Heliyon 9 (2023) e18320

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

# Assessing aflatoxin safety awareness among grain and cereal sellers in greater Accra region of Ghana: A machine learning approach

Vincent Owusu Kyei-Baffour [a,b,*], Hilary Kwesi Ketemepi [a], Nancy Nelly Brew-Sam [a], Ebenezer Asiamah [a], Leonora Charlotte Baffour Gyasi [a], Wisdom Kofi Amoa-Awua [b]

[a] *Council for Scientific and Industrial Research – Food Research Institute (CSIR-FRI), P.O Box M20, Accra, Ghana*
[b] *Department of Agro-Processing Technology and Food Bio-Sciences, CSIR College of Science and Technology, Josip Broz Tito Avenue, Accra, Ghana*

A R T I C L E I N F O

*Keywords:*
Aflatoxin awareness
Grains
Cereals
Market sellers
Machine learning
CART

A B S T R A C T

Studies have established high prevalence of aflatoxin contamination in grains and cereals produced in Ghana. Mitigation strategies have focused mainly on capacity building for farmers, agricultural extension officers, bulk distributors and processors to the detriment of the market women who act as the final link between consumers and producers. This study used supervised machine learning algorithms by means of Classification and Regression Trees (CART) to investigate aflatoxin knowledge and awareness of market women in Greater Accra Region of Ghana. A cross-sectional survey and probability sampling methods were employed for data collection. Ninety-two (92%) of participants had never heard about aflatoxins and yet, 62% reported that they usually observe mould growth in their cereals/grains. Unsurprisingly, 97% of participants indicated that they had no knowledge of the aflatoxin bill passed by the government of Ghana parliament. Despite participants not being aware of aflatoxin menace, the percent correctness of their aflatoxin safety measure score was 40%. A regression tree algorithm showed that, participant's ethnic group was the most significant parameter to consider regarding their aflatoxin safety knowledge. Their educational background and age were 95.5% and 72.5% as significant as their ethnic group. A classification tree algorithm showed that, educational level was the most significant parameter to consider when it comes to sorting of grains/cereals. Their ethnic group and marital status were 92.4% and 89.3% as important as educational level. It is therefore imperative for the Ghana government to extend sensitization and awareness programs to these market women, targeting the uneducated and specific age and ethnic groups.

## 1. Introduction

Aflatoxins are fungal secondary metabolites produced primarily by toxigenic strains of the fungi *Aspergillus flavus* and *Aspergillus parasiticus*. Aflatoxin-producing fungi are found in areas with a hot, humid climate and their presence in food are as a result of both pre- and post-harvest fungal contamination [1]. Aflatoxins B1 (AFB1), B2 (AFB2), G1 (AFG1) and G2 (AFG2) are the four (4) main types of aflatoxins frequently found in contaminated food and AFB1 is the most virulent of the four [2]. They have been associated with liver

* Corresponding author. Council for Scientific and Industrial Research – Food Research Institute (CSIR-FRI), Ghana.
*E-mail address:* vok_baffour@yahoo.com (V.O. Kyei-Baffour).

https://doi.org/10.1016/j.heliyon.2023.e18320
Received 15 June 2023; Received in revised form 9 July 2023; Accepted 13 July 2023
Available online 17 July 2023
2405-8440/

cancer and have been classified as class 1 carcinogen, with peanut, maize and their derivatives being the main vehicles [3].

In Ghana, studies have established high prevalence of aflatoxin contamination in grains and cereals. Ref. [4] analyzed 180 maize samples from different agro-ecological zones in Ghana. They concluded that, 131 of the maize samples were contaminated with aflatoxins with 127 (70.50%) and 116 (64.44%) exceeding the EFSA and Ghana Standard Authority (GSA) limits respectively. Another study by Baah-Tuahene [5] showed that groundnut oil and its by-products (*kulikuli* and khebab powder) were highly contaminated with aflatoxins and the levels far exceeded the set limit of 4 μg/kg total aflatoxin by the European Union and the adopted limit of 15 μg/kg by Codex in peanuts for further processing. Ref [6,7,8], have all associated high levels of aflatoxin with grains and cereals in Ghana.

There have been efforts to mitigate against the occurrence of aflatoxins in maize, other cereals and groundnuts, since the study by Kpodo, Sørensen and Jakobsen [9] on the occurrence of aflatoxins in kenkey, a traditional fermented maize product became a national issue. These measures have targeted both pre- and post-harvest stages of the food production chain. In some countries crop seeds have been genetically modified to be stress tolerant and more resistant to mould infestation and subsequent aflatoxin production [2]. Biological control measures have been developed where moulds incapable of producing aflatoxins are allowed to colonize crops to prevent the aflatoxin producing strains from infesting the crops [10]. Good agricultural practices such as crop rotation, moisture management and timely harvesting of crops have also been used as mitigation measures [2]. Some post-harvest strategies including quick drying of crops after harvest, cleaning of crops, and proper storage have also been adopted. Other interventions have focused on diet diversification of consumers and post exposure management including enterosorption, where products that bind aflatoxins in the gut and prevent their uptake have been introduced [11].

Majority of these interventions however, have focused mainly on capacity building for the farmer, agricultural extension officers, bulk distributors and processors to the detriment of the market women who act as the final link between consumers and farmers/processors of food crops. The level of aflatoxin safety knowledge and awareness of market women especially on post-harvest handling of grains and cereals before final consumer purchase, will be very vital in the entire mitigation strategies. To appreciate the broad picture and effectively curtail the situation, the safety knowledge and awareness of the situation by market women should also be investigated to identify gaps to which appropriate strategies can be recommended.

Machine learning is one of the current rapidly growing technical fields and in recent years, has been widely used in various fields including epidemiology, nutrition and food safety [12]. It is a powerful statistical method for collecting, summarizing, and analyzing data from different perspectives into valuable and practical information to identify useful relationships [13]. As a representative machine learning method, the Classification and Regression Tree (CART) has considerable advantages compared to traditional statistical modeling methods. It can achieve more accurate results, handle larger and more complex data. Machine learning methods have therefore become popular methods to solve problems of food safety [14]. It is a nonparametric method assuming no predefined data probability distributions or variable relationships, making it appropriate for solving complex, dynamic problems [15]. Additionally, machine learning models are able to handle missing attribute values and outliers which may ruin a model [16]. Machine learning therefore holds potential in leveraging large, emerging data sets to improve the safety of food supply and mitigate the impact of food safety incidents.

Ref [17] used machine learning (ML) models, which included weather-based mechanistic model predictions for aflatoxin occurrence in maize. Work done by Yoo et al. [18] used CART models to study the potential hazards of urban airborne bacteria during Asian dust events. Ref [19] employed CART models to study the socio-economic and lifestyle parameters associated with diet quality of children and adolescents. Machine learning models have again been used to analyze poultry data to improve food safety and production efficiencies [20]. The aim of this study was to use machine learning algorithms by means of CART to investigate the safety knowledge and awareness of fungal secondary metabolites (aflatoxins) contamination in grains and cereals of market women in Greater Accra Region of Ghana and to know the principal factors that influence their level of knowledge and awareness.

