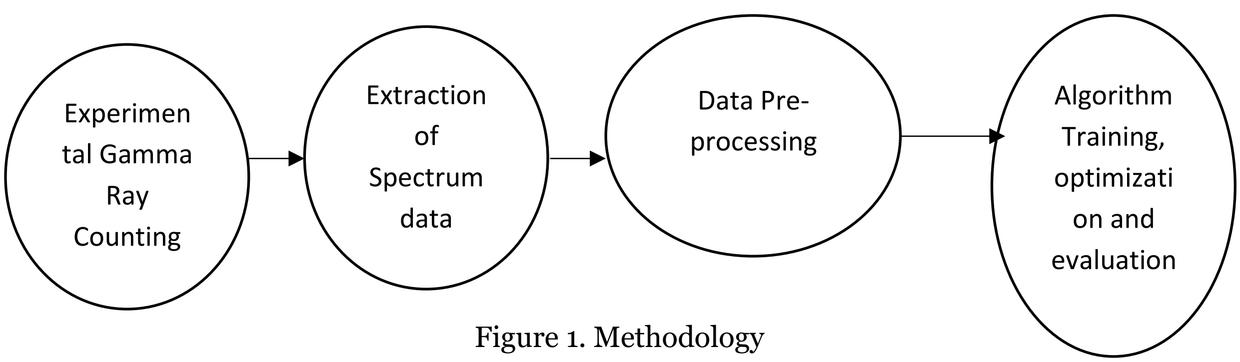

Figure 1. Methodology

Before γ-ray counting, the detector was energy-calibrated to obtain a channel-energy relationship that would help in identifying radionuclide full energy peaks in the spectrum. IAEA certified reference materials were used for both resolution and energy calibration. The sample was placed at the centre of the detector area and the shield covered with its lead lid. A total of 5 spectrums was obtained for measurement times; 4.5hrs, 6.5hrs, 7.5hrs and 7.6hrs. Since the detector has 1024 channels, each spectrum had 1024 instances of gamma-ray counts. As a way of data cleaning, full energy peaks' counts for radionuclides of interest were extracted from the spectrums, i.e., K-40 at 1460keV, Pb-214 at 352keV representing U-238 and Pb-212 at 239keV representing Th-232. Each column in the dataset represented counts of energies at different measurement durations and different channels. Thus, across the row were energy counts on the same channel for different durations. The last column in each dataset was set as the target in the Python program written to implement the XGB regression on the dataset. 80% and 20% of the dataset was used to train and test the model performance respectively. Further, XGB hyperparameter tuning was done to improve the performance of the model after each training. The R-squared value was the main metric to evaluate the model's prediction performance. The optimal model hyperparameters were set as;

Colsample_bytree:0.3

Learning rate: 0.1

Max_depth :5

Alpha:11

N_estimators=3000

The model flow chart is shown in figure 2.

