

Building Predictive Model of Covid 19 Quarantine Impact on the Purchase of Environmentally Green Products by Using J48 and LMT Algorithms

Dr. Mutasim Al Sadig¹

Assistant Professor, Department of CSI, College of Science, Majmaah University,
Al Majmaah 11952, Saudi Arabia.

mm.elkir@mu.edu.sa

ORCID: 0000-0002-5614-4527

Dr. Nahid Osman Ali Babikir²

Department of Business Administration
College of Sciences and Human Studies in Hotat Sudair Majmaah University, Saudi Arabia

na.ali@mu.edu.sa

Dr. Faisal Mohammed Nafie Ali³

Assistant Professor, Department of Computer Science, College of Science and Humanities at Alghat,
Majmaah University, Al Majmaah 11952, Saudi Arabia.

fm.ali@mu.edu.sa

Dr. Khalid Nazim Abdul Sattar⁴

Assistant Professor, Department of CSI, College of Science, Majmaah University,
Al Majmaah 11952, Saudi Arabia.

k.sattar@mu.edu.sa

ORCID: 0000-0002-0759-0512

Abstract

When COVID 19 pandemic appeared, World Health Organization officially announced on January 30, 2020, that the outbreak of the virus constituted a health emergency, and most countries of the world announced a quarantine of all citizens as one of the precautionary to limit the epidemic spread. The Kingdom of Saudi Arabia is one of the countries that announced quarantine, which had an impact on all aspects of life. This study is related to the impact of quarantine on the purchase of green products and effects on the environment, that by using machine learning algorithms. J48 and ML algorithms were used to Build a predictive model to estimate the effect of the quarantine for covid 19 on the purchase of environmentally friendly food products that in Zulfi region of the Kingdom of Saudi Arabia and found that J48 algorithm have highest performance compared to LMT algorithm.

Keywords: *green products; COVID-19; Classification; WEKA; environment*

1. Introduction

The concept of environmentally green products refers to any product or service that is not harmful to the atmosphere or the environment[1]. behavioral change in food consumption has a powerful influence in reducing natural resource use[2], also food consumption is one of the most important side that affect the environment, food that does not take into account the environment is harmful

to both health and the climate[3]. Economic development, continuous population growth, rapid urban development and increased demand for food create environmental challenges all over the world[4]. This is in addition to the increased demand for water and energy. On June 20 of the year 202, the Saudi Ministry of Interior announced the tightening of precautionary measures to confront the Corona epidemic, with the emergence of indications of a rise in the epidemiological curve in some regions of the Kingdom. One of these precautionary measures is quarantine which had an was all having impact on all areas of life, Behavioral, economic, social and health including food consumer behavior[5]. This paper comes on the context of researching quarantine impact on the purchase environmentally friendly food products, by using Data mining techniques and machine learning algorithms using weka tools. supervised learning method was used to Building a predictive model to estimate the impact of Covid 19 quarantine on the purchase of environmentally friendly food products by using J48 and LMT algorithms and compare them in terms of performance. Based on dataset taken from a group of residents of the Zulfi region in the Kingdom of Saudi Arabia, during the period of complete quarantine imposed by the Saudi authorities during the Covid 19 pandemic in the middle of the year 2020.