## 2. Materials and methods

### 2.1. Study settings

The current study was conducted among market sellers of grains and cereals within selected markets in the Greater Accra Region of Ghana. The study took place from February to July 2022. Eighteen (18) markets in the Greater Accra Region considered as the "hub" for commercial grain and cereal activities were identified. Ten (10) markets were systematically and randomly selected from the 18 major markets in the region.

### 2.2. Study design, participants and sampling

A cross-sectional survey and probability sampling methods were employed for data collection. Ten (10) markets in the Greater Accra region of Ghana were systematically and randomly selected. A sampling interval was calculated to generate a systematic order for sampling the 10 markets.

Sampling Interval = total markets/markets to be selected.

All markets were numbered from 1 to 18 and a number between the sampling interval randomly selected. Markets were counted by their order until the selected number was reached and was chosen as the first market for data collection. The sampling interval was added to the selected market's number to get the next systematic order for which the second market was selected. This process was repeated until all 10 markets were selected from the 18 major markets.

Cochran's formula was employed to calculate the sample size of respondents (sellers of grains and cereals) required for the survey.

$$N = \frac{\boldsymbol{P}[\mathbf{1} - \boldsymbol{p}]z^2}{\boldsymbol{e}^2}$$

Where;

N = Sample size,

P = Proportion of the population of market women and men with knowledge on aflatoxins,

e = The error margin for desired precision

z = The standard score (z value) for a normal distribution at a specific confidence level.

A non-responsive rate was used to inflate the sample size to accommodate for an unforeseen non-responsiveness. A non-responsive rate was calculated using the expression;

Final N = [Calculated N]/[1 – NRR] where NRR is the none responsive rate.

There was no literature estimating the proportion of market sellers with knowledge of aflatoxins. A 50% proportion was therefore assumed. At a 5% error margin for desired precision, four hundred and four (404) market sellers were surveyed. Out of this total, forty (40) sellers of grains and cereals were interviewed from each of the 10 markets.

### 2.3. *Data collection tool and procedure*

The questionnaire used in this study consisted of several parts including: socio-demographic information, post-harvest practices, aflatoxin safety knowledge and awareness and, aflatoxin regulation. The socio-demographic information included gender, marital status, age, ethnic group, religious background and educational level of participants. A set of eight (8) questions with closed end options to choose from was used to assess participants post-harvest practices regarding grains and cereals. The post-harvest section included information on duration of storage, storage conditions, mode of storage and sorting of grains/cereals. Information on aflatoxin safety knowledge and awareness included thirteen (13) questions centering on visible mould contamination, how they are handled, aflatoxin susceptible grains and periodic laboratory testing of grains/cereals. The aflatoxin safety knowledge of grain/cereals was scored based on the thirteen (13) questions asked. A participant scored one (1) mark if he/she answered the question correctly, otherwise scored a zero (0). An aggregate top score of 10 was expected. Research staff visited different locations of the 10 selected major markets in the Greater Accra region of Ghana where grains and cereals are sold. Research staff randomly approached sellers 20 years or older (visual estimation) who sold grains and cereals in bulk quantities and invited them to participate in the study. Research staff reviewed the objectives of the study to all participants and asked those willing to participate to sign a consent form. Data were collected by face-to-face interview using a structured questionnaire.

### 2.4. *Validity and reliability of questionnaire*

The English-version of questionnaire used was translated into the local language (*Twi*), one of the predominant Ghanaian languages during data collection. Translation of the questionnaire from English to *Twi* was executed by a bilingual translator, which was checked by an independent research scientist. Back-translation of the questionnaire was also done by a separate independent bilingual translator to check for consistencies and to avoid any bias in the questionnaire. Prior to questionnaire administration, it was pre-tested and the internal consistency of components were assessed. A Cronbach's alpha of 0.8142 was obtained ($n = 20$) indicating an acceptable reliability of questionnaire used in this study.

### 2.5. *CART analytical models*

The relationships between aflatoxin safety knowledge, sorting behaviour of grain and cereal sellers, and their socio-demographic parameters was studied using classification and regression tree algorithms as described by Zacharis [16]. For the given predictors, the response data set was split into two parts using homogeneity of data as criterion. In order to decide which attribute to split, the Gini impurity measure was used.

The predictors were treated as either continuous or categorical variables and based on this, an appropriate splitting was done. For a continuous predictor variable 'X' and a value "c", a split was defined by sending all records with the values of 'X' less than or equal to "c" to the left branch node, and all remaining records to the right branch node. The average of two adjacent values was then used to compute 'c'. A continuous variable with N distinct values would generate up to N–1 potential splits of the root node. For a categorical predictor variable 'X' with distinct values ($c_1, c_2, \ldots, c_k$), a split was defined as a subset of levels that were sent to the left branch node. A categorical variable with K levels would therefore generate up to $2^{K-1} - 1$ split.

#### 2.5.1. *Pruning of CART*

To prevent data overfitting and one data in each leaf node, trees were pruned to enhance predictive power of the classification. Since the data set did not exceed 5,000, a v-fold cross validation of the data was done and included independent test sets.

All the data set were used to fit an initial overly large tree. The data was then divided into $v = 10$ subgroups, and 10 separate models fitted. The first model used subgroups 1–9 for training, and subgroup 10 for testing. The second model used groups 1–8 and 10 for training, and group 9 for testing. In all cases, an independent test subgroup was available. These 10 test subgroups were then combined to give independent error rates for the initial overly large tree which was fitted using all data set.

#### 2.5.2. CART model adequacy

The most accurate classification tree was selected based on misclassification cost for classification trees and R-squared for regression trees. The lowest misclassification cost within one standard error was selected and receiver operating characteristic (ROC) estimated. This was done by plotting the true positive rates (power of model) against false positive rates (type I error). The performance of training data set was also compared to that of test data to assess overall model adequacy.

### 2.6. Statistical analysis

Descriptive statistics (response frequencies, median, percentages and graphs) were used to summarize variables of interest. A Chi square goodness of fit test (N outcomes) was employed to test for significance among parameters assessed. CART algorithms were employed to assess the relationship between aflatoxin safety knowledge scores, grain/cereal sorting behaviour and socio-demographic variable of participants. Misclassification cost was used to select most appropriate tree. Significance was accepted at 5% type I error rate for responses. All data were analyzed using Minitab statistical software version 21.