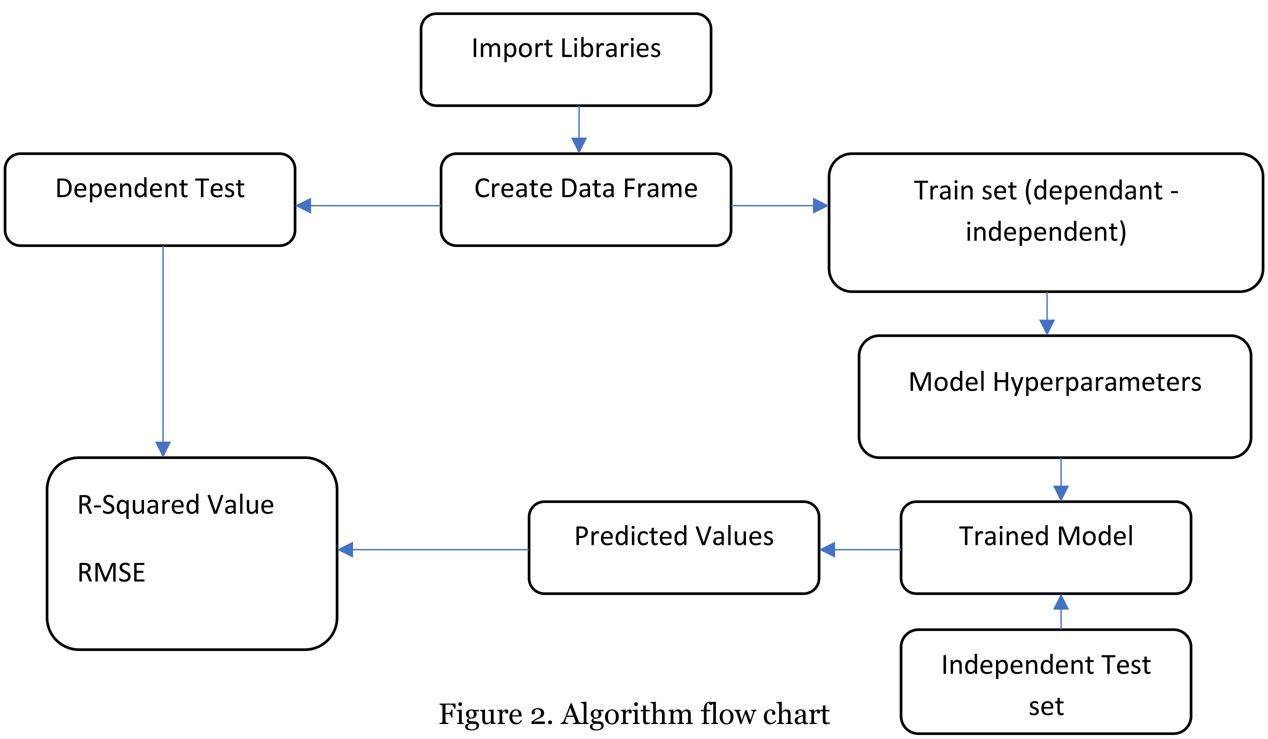

Figure 2. Algorithm flow chart

4. Results and discussion

Peak-wise prediction of the counts shows that the K-40 peak at 0.99 had a higher R-squared value while Pb-212 was the least with 0.92. Resultantly, there is a good agreement between the experimental and the predicted counts. With combined datasets for all the peaks, the R-square is equal to that of K-40. Generally, the R-square value increased with peak energy from 239keV to 1460keV. On the other hand, the RMSE value decreased with increasing peak energy from 59.62 at 239keV to 10.36 at 1460keV. The RMSE value for the combined dataset lies between that of K-40 and Pb-214 where a summary of the model performance is found in tables 1 and 2. The energy resolution, ER, of the detector, was determined according to equation 1 for three gamma-ray energies at their respective full width at half-maximum of the peak height.

$$ER = \frac{FWHM}{(PhotoPeak\ Energy)} \qquad 1$$

The energy resolution reduced from 8.43% at 239keV to 4.38% at 1460keV which compares with (Akkurt, Gunoglu & Arda, 2014) as shown in figures 3 and 4. Among the prediction errors, the greatest is 8.2% for the combined dataset while the least is 0.1% for the K-40 peak. Generally, the maximum and minimum error statistics for the three datasets exhibit a cyclical trend i.e., start slightly high at Pb-212, drop at Pb-214 and K-40, then increase for the combined dataset. This cyclical nature is similar to what is observed in the R-squared value as peak energies increase. On average, the model performed best at the K-40 peak with the lowest average error of 2%. The low energy resolution at 1460keV produced a broad peak for K-40 providing more instances given that larger datasets yield better performance (Althnian, AlSaeed, Al-Baity, Samha, Dris, Alzakari, Abou, Elwafa & Kurdi, 2021). Since the number of channels for the detector is relatively small, 1024, the resulting photopeak datasets were also small in size which explains the overfitting observed in the R-squared values.

Table 1. XGB Regression model performance

Radionuclide	**keV**	**R2**	**RSME**	**Resolution (%)**
Pb-212	239	0.97	59.62	8.43
Pb-214	352	0.92	29.59	5.60
K-40	1460	0.99	10.36	4.38
Datasets combined	na	0.99	24.95	na

Table 2. Prediction Errors

	Prediction Absolute Errors			
	Combined Datasets	**K-40 Peak**	**Pb-212 Peak**	**Pb-214 Peak**
MAX	8.2%	5.0%	5.5%	4.4%
MIN	0.3%	0.1%	1.6%	1.4%
Average	3.5	2.0	3.5	2.5