2. Related work

The Ketelsen, M., M. Janssen, and U. Hamm study, published in the Journal of Cleaner Production, 2020, indicates the importance of food packaging in the purchase of green food products. and the study concluded that consumers' response to green food packaging is not yet well understood[6]. Also, Study Indriani, I.A.D., M. Rahayu, and D. Hadiwidjojo, concluded that environmental knowledge does not have a direct, significant impact on the intent to purchase environmentally friendly products[7]. The research presented by Wang, Y., et al., which was published under the title "How impacting factors affect Chinese green purchasing behavior based on Fuzzy Cognitive Maps", focused on the interaction between the basic functional value of environmentally friendly food, their social value and the cognitive value in the online environmentally friendly food purchase behavior using a fuzzy cognitive map. Consumers paying more is the biggest factor hindering green behavior[8]. Martínez, A., et al. Apply machine learning algorithms to predict future consumer behavior and take advantage of the results to efficiently direct resources in sales and marketing departments[9]. A study focused Awadh, W.A., A.S. Alasady, and H.I. Mustafa. on the impact of quarantine for Covid on demographic, psychological, health, cognitive, awareness and preventive factors. The results of the study showed that quarantine had an important role in controlling the spread of the disease with some negative effects. By comparing the accuracy of the algorithms used, it was found that the best algorithm is J48 [10][15].

3. Weka application

weka is open-source application that include collection of machine learning algorithms for data mining tasks. It contains tools for data preparation, regression, classification, association rules mining, clustering and visualization[6]. It used in supervised and unsupervised classification. Supervised learning is a machine learning process to learn a task that maps inputs to outputs based on the example of input-output pairs. Unsupervised learning is the task of machine learning to infer a function to describe the hidden structure from unclassified data. The main difference between supervised and unsupervised machine learning is that supervised learning uses classified data while unsupervised learning uses unclassified data[7].

4. Research Methodology

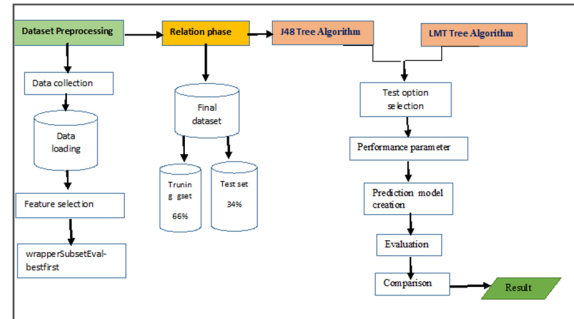


Fig. 1 Model Framework

J48 and LMT decision tree algorithms are used to predict the classification model by using WEKA tools. Figure (1) shows the framework of the prediction model

a) Data preprocessing:

Dataset was taken from random sample of Sudair region population, Kingdom of Saudi Arabia during Coronavirus Quarantine period in 2020 collected through questionnaire with 12 attributes as dependent variables which include sex, Educational level, income level, Buying food products with eco-friendly packaging, Buying products made from natural materials, Increase awareness of green products, Identify companies that support environmental protection organizations, Use of materials that are less harmful to the environment, Understand environmental laws and regulations, Establishing the values of preserving the environment in my buying behavior, My interest in buying environmentally friendly food products ends with the end of the Corona pandemic, Buy products that are associated with environmentally friendly brands and labels that do not cause pollution. Environment conservation attribute used as class (independent attribute). table () Number of instances of dataset are 305 instances. Dataset were collected in csv files format. To improves the quality of data mining tool better Data preprocessing was being taken, making data appropriate for analysis[8]. After data preprocessing, csv file was loaded in weka application and features selection were taken by using evaluates attribute sets by using a learning scheme.

TABLE 1 Attributes Description		
No	attributes	type
1	sex	Nominal
2	Educational level	Nominal
3	income level	Numeric
4	Buying food products with eco-friendly packaging	Nominal
5	Buying products made from natural materials	Nominal
6	Increase awareness of green products	Nominal
7	Identify companies that support environmental protection organizations	Nominal
8	Use of materials that are less harmful to the environment	Nominal
9	Understand environmental laws and regulations	Nominal
10	Establishing the values of preserving the environment in my buying behavior	Nominal
11	My interest in buying environmentally friendly food products ends with the end of the Corona pandemic	Nominal
12	Buy products that are associated with environmentally friendly brands and labels that do not cause pollution	Nominal
13	Class (environment conservation)	Nominal

b) Relation Phase:

in relation phase dataset split into 66% training dataset and 34% for test dataset

c) Classification Phase:

two classification process were taken one classification by using j48 algorithm and the other by using LMT algorithm.