## 3. Results and discussions

### 3.1. Socio-demographic information

The socio-demographic information is summarized in Table 1. There were 340 females in the study accounting for 89% of total participants. Majority were married and accounted for 71% of total participants. This proportion was significantly higher ($p < 0.05$) than those participants who were single, divorced or widowed. The majority of participants were in the age range of 30–39 years. Mole Dagbani ethnic group dominated the study with a proportion of 47% which showed significant difference ($p < 0.05$) from other ethnic groups. Muslims participants (82%) were more than Christian participants and the difference was significant ($p < 0.05$). Middle school/Junior High School leavers dominated the study with a proportion of 31%.

Gender segregation in the labour market continues to exist in a growing economy like Ghana [21]. The proportion of females (89%) to males (11%) confirms the assertion that certain sectors of the market are assigned to specific genders. Northern Ghana is the main food basket in terms of grain cereals for Ghana and the region account for ca. 97% of sorghum and millet production in the country [22]. It was therefore not surprising to have Muslims and Mole Dagbani, a strong and influential Muslim community native to the northern parts of Ghana [23], dominating the study.

### 3.2. Postharvest handling and practices

Questions were asked to assess participants post-harvest handling of grains and cereals. The result is presented in Table 2. Forty two percent (42%) of respondents gave pest infestation as a significant cause of losses during post-harvest handling of cereals and grains. Majority (42%) of the participants stored their grains for a period of 1–3 months before sales, whereas 32% stored under a month. It was interesting to note that, 76% of participants stored their grains in sacks and on wooden pellets while 3% poured grains on bare

**Table 1**
Socio-demographic parameters of grains/cereal sellers from selected markets in the Greater Accra region.

| Parameters | Levels | Counts | Proportion | Chi-square statistic ($\chi^2$) | p-value ($\alpha = 0.05$) |
|---|---|---|---|---|---|
| Gender | Female | 340 | 0.89 | 273 | <0.001 |
| | Male | 40 | 0.11 | | |
| Marital Status | Single | 50 | 0.13 | 441 | <0.001 |
| | Married | 270 | 0.71 | | |
| | Divorced | 50 | 0.13 | | |
| | Widowed | 10 | 0.02 | | |
| Age | 20–29 | 90 | 0.23 | 13.7 | 0.003 |
| | 30–39 | 120 | 0.31 | | |
| | 40–49 | 70 | 0.18 | | |
| | >50 | 100 | 0.26 | | |
| Ethnic groups | Akan | 10 | 0.02 | 419 | <0.001 |
| | Ga Dangme | 20 | 0.05 | | |
| | Ewe | 70 | 0.18 | | |
| | Guan | 20 | 0.05 | | |
| | Mole Dagbani | 180 | 0.47 | | |
| | Other | 80 | 0.21 | | |
| Religion | Muslim | 310 | 0.82 | 152 | <0.001 |
| | Christian | 70 | 0.18 | | |
| Education | None | 100 | 0.26 | 56.8 | <0.001 |
| | Primary | 60 | 0.15 | | |
| | Middle School/JHS | 120 | 0.31 | | |
| | Secondary | 60 | 0.15 | | |
| | Tertiary | 40 | 0.10 | | |

**Table 2**
The post-harvest handling assessment of grain/cereal sellers from selected markets in the Greater Accra region of Ghana.

| Questions | Levels | Counts | Proportion | Chi-square statistic ($\chi^2$) | p-value (α = 0.05) |
|---|---|---|---|---|---|
| What causes losses of grains/cereals | Pest infestation | 160 | 0.42 | 253 | <0.001 |
| | Poor storage | 50 | 0.13 | | |
| | Bad weather | 90 | 0.24 | | |
| | Poor storage; Bad weather | 20 | 0.05 | | |
| | Other reasons | 60 | 0.16 | | |
| How long do you store grains/cereals | Less than 1 month | 120 | 0.32 | 93.7 | <0.001 |
| | 1–3 months | 160 | 0.42 | | |
| | 4–6 months | 50 | 0.13 | | |
| | Over 6 months | 50 | 0.13 | | |
| How do you store grains/cereals | Bare floor | 10 | 0.03 | 767 | <0.001 |
| | Terrazzo/Tiles | 20 | 0.05 | | |
| | Line floor with empty sacks | 10 | 0.03 | | |
| | In sacks on floor | 50 | 0.13 | | |
| | In sacks on structures | 290 | 0.76 | | |
| Do you sort grains/cereals | Yes | 360 | 0.95 | 304 | <0.001 |
| | Sometimes | 20 | 0.05 | | |
| Why do you sort | Better price | 290 | 0.76 | 540 | <0.001 |
| | Better price; Safety reasons | 10 | 0.03 | | |
| | To attract customers | 40 | 0.11 | | |
| | Convenience | 40 | 0.11 | | |
| How do you handle sorted grains/cereals | Discard | 130 | 0.34 | 289 | <0.001 |
| | Feed | 150 | 0.39 | | |
| | Feed; Discard | 30 | 0.08 | | |
| | Food for home | 20 | 0.05 | | |
| | Feed; Sold separately | 40 | 0.11 | | |
| | Sold separately | 10 | 0.03 | | |

floor during storage. Remarkably, 95% of participants indicated that they always sorted their grains before sales, whereas 5% indicated they sorted only at certain times and not always. In Tanzania, Magembe et al. [24] who carried out an assessment of awareness of mycotoxin infections in stored maize and groundnut in Kilosa District, Tanzania reported that 50% of the participants sorted out their grains before sales while the other 50% did not sort at all. In the present study the respondents explained that they sorted their cereals/grains because it gave a better price, attracted buyers and also for safety reasons. Other studies in Ghana have also reported that sorting of grains before selling maintains grain quality as well as the grade, which will eventually result in better pricing [25,26]. When participants were asked how they handled the sorted/rejected grains/cereals, 34% claimed that they discarded the bad grains, 39% said that they used it as feed for animals and 11% indicated that they sold the bad grains separately or used them for animal feed. Ref [25] emphasized that, shriveled and insect damaged peanut kernels are probable source of aflatoxin contamination and must be sorted out of healthy kernels and discarded. Contrary to this assertion, half (50%) of the participants surveyed sorted out bad grains but used it as either animal feed or sold them separately.