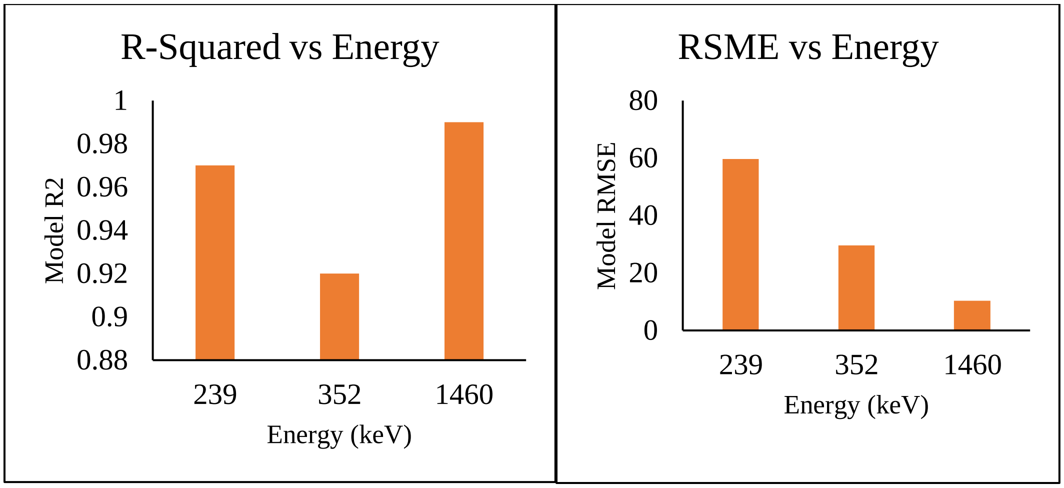

Figure 3. Prediction R-squared AND RMSE Values

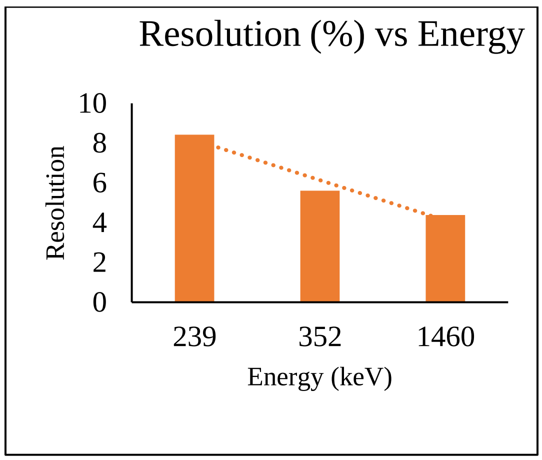

Figure 4. NaITi Energy Resolution (Present data)

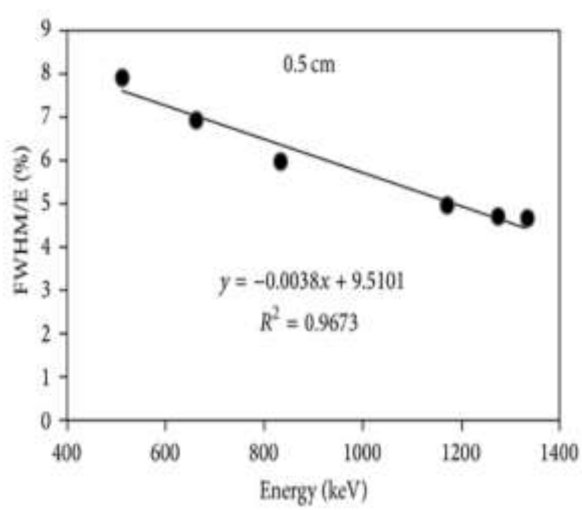

Figure 5. NaITi energy resolution (Akkurt, Gunoglu & Arda, 2014)

5. Conclusions

The prediction performance of the XGB regression algorithm has been evaluated based on experimental and predicted values. The model performs best at higher gamma-ray energies compared to lower ones. The algorithm exhibited excellent fitting capabilities for the gamma-ray counts for 6 mins. It would be important for another study to be done with a similar detector that has a larger number of channels offering larger datasets and investigating overfitting. Additionally, a similar study can be done using a hyper-purity germanium gamma ray spectrometry system to compare the performance of the model between the two systems. Also, other ML algorithms can be tested and their performances compared to the findings of this study. Further research into incorporating machine learning algorithms in scientific works may pave way for the development of more intelligent scientific research software. This may yield rapid research and development across many sectors.

Acknowledgements

This work has been supported by the South Eastern Kenya University, by providing the gamma-ray spectrometry system for the gamma-ray measurements.

The authors declare no competing interests.

References

Akkurt, I., Gunoglu, K., & Arda, S. (2014). Detection efficiency of NaI(Tl) Detector in 511–1332 keV energy range," *Science and Technology of Nuclear Installations,* vol. 2014, 2014.

Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A., Alzakari, N., Abou, A., Elwafa, & Kurdi, H. (2021). Impact of Dataset Size on Classification Performance: An empirical evaluation in the medical domain. *Appl. Sci.*, 11, p. 796.

Asaduzzaman, K., Mannan, F., Khandaker, M., Farook, M., Elkezza, A., & Amin, Y. (2015). Assessment of natural radioactivity levels and potential radiological risks of common building materials used in Bangladeshi dwellings. *PLOS ONE*, 10.

Aslam, M., Gul, R., Ara, T., & Hussain, M. (2012). Assessment of radiological hazards of naturally occurring radioactive materials in cement industry. *Radiation Protection Dosimetry*, *3*(151), 483-488.