One of the eminent features in the process of data mining is the evaluation of the algorithms. to evaluate the performance accuracy of the proposed classifier when using J48 and LMT algorithms in weka, confusion matrix, learning curves and receiver operating curves (ROC) have been measured. The various evaluation parameters were measured in terms of measured in terms of precision (1), recall (2), F-measure (3), accuracy (4), time taken to build a model, correctly classified instances percent, and incorrectly classified instances percent

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP}) \tag{1}$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \tag{2}$$

$$\text{F-measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{3}$$

$$\text{Accuracy} = (\text{TP}+\text{TN}) / (\text{TP}+\text{TN}+\text{FP}+\text{FN}) \tag{4}$$

Where:

TP = the number of positive cases is classified as positive

TN = the number of negative cases is classified as positive

FP = the number of negative cases is classified as negative

FN = the number of positive cases is classified as negative

5. Experimental Results

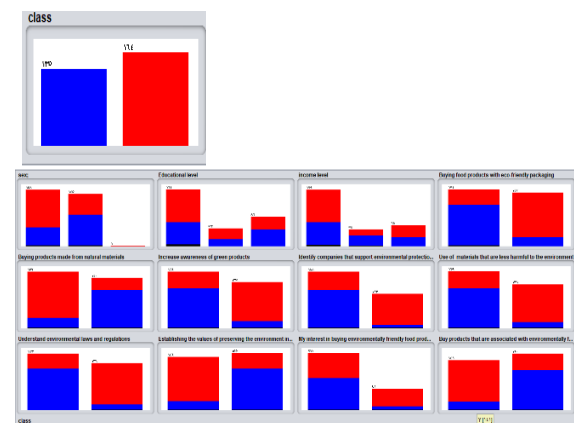


Fig. 2 Attributes Visualizer

1. J48 algorithm classifier

TABLE 2 J48 Run information
Instances: 305
Attributes: 13
sex;
Educational level
income level
Buying food products with eco-friendly packaging
Buying products made from natural materials
Increase awareness of green products
Identify companies that support environmental protection organizations
Use of materials that are less harmful to the environment
Understand environmental laws and regulations
Establishing the values of preserving the environment in my buying behavior
My interest in buying environmentally friendly food products ends with the end of the Corona pandemic
Buy products that are associated with environmentally friendly brands and labels that do not cause pollution
class
Test mode: split 66.0% train, remainder test

TABLE 3 J48 Algorithm pruned tree	
Understand environmental laws and regulations = N Buying products made from natural materials = Y Increase awareness of green products = N Identify companies that support environmental protection organizations = N Buying food products with eco-friendly packaging = N: N (18.34/5.0) Buying food products with eco-friendly packaging = Y: Y (10.0/3.0) Identify companies that support environmental protection organizations = Y: Y (6.0/1.0) Increase awareness of green products = Y Educational level = university Graduate: Y (14.55/1.0) Educational level = Postgraduate : N (2.0) Educational level = Undergraduate : Y (4.89/1.0) Buying products made from natural materials = N: N (108.32/9.66) Understand environmental laws and regulations = Y Use of materials that are less harmful to the environment = N sex; = F Buying food products with eco-friendly packaging = N income level = Middle Increase awareness of green products = N: N (3.0) Increase awareness of green products = Y: Y (2.0) income level = poor: Y (0.0) income level = high: Y (2.0) a) Buying food products with eco-friendly packaging = Y: Y (13.27/1.0) sex; = M Identify companies that support environmental protection organizations = N: N (8.0/1.0) Identify companies that support environmental protection organizations = Y Buying products made from natural materials = Y: Y (4.0) Buying products made from natural materials = N: N (4.0/1.0) sex; = DÑ: Y (0.0) Use of materials that are less harmful to the environment = Y: Y (98.63/1.0) Number of Leaves : 17 Size of the tree : 30	