### 3.3. *Aflatoxins awareness and safety knowledge*

The responses to questions asked to assess participants aflatoxin awareness and safety knowledge is presented in Table 3. Surprisingly, 92% of participants had not heard about aflatoxins. However, 66% reported that they usually experience mould growth in

**Table 3**
Aflatoxin awareness and safety knowledge assessment of grain/cereal sellers from selected markets in the greater Accra region of Ghana.

| Questions | Levels | Counts | Proportion | Chi-square statistic ($\chi^2$) | p-value (α = 0.05) |
|---|---|---|---|---|---|
| Have you heard about aflatoxins | Yes | 30 | 0.08 | 269 | <0.001 |
| | No | 350 | 0.92 | | |
| Do you get mould contamination in grains/cereals | Yes | 250 | 0.66 | 37.9 | <0.001 |
| | No | 130 | 0.34 | | |
| Do you use chemical to prevent moulds | Yes | 170 | 0.45 | 4.21 | 0.04 |
| | No | 210 | 0.55 | | |
| Do regulatory authorities come to check your grains | Yes | 130 | 0.34 | 37.9 | <0.001 |
| | No | 250 | 0.66 | | |
| If yes how often do they come | Once in 6 months | 40 | 0.31 | 56.9 | <0.001 |
| | Once in 12 months | 80 | 0.62 | | |
| | Others-once in a while | 10 | 0.08 | | |
| Have you ever tested your products for aflatoxins | No | 380 | 1.00 | N/A | N/A |
| | yes | 0 | 0.00 | | |
| Are you aware of the current aflatoxin bill | Yes | 10 | 0.0263 | 341 | <0.001 |
| | No | 370 | 0.9737 | | |

their grains. In an attempt to control mouldy grains, 45% indicated they used chemicals to prevent mould growth. Obviously, the grain sellers had no knowledge that mouldy grains could lead to the production of carcinogenic toxins including aflatoxins. As such, they assumed that the mere treatment of mouldy grains with chemicals/fungicides rendered them safe for consumption. Agbetiameh et al. [6] reported on aflatoxin contamination in maize and groundnut in major producing regions across three agroecological zones (AEZs) in Ghana. They found that there was a high prevalence of aflatoxin contamination in maize and groundnut, even when there was no visible mould growth. Similar assertion was made by Ortega-Beltran and Bandyopadhyay [27] who stressed on the importance of preventing mould growth and subsequent aflatoxin production through proper storage and Laboratory testing. When participants in the current study were asked about regulation by authorities, 34% indicated that regulatory authorities seldomly inspected their grains and of the 34% who indicated so, 62% stated that the visit was just once in a year. Majority (66%) however indicated that regulatory authorities have never inspected their grains. It was therefore not surprising that all participants indicated they have never tested their cereals/grains for aflatoxins. Consequently, over 97% of the sellers had no knowledge of the aflatoxin bill passed by the government of Ghana parliament. The absence of regulatory authorities from the various markets to survey and monitor grains and cereals handling confirm Lawal [28] avowal that the Food Inspectorate Department of the Food and Drugs Authority, Ghana, are faced with some challenges which makes market surveillance difficult. These he claimed include inadequacy in personnel, lack of equipment and a single main food laboratory located in the Accra main office serving the entire country.

The result of aflatoxins safety knowledge scores is presented in Fig. 1. The aflatoxin knowledge scores of participants ranged from 0 to 10 with a median score of 4. The lower 25% of participants scored a zero whereas the top 25% scored a 10. The lower 50% of participants scored between 0 and 4 whereas the top 50% scored between 4 and 10. The interquartile range was the same as the minimum (0) and maximum (10) scores since the lower 25% scored 0 and the top 25% scored 10. The median score of 4 represented a 40% aflatoxin safety knowledge of grain/cereal sellers. The below average safety knowledge score on aflatoxins confirms what Ortega and Tschirley [29] concluded in their work. The authors found that the overall awareness of food safety issues in Sub Saharan Africa is low relative to Asia. Moreover, knowledge of producer behavior and consumer demand for food safety in developing countries is very limited.

### 3.4. Regression tree algorithm of aflatoxin safety knowledge scores

To further understand how socio-demographics of grain/cereals sellers are associated with aflatoxin safety knowledge, a 16 terminal node regression tree (Fig. 2) was built to study associations and for future predictions. The tree algorithms used 262 data set as training data and 118 as test data. An initial root node (node 1) showed that aflatoxin safety knowledge prediction for sellers will range from 0.8 to 9.4 where a score of 10 indicates a maximum (100%) aflatoxin safety knowledge. The 16 terminal node tree had an R squared of 83.54% (Fig. 3) which is an indication of the predictive power of the tree. This meant that the tree algorithm was adequate and that almost 84% of total variations in aflatoxin safety knowledge score was explained by the regressors.

From the root node (node 1), the ethnic group of grain/cereal sellers was the predictor with the least Gini impurity (the criteria employed for node splitting) and was therefore used to split the root node. When sellers are of the Akan or Guan ethnic group, their level of education, gender, marital status and age (in decreasing order of importance), will determine their aflatoxin safety knowledge scores. An Akan or Guan who has no formal education or has attained only primary education will have aflatoxin safety knowledge score between 0% and 27% (terminal node 1). Comparably, a female of the Akan or Guan ethnic group with tertiary education will have higher safety knowledge scores than a male (terminal node 2 and node 4). If these females are married and fall in the age bracket of 50 years and above, their aflatoxin safety knowledge score will be a 100% (terminal node 5) and between 39% and 59% if they are below 50 years (terminal node 4).

On the right branch node from the root node shows the relationship between socio-demographics and aflatoxin safety knowledge score when sellers are of the Ewe, Ga Dangme or Mole Dagbani ethnic group. Considering nodes where age was used for splitting, younger sellers will always score less on aflatoxin safety knowledge than older sellers (terminal nodes 6, 7 and 8) except for female Mole Dagbanis with no formal education. Younger sellers (20–29 years) will score more on aflatoxin safety knowledge (71.4%) than older sellers (8.6%) who are 30 years and above (terminal nodes 13 and 14). When a relative variable importance chart was considered

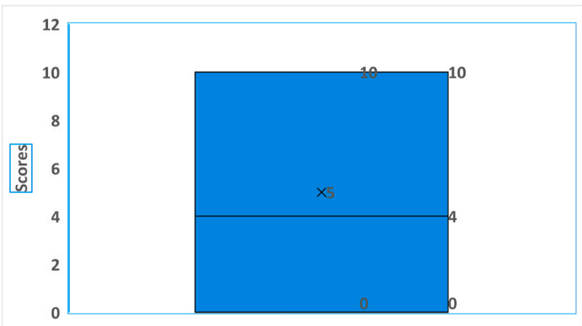


**Fig. 1.** A box and whisker plot showing the aflatoxin safety knowledge scores distribution among grain/cereal sellers from selected markets in the Greater Accra Region of Ghana (N = 380).