Connell, C., & Pike, G. (2005). *Encyclopedia of Analytical Science* (Second Edition).

Ebbing, D., & Wentworth (1995). In *Experiments in Introductory Chemistry*, p. 221.

Hossain, I., Sharip, N., & Viswanathan, S. (2012). Efficiency and resolution of HPGe and NaI(Tl) detectors using gamma-ray spectroscopy. *Scientific Research and Essays*, *7*(1), 86-89.

Keeble, J., & Rios, A. (2020). Machine learning the deuteron. *Physics Letters B*.

Klaus, G., & John, C. (1995). Neural networks that learn to predict probabilities: Global models of nuclear stability and decay. *Neural Networks*, 8(2), 291-311.

James, P., & Christine, C. (2015). *Compton scattering of Cs-137 Gamma rays*. Department of Physics and Astronomy, The University of Tennessee, 401 Nielsen Physics Building, Knoxville.

Niu, M., Liang, Z., Sun, H., Long, W., & Niu, F. (2019). Predictions of nuclear β-decay half-lives with machine learning and their impact on r-process nucleosynthesis, *Physics Review C*.

NRCC (1999). *Natural radioactivity and radiation*. National Academies Press (US), Washington (DC).

S., P., Freitas, A., & John, C. (2019). Experiments in machine learning of alpha-decay half-lives. *Nuclear Theory*.

Saxena, G., Sharma, K., & Prafulla, S. (2021). Modified empirical formulas and machine learning for α-decay systematics. *Journal of Physics G: Nuclear and Particle Physics*, *48*(5).

Sharma, N., Singh, J., Esakki, C. A., & Tripath, R. (2016). A study of the natural radioactivity and radon exhalation rate in some cements used in India and its radiological significance. *Journal of Radiation Research and Applied Sciences*, *9*(1), 47-56.

Viruthagiri, G., Rajamannan, B., & S. J. K. (2013). Radioactivity and associated radiation hazards in ceramic raw materials and end products. *Radiation Protection Dosimetry*, *3*(157), 383-391.

AIMS AND SCOPE

The OJIT, as an international multi-disciplinary peer-reviewed ***online open access academic journal***, publishes academic articles deal with different problems and topics in various areas of information technology and close scientific disciplines (information society, information communication technology - ICT, information architecture, knowledge organisation and management, information seeking, information management, electronic data processing – hardware and software, philosophy of information, communication theory and studies, mass communication, information ethics, library and information science, archival science, intellectual property, history of computer technology, development of digital competencies, ICT in education and learning, ICT education, etc.).

The OJIT provides a platform for the manuscripts from different areas of research, which may rest on the full spectrum of established methodologies, including theoretical discussions and empirical investigations. The manuscripts may represent a variety of theoretical perspectives and different methodological approaches.

The OJIT is already indexed in Crossref (DOI), BASE (Bielefeld Academic Search Engine), Google Scholar, J-Gate, ResearchBib and WorldCat - OCLC, and is applied for indexing in the other bases (Clarivate Analytics – SCIE, ESCI, and SCI, Scopus, Ulrich's Periodicals Directory, Cabell's Directory, SHERPA/RoMEO, EZB - Electronic Journals Library, etc.).

The authors of articles accepted for publishing in the OJIT should get the ORCID number (www.orcid.org).

The journal is now publishing 2 times a year.

PEER REVIEW POLICY

All manuscripts submitted for publishing in the OJIT are expected to be free from language errors and must be written and formatted strictly according to the latest edition of the APA style. Manuscripts that are not entirely written according to APA style and/or do not reflect an expert use of the English language will ***not*** be considered for publication and will ***not*** be sent to the journal reviewers for evaluation. It is completely the author's responsibility to comply with the rules. We highly recommend that non-native speakers of English have manuscripts proofread by a copy editor before submission. However, proof of copy editing does *not* guarantee acceptance of a manuscript for publication in the OJIT.

The OJIT operates a double-blind peer reviewing process. The manuscript should not include authors' names, institutional affiliations, contact information. Also, authors' own works need to be blinded in the references (see the APA style). All submitted manuscripts are reviewed by the editors, and only those meeting the aims and scope of the journal will be sent for outside review. Each manuscript is reviewed by at least two reviewers.

The editors are doing their best to reduce the time that elapses between a paper's submission and publication in a regular issue. It is expected that the review and publication processes will be completed in about 2-3 months after submission depending on reviewers' feedback and the editors' final decision. If revisions are requested some changing and corrections then publication time becomes longer. At the end of the review process, accepted papers will be published on the journal's website.

Open Journal for Information Technology, 2022, *5*(1), 1-40.
ISSN (Online) 2620-0627

OPEN ACCESS POLICY