TABLE 4 J48 algorithm Performance parameters		
S. No.	Parameters	values
1	Correctly Classified Instances	81.3725%
2	Incorrectly Classified Instances	18.6275%
3	Kappa statistic	62.69%
4	Mean absolute error	23.95%
5	Root mean squared error	41.41%
6	Relative absolute error	48.273%
7	Root relative squared error	82.274%
8	TP Rate Weighted Avg.	81.4%
9	Precision Weighted Avg.	81.5%
10	Recall Weighted Avg.	81.4%
11	Time taken to build model	0.05 Sec

12	Time taken to test model on test split (in Sec)	0.03 ec
----	---	---------

i. Confusion Matrix
 Confusion matrix is a tabular chart that allows visualization of the performance of the algorithm according to the calculation of test data[9]. table () shows the Confusion matrix for J48 algorithm

TABLE 5 J48 algorithm Confusion matrix		
a	b	<-- classified as
39	11	a=N
8	44	b=Y

1. LMT algorithm

TABLE 6 LMT algorithm Run information	
Instances:	305
Attributes:	13
	sex;
	Educational level
	income level
	Buying food products with eco-friendly packaging
	Buying products made from natural materials
	Increase awareness of green products
	Identify companies that support environmental protection organizations
	Use of materials that are less harmful to the environment
	Understand environmental laws and regulations
	Establishing the values of preserving the environment in my buying behavior
	My interest in buying environmentally friendly food products ends with the end of the Corona pandemic
	Buy products that are associated with environmentally friendly brands and labels that do not cause pollution
	class

i. Logistic model tree

TABLE 7 LMT algorithm Logistic model	
Understand environmental laws and regulations = N	
Buying products made from natural materials = Y	
Increase awareness of green products = N	
My interest in buying environmentally friendly food products ends with the end of the Corona pandemic = N	
Identify companies that support environmental protection organizations = N	
Use of materials that are less harmful to the environment = N	
Buying food products with eco-friendly packaging = N: LM_1:22/176 (16)	
Buying food products with eco-friendly packaging = Y: LM_2:22/176 (9)	
Use of materials that are less harmful to the environment = Y: LM_3:0/132 (4)	
Identify companies that support environmental protection organizations = Y: LM_4:0/110 (4)	
My interest in buying environmentally friendly food products ends with the end of the Corona pandemic = Y: LM_5:0/88 (2)	
Increase awareness of green products = Y: LM_6:22/88 (23)	
Buying products made from natural materials = N	

		Identify companies that support environmental protection organizations = N: LM_7:22/88 (102)
		Identify companies that support environmental protection organizations = Y: LM_8:22/88 (5)
		Understand environmental laws and regulations = Y
		Use of materials that are less harmful to the environment = N: LM_9:22/66 (37)
		Use of materials that are less harmful to the environment = Y: LM_10:22/66 (97)
Number of Leaves :		10
Size of the Tree :		19

TABLE 8 LMT algorithm Performance parameters

S. No.	Parameters	values
1	Correctly Classified Instances	78.4314%
2	Incorrectly Classified Instances	21.56.86%
3	Kappa statistic	56.78%
4	Mean absolute error	21.65%
5	Root mean squared error	36.54%
6	Relative absolute error	43.4062%
7	Root relative squared error	72.6113%
8	TP Rate Weighted Avg.	78.4%
9	Precision Weighted Avg.	78.6%
10	Recall Weighted Avg.	78.4%
11	Time taken to build model	0.84 Sec
12	Time taken to test model on test split (in Sec)	0.02 ec