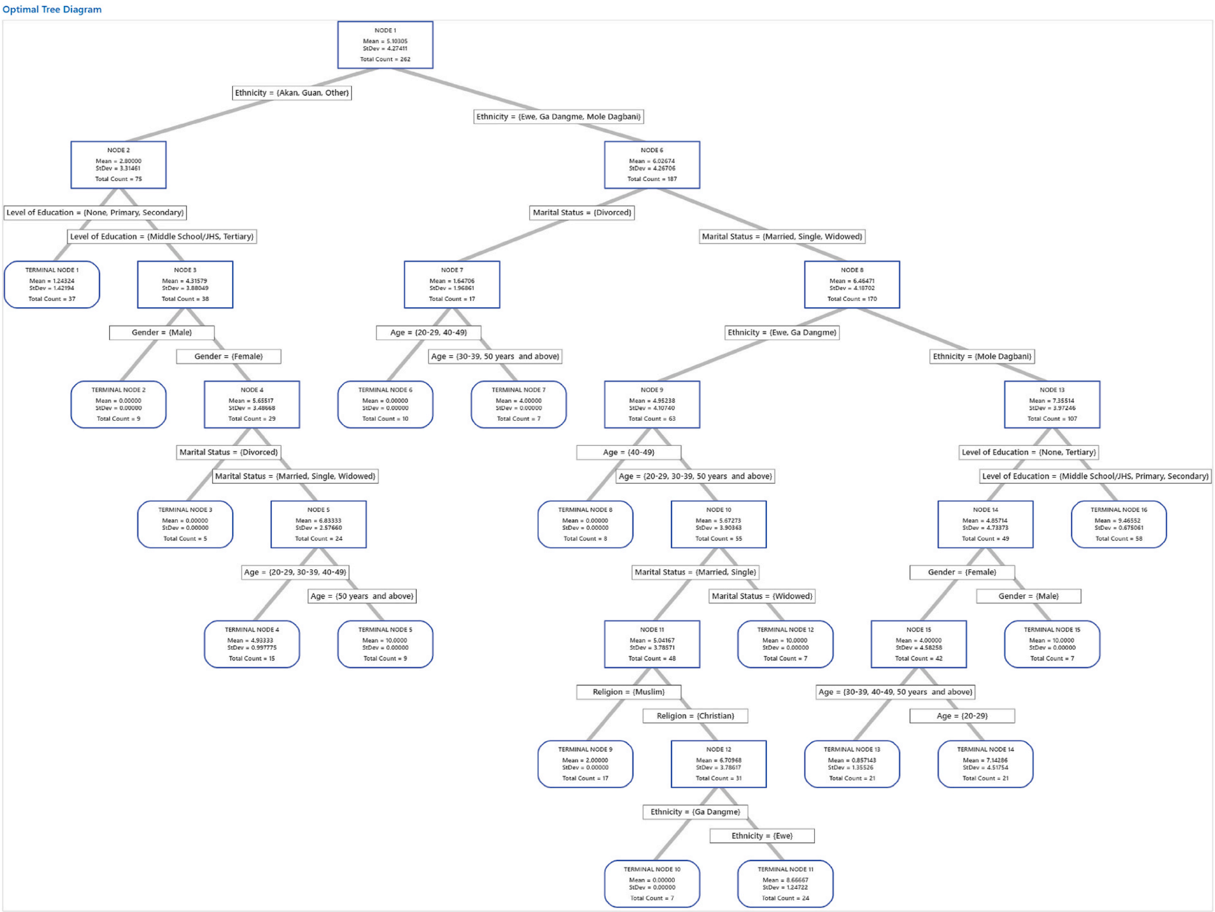


**Fig. 2.** A 16 terminal node regression tree algorithm built using aflatoxin safety knowledge scores as a discrete response and socio-demographic parameters as categorical predictors (N = 262).

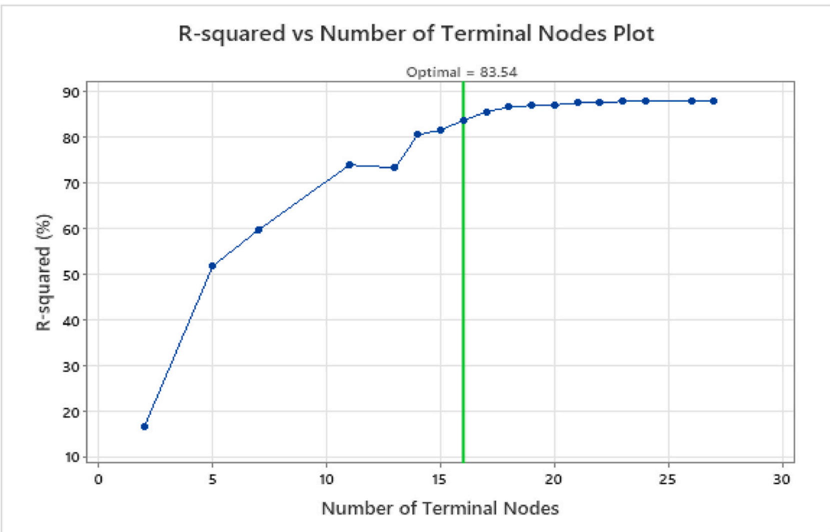


**Fig. 3.** The R-squared versus number of terminal nodes graph used to select the optimal tree that maximizes the R-squared within 1 standard error.

for the 16 terminal node tree (Fig. 4), the ethnic group of sellers emerged as the most important parameter to consider during mitigation strategies. Seller's level of education, age, marital status and gender were equally significant parameters to consider in decreasing order of importance (Fig. 4).

Kamano and his colleagues sought to investigate the influence of knowledge, attitude and practices of farmers on aflatoxin contamination of maize in Makueni and Baringo counties in Kenya [30]. They used socio-demographics as linear regressors of aflatoxin knowledge. They established that age was a significant predictor of aflatoxin knowledge scores and that younger age groups were more knowledgeable than older age groups. However, the present study carried out in Ghana predicts that older sellers are likely to be more knowledgeable in aflatoxin safety than younger sellers. It should also be noted that whereas the study in Kenya focused on farmers, the present study focused on grain and cereals sellers in the market i.e. a different focus group. In Nigeria, Adekoya et al. [31] studied consumer awareness and prevalence of mycotoxin contamination in selected Nigerian fermented foods. They concluded that, 98% of consumers were unaware of mycotoxins and that educational status of consumers was a significant predictor of mycotoxin awareness. A similar trend was observed in the present study where 92% of sellers were unaware of aflatoxins and educational status of sellers was an important determinant of their aflatoxin safety knowledge score. For both studies, participants with higher level of education are more likely to have a better score for aflatoxin safety knowledge.

### 3.5. *Classification tree algorithm of cereal sorting*

To study the associations and relationships between socio-demographics and the sorting behaviour of grain/cereal sellers, a five terminal node classification tree algorithm was built (Fig. 5). For the initial root node (node1), responses from 262 respondents used in model training was considered. Out of this, 242 representing 92.4% contended that they sorted their grains. Considering socio-demographics, the ethnic group of respondents emerged as the parameter with the least Gini impurity and was used to split the root node. Sellers from the Akan, Ewe, Ga Dangme and Guan ethnic groups will always sort their grains with a 98.9% probability (terminal node 5). The sellers from the Mole Dagbani ethnic group with some tertiary education will sort their grains with a probability of 98.1% (terminal node 4). However, those with no formal education or have had up to only middle school education and are in the middle age bracket (30–39 years) are less likely to sort their grains with only 50% chance of sorting (terminal node 1). Those at the upper age bracket (40 years and above) who are married have a higher probability of sorting their grains than those who are single with probabilities of 95.3% and 61.1% respectively (terminal nodes 2 and 3).