TABLE 9 LMT algorithm Confusion matrix

a	b	<-- classified as
37	13	a=N
9	43	b=Y

6. Discussion

Table 2 and Table 3 show that the run information for weka. classifiers. trees.J48 -C 0.25 -M 2 and weka .classifiers .trees .LMT -I -1 -M 15 -W 0.0 respectively, shows that the total number of instances is 305 and the number of attributes is 13 with all attributes shown, also shows that the test mode is split by 66.0% for the test training and 34 % for the remainder test. Table 3 shows J48 Algorithm pruned tree Which shows that the effective attribute is Understand environmental laws and regulations, which represents the root of the tree, followed in importance by attribute that shows Use of materials that are less harmful to the environment. Table 7 shows LMT Logistic model tree which emphasizes that the effective attribute is Understand environmental laws and regulations, which represents the root of the tree, followed in importance by attribute that shows Use of materials that are less harmful to the environment. This indicates that the most important factors affecting the purchase of green products (environmentally friendly products) during the

quarantine of the Corona virus that were applied in the Kingdom of Saudi Arabia during the year 2020 are Understand environmental laws and regulations.

Table 10 J48 classification algorithm Vs LMTclassification Algorithm

S. No.	Parameters	J48 algorithm	LMTAlgorithm
1	Correctly Classified Instances	81.372 5%	78.4314 %
2	Incorrectly Classified Instances	18.627 5%	21.56.86 %
3	Kappa statistic	62.69 %	56.78%
4	Mean absolute error	23.95 %	21.65%
5	Root mean squared error	41.41 %	36.54%
6	Relative absolute error	48.273 %	43.4062 %
7	Root relative squared error	82.274 %	72.6113 %
8	TP Rate Weighted Avg.	81.4%	78.4%
9	Precision Weighted Avg.	81.5%	78.6%
10	Recall Weighted Avg.	81.4%	78.4%
11	Time taken to build model	0.05 Sec	0.84 Sec
12	Time taken to test model on test split	0.04 Sec	0.03 Sec

Table 10 shows the Performance parameters and their values for J48 classification algorithm Vs LMTAlgorithm, it was found that Algorithm J48 is preferable than LMTAlgorithm in model building because it has higher Correctly Classified Instances (81.3725%), TP Rate Weighted Avg. (81.4%) and less Time taken to build model (0.05 Sec) compared to LMTAlgorithm which has 78.4314% Correctly Classified Instances, 78.4% TP Rate Weighted Avg. and 0.84 Sec Time taken to build model.

Table 5 and Table 9 show the confusion matrix for each of the J48 algorithm and LMT algorithm respectively, as there is no significant difference between them.

7. Conclusion and Recommendation

The contribution of this paper comes to knowing the impact of the COVID 19 quarantine that was enforcement in Kingdom of Saudi Arabia during the year 2020 on the behavior of consumers in purchasing green food products and the extent of its impact on the environment. A predictive model was built to know the impact of quarantine on the purchase of green products using machine learning algorithms, where the J48 and LMT algorithms were used through the Weka application. Experimental result shows that Understand environmental laws and regulations is the important factor for purchasing environmentally friendly food. Algorithms used performance were compared, and the test shows that

the J48 algorithm has best performance compared to LMT algorithm. The study recommends spreading and intensifying awareness of environmental laws and regulations, in order to preserve the environment.