A relative variable importance chart (Fig. 6) for sellers sorting behaviour showed that, educational level of sellers was the most important demographic factor to consider. Their ethnic groups and marital statuses were 92.4% and 89.3% as important as their educational level when sorting of cereals and grains is concerned. The 5 terminal node algorithm was used to explain sorting behaviour since it was the optimal tree with the least misclassification cost (0.26) within one standard error (Fig. 7). The Receiver Operating Characteristic (ROC) curve (Fig. 8) shows how well the 5 terminal node tree classifies the data set. This is a plot of power of prediction (sensitivity) against type I error (1-specificity). A perfect classification model will always have an area of 1 under the ROC curve. An area of 86.1% and 83.9% for training and test data sets under the ROC curve indicates the model adequately fits the data set and as a result, classification of responses was not random.

Shriveled and immature grains are known to be the more susceptible to aflatoxin contamination than healthy grains [25]. Sorting out these grains has been found to significantly reduce aflatoxin levels in groundnut and maize [25,32,33]. In effect, regular sorting of

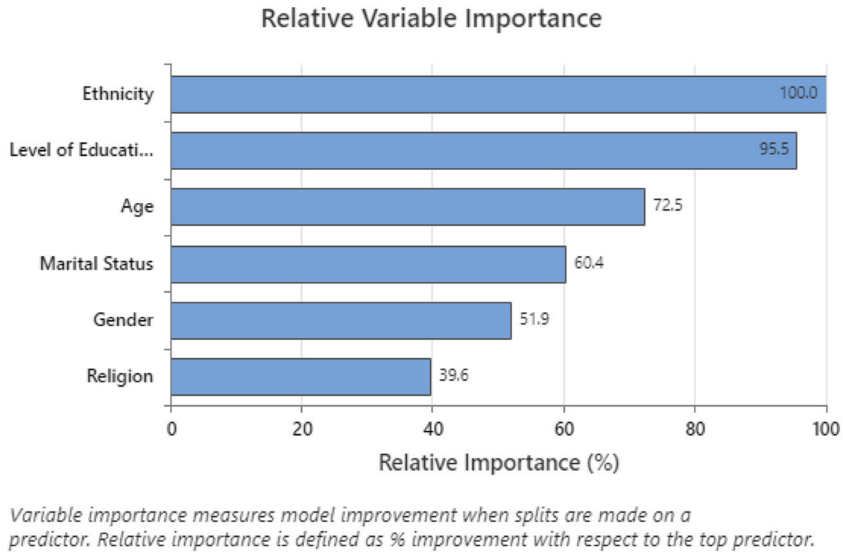


**Fig. 4.** A socio-demographic variable importance chart for the 16 terminal node regression algorithms showing the most important parameter on aflatoxin safety knowledge scores.

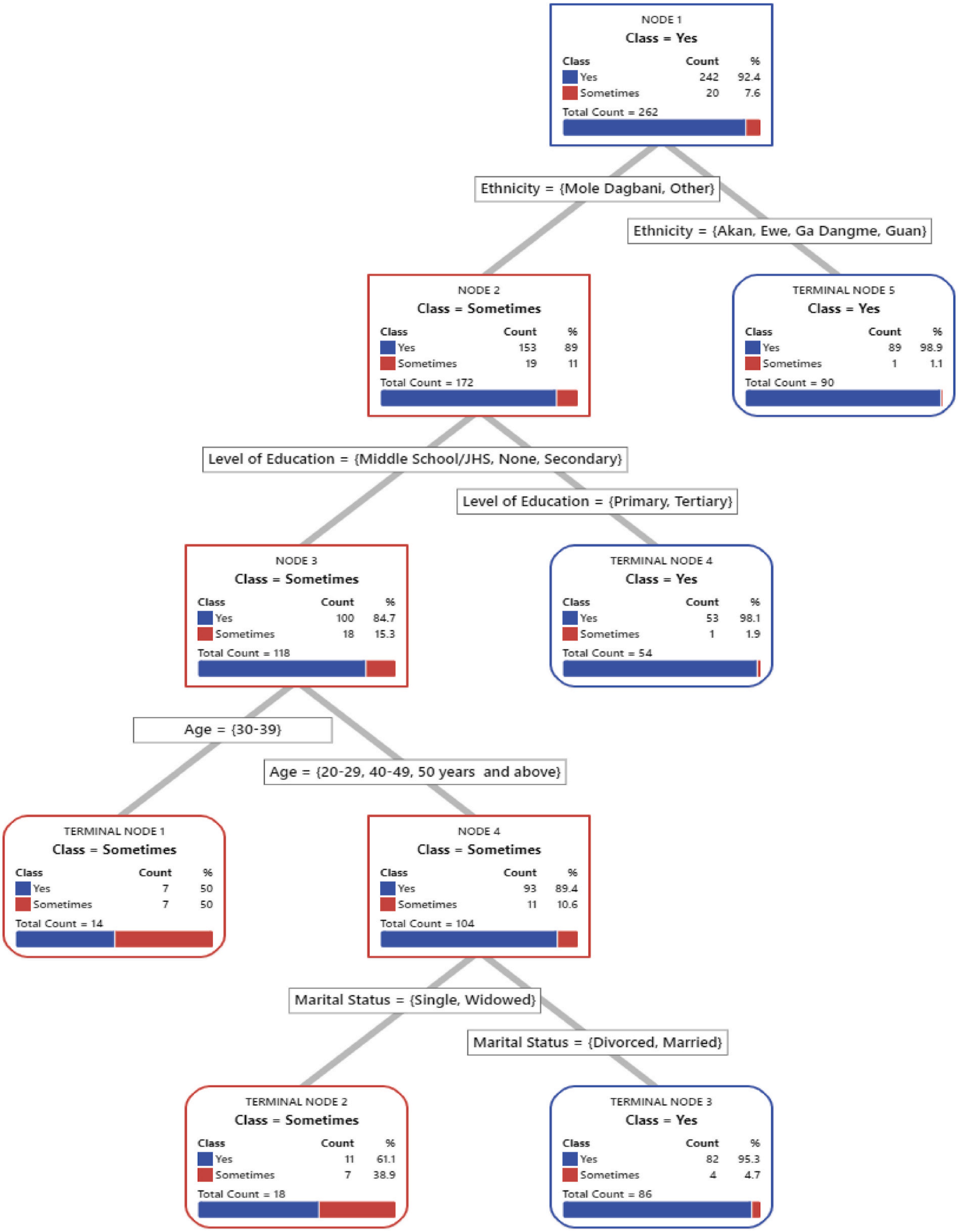


**Fig. 5.** A five terminal node classification tree algorithm built using the sorting behaviour of grain/cereal sellers and their socio-demographic variables.

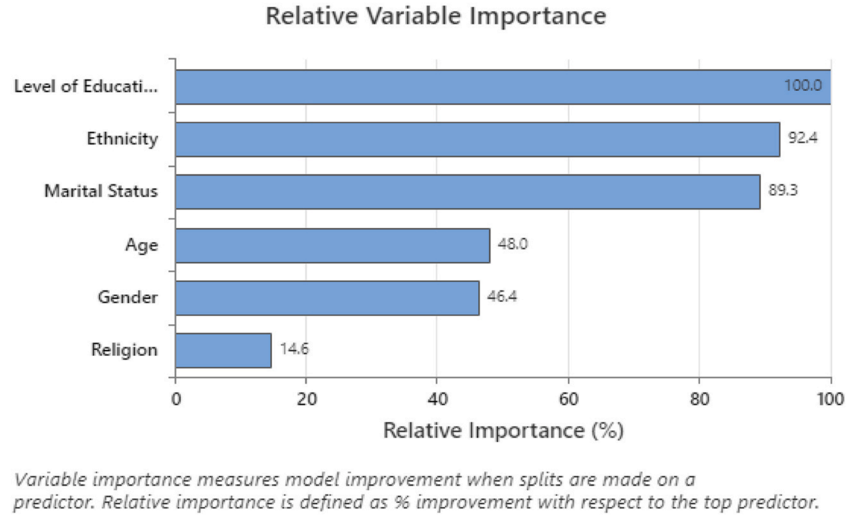


**Fig. 6.** A socio-demographic variable importance chart for the 5 terminal node classification tree algorithm showing the most important parameter for grain/cereal sorting behaviour.

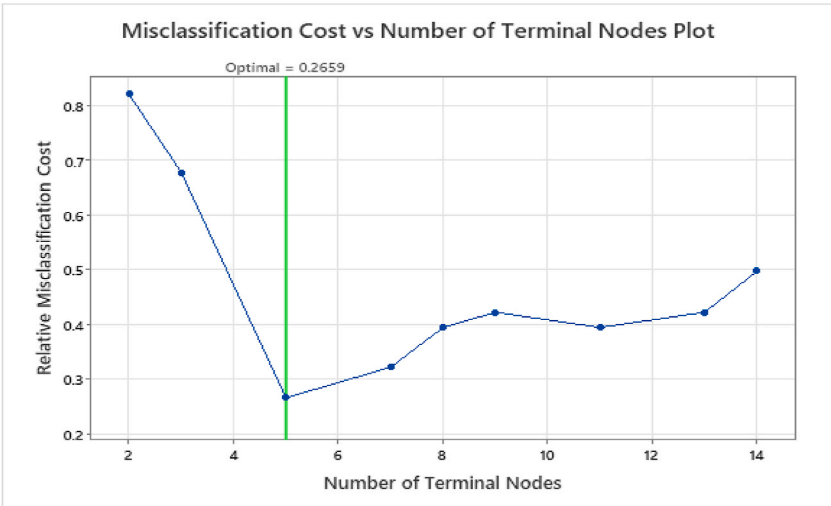


**Fig. 7.** The relative misclassification cost of the 5 terminal node classification tree within 1 standard error.

grains or cereals by market women could be an effective measure to reduce the health implications of aflatoxins on consumers. According to Tyroler [34], sorting is generally considered a women's role and that market women often allocate their best peanuts (sorted) up for sale and often consume peanuts of lower quality that are unknowingly contaminated with aflatoxins. The findings of this current work confirm Tyroler's assertions in that, gender was found to be one of the significant parameters to consider when the sorting behavior of sellers is concerned. The classification tree algorithm revealed that women are more likely to sort their grains than men. It was however worrying to note that sorted grains and cereals which are to be discarded are mostly consumed by sellers. This indicates that understanding the high risks of aflatoxin contamination is important and that information is not sufficiently disseminated to sellers, particularly women who are single and within the 30–39 years age bracket. Male sellers require even more understanding of the high risks of aflatoxin contamination and mitigation actions should target such groups.

## 4. Conclusion

The level of aflatoxin safety knowledge and awareness of grain and cereal sellers in the Greater Accra Region of Ghana was determined. Machine learning algorithm by means of Classification and Regression Trees was used to study how socio-demographic parameters affected aflatoxin safety knowledge and sorting behaviour of sellers. The relative variable importance chart determined which predictors were the most important and the variable with the highest improvement score on the model was set as the most

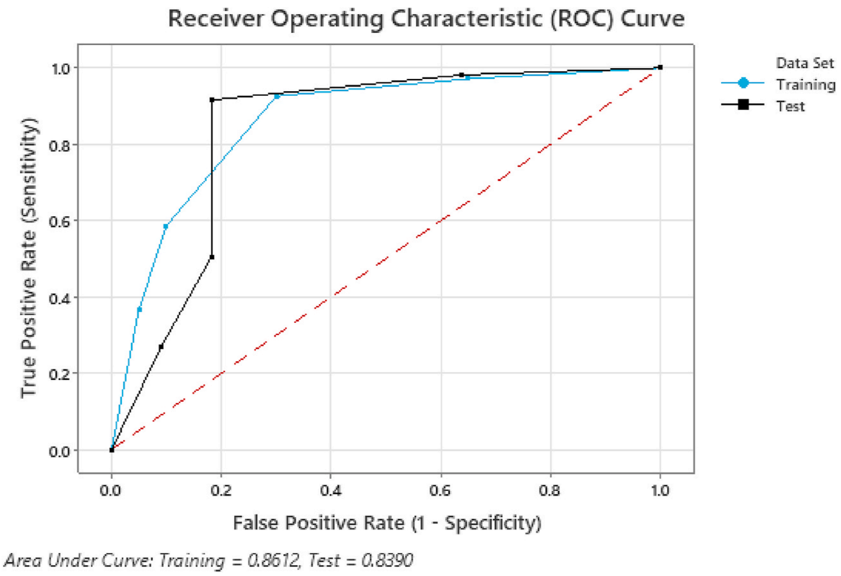


**Fig. 8.** Receiver operating characteristic (ROC) curve for the 5 terminal node classification tree algorithm.

important variable to affect aflatoxin safety knowledge score and the sorting behaviour of sellers. The aflatoxin safety knowledge score of sellers was below average. The ethnic group and educational background of sellers were the most important variables to consider during mitigation strategies. Regarding the sorting behaviour of sellers, their educational background was of utmost importance. Findings from this research highlight the necessity for more education and awareness creation of proper post-harvest handling and storage procedures for grains and cereals in order to prevent the production of carcinogenic toxins and assure food safety. Regulatory authorities must intensify their efforts to monitor and assure compliance to food safety standards within the Greater Accra Region of Ghana.

## Funding

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

## Author contribution statement

Vincent Owusu Kyei-Baffour: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Hilary Kwesi Ketemepi, Ebenezer Asiamah: Performed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Nancy Nelly Brew-Sam: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Leonora Charlotte Baffour Gyasi: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Wisdom Kofi Amoa-Awua: Conceived and designed the experiments; Wrote the paper.