8. References

- [1]. Kantatasiri, P., P. Jaroenwanit, and R. Brown, *The influencing of young consumers shopping style on attitude toward the environmentally friendly food products in Thailand*. International Business Management, 2015. **9**(1): p. 105-110.
- [2]. Arnaudova, M., T.A. Brunner, and F. Götze, *Examination of students' willingness to change behaviour regarding meat consumption*. Meat Science, 2022. **184**: p. 108695.
- [3]. Granheim, S.I., et al., *Mapping the digital food environment: A systematic scoping review*. Obesity Reviews, 2022. **23**(1): p. e13356.
- [4]. Norton, G.W., J. Alwang, and W.A. Masters, *Economics of agricultural development: World food systems and resource use*. 2021: Routledge.
- [5]. Anastasiadou, E., et al., *The coronavirus' effects on consumer behavior and supermarket activities: insights from Greece and Sweden*. International Journal of Sociology and Social Policy, 2020.
- [6]. Ketelsen, M., M. Janssen, and U. Hamm, *Consumers' response to environmentally-friendly food packaging-A systematic review*. Journal of Cleaner Production, 2020. **254**: p. 120123.
- [7]. Indriani, I.A.D., M. Rahayu, and D. Hadiwidjojo, *The influence of environmental knowledge on green purchase intention the role of attitude as mediating variable*. International Journal of Multicultural and Multireligious Understanding, 2019. **6**(2): p. 627-635.
- [8]. Wang, Y., et al., *How impacting factors affect Chinese green purchasing behavior based on Fuzzy Cognitive Maps*. Journal of Cleaner Production, 2019. **240**: p. 118199.
- [9]. Martínez, A., et al., *A machine learning framework for customer purchase prediction in the non-contractual setting*. European Journal of Operational Research, 2020. **281**(3): p. 588-596.
- [10]. Awadh, W.A., A.S. Alasady, and H.I. Mustafa. *Predictions of COVID-19 Spread by Using Supervised Data Mining Techniques*. in *Journal of Physics: Conference Series*. 2021. IOP Publishing.
- [11]. Kotthoff, L., et al., *Auto-WEKA: Automatic model selection and hyperparameter optimization in WEKA*, in *Automated Machine Learning*. 2019, Springer, Cham. p. 81-95.
- [12]. El Bouchefry, K. and R.S. de Souza, *Learning in big data: Introduction to machine learning*, in *Knowledge Discovery in Big Data from Astronomy and Earth Observation*. 2020, Elsevier. p. 225-249.
- [13]. Sun, W., et al., *Data processing and text mining technologies on electronic medical records: a review*. Journal of healthcare engineering, 2018. **2018**.
- [14]. Strilets, V., et al., *Application of artificial neural networks in the problems of the patient's condition diagnosis in medical monitoring systems*, in *Integrated Computer*

Technologies in Mechanical Engineering. 2020, Springer. p. 173-185.

- [15] Dr. Khalid Nazim Abdul Sattar et. al, *Developing a Prediction Model Using J48 Algorithm to Predict Symptoms of COVID-19 Causing Death*, , IJCSNS, International Journal of Computer Science and Network Security, Vol. 20 No. 8 pp. 80-83, August 2020.

Authors Profile



Dr. Mutasim Mohamed Al Sadig, Assistant Professor, Department of CSI, College of Science, Majmaah University, Az Zulfi, Kingdom of Saudi Arabia, received his Bs in Science and Technology & M. Ec(CS) from GAZERA University, Sudan. PhD in Information Technology, from Academic Sudan University. He has 20 + years of research and teaching experience and his research expertise encompasses Data Mining, Biometric Identification System , Software Engineering, Artificial Intelligence, and Computer Networks.



Dr. Faisal Mohammed Nafie is an Assistant Professor, at Department of Computer Science, College of Science and Humanities at Alghat, Majmaah University, Majmaah 11952, Saudi Arabia. He has doctorate in Computer Science. He has 18+ years of teaching experience. His research interests include Data Mining, Process

Mining, Workflow Mining, Computer Networks and telecommunication, Information Security, Bioinformatics, Software Engineering, IoT and Smart Cities, and e- learning in higher education institutions.



Dr. Khalid Nazim Abdul Sattar, Assistant Professor, Department of CSI, College of Science, Majmaah University, Az Zulfi, Kingdom of Saudi Arabia, received his B. E. Degree from Bangalore University, M. Tech from VTU Belagavi, and Ph.D. in Computer Science & Engineering from Singhania University, Rajasthan, India. He has 07 years of research and 14

years of teaching experience and his research expertise include but are not limited to Image processing, digital signal processing, Artificial Intelligence, Internet of Things, Cryptography, Data & Network Security and Data mining.