## Data availability statement

Data included in article/supplementary material/referenced in article.

## Declaration of competing interest

Data of this study has been included in article/supplementary material/referenced in article. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We sincerely are grateful to all study participants who dedicated their time to respond to all questions posed at them. We are again thankful to the bilingual translator from the University of Ghana who assisted to translate the English version of questionnaire to the local language (*Twi*) and responses from *Twi* to English.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e18320.

## References


[1] European Food Safety Authority Panel on Contaminants in the Food Chain (CONTAM), D. Schrenk, M. Bignami, L. Bodin, J. K. Chipman, J. del Mazo, H. Wallace, Risk assessment of aflatoxins in food, EFSA J. 18 (3) (2020).

[2] R. Omari, E.K. Tetteh, S. Baah-Tuahene, R. Karbo, A. Adams, I.K. Asante, Aflatoxins and Their Management in Ghana: A Situational Analysis, FARA research report, 2020.

[3] N. Benkerroum, Chronic and acute toxicities of aflatoxins: mechanisms of action, Int. J. Environ. Res. Publ. Health 17 (2) (2020) 423.

[4] N.K. Kortei, T. Annan, V. Kyei-Baffour, E.K. Essuman, H. Okyere, C.O. Tettey, Exposure and risk characterizations of ochratoxins A and aflatoxins through maize (Zea mays) consumed in different agro-ecological zones of Ghana, Sci. Rep. 11 (1) (2021) 1–19.

[5] S. Baah-Tuahene, Assessment and Safety Improvement of Peanut Based By-Products from Peanut Oil Processing: Kulikuli and Khebab Powder, Doctoral dissertation, University of Ghana), 2015.

[6] D. Agbetiameh, A. Ortega-Beltran, R.T. Awuah, J. Atehnkeng, P.J. Cotty, R. Bandyopadhyay, Prevalence of aflatoxin contamination in maize and groundnut in Ghana: population structure, distribution, and toxigenicity of the causal agents, Plant Dis. 102 (4) (2018) 764–772.

[7] G.K. Blankson, F.C. Mills-Robertson, I.W. Ofosu, Survey of occurrence levels of aflatoxins in selected locally processed cereal-based foods for human consumption from Ghana, Food Control 95 (2019) 170–175.

[8] J. Kumi, N.J. Mitchell, G.A. Asare, E. Dotse, F. Kwaa, T.D. Phillips, N.A. Ankrah, Aflatoxins and fumonisins contamination of home-made food (weanimix) from cereal-legume blends for children, Ghana medical journal 48 (3) (2014) 121–126.

[9] K. Kpodo, A.K. Sørensen, M. Jakobsen, The occurrence of mycotoxins in fermented maize products, Food Chem. 56 (2) (1996) 147–153.

[10] F. Peles, P. Sipos, S. Kovács, Z. Győri, I. Pócsi, T. Pusztahelyi, Biological control and mitigation of aflatoxin contamination in commodities, Toxins 13 (2) (2021) 104.

[11] M. Wang, S.E. Hearon, T.D. Phillips, A high capacity bentonite clay for the sorption of aflatoxins, Food Addit. Contam. 37 (2) (2020) 332–341.

[12] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science 349 (6245) (2015) 255–260.

[13] M.J. Berry, G.S. Linoff, Data Mining Techniques: for Marketing, Sales, and Customer Relationship Management, John Wiley & Sons, 2004.

[14] Q. Wang, W. Jiao, P. Wang, Y. Zhang, A tutorial on deep learning-based data analytics in manufacturing through a welding case study, J. Manuf. Process. 63 (SI) (2021) 2–13.

[15] Y. Yoo, L.W. Tang, T. Brosch, D.K. Li, L. Metz, A. Traboulsee, R. Tam, Deep learning of brain lesion patterns for predicting future disease activity in patients with early symptoms of multiple sclerosis, in: Deep Learning and Data Labeling for Medical Applications: First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 21, 2016, pp. 86–94. Proceedings 1.

[16] N.Z. Zacharis, Classification and regression trees (CART) for predictive modeling in blended learning, IJ Intelligent Systems and Applications 3 (2018) 1–9.

[17] M. Camardo Leggieri, M. Mazzoni, P. Battilani, Machine learning for predicting mycotoxin occurrence in maize, Front. Microbiol. 12 (2021) 782.

[18] K. Yoo, H. Yoo, J.M. Lee, S.K. Shukla, J. Park, Classification and regression tree approach for prediction of potential hazards of urban airborne bacteria during Asian dust events, Sci. Rep. 8 (1) (2018) 1–11.

[19] M. Yannakoulia, A. Lykou, C.M. Kastorini, E.S. Papasaranti, A. Petralias, A. Veloudaki, DIATROFI Program Research Team, Socio-economic and lifestyle parameters associated with diet quality of children and adolescents using classification and regression tree analysis: the DIATROFI study, Publ. Health Nutr. 19 (2) (2016) 339–347.

[20] M. Pitesky, J. Gendreau, T. Bond, R. Carrasco-Medanic, Data Challenges and Practical Aspects of Machine Learning-Based Statistical Methods for the Analyses of Poultry Data to Improve Food Safety and Production Efficiency, CABI Reviews, 2020.

[21] R. Abukari, R.O. Odai, Gender and the labour market in Ghana: the relationship in terms of the family, the market and the State, Adv. Appl. Sociol. 8 (4) (2018) 285–294.

[22] T.N. Wood, Agricultural Development in the Northern Savannah of Ghana, The University of Nebraska-Lincoln, 2013.

[23] S.K. Twumasi, Understanding the Folk Islam of the Dagbani-Speaking People: a Prerequisite to Evangelism in North Ghana, Andrews University Digital Commons. Graduate Research, 1996.

[24] K.S. Magembe, M.W. Mwatawala, D.P. Mamiro, E.E. Chingonikaya, Assessment of awareness of mycotoxins infections in stored maize (Zea mays L.) and groundnut (arachis hypogea L.) in Kilosa District, Tanzania, Int. J. Flow Control 3 (1) (2016) 1–8.

[25] G.A.A. Anyebuno, V. Kyei-Baffour, D. Narh, Effect of manual sorting on Aflatoxins content in peanuts (Arachis Hypogaea, L.) from a Ghanaian market, Ghana J. Agric. Sci. 52 (2018) 5–15.

[26] B. Darfour, K.A. Rosentrater, Pre-harvest and post-harvest farmer experiences and practices in five maize growing regions in Ghana, Front. Nutr. (2022) 9.

[27] A. Ortega-Beltran, R. Bandyopadhyay, Contributions of integrated aflatoxin management strategies to achieve the sustainable development goals in various African countries, Global Food Secur. 30 (2021), 100559.

[28] M. Lawal, Prioritizing Food Safety in Ghana, Scientect Website (Issue April), 2021. https://scientect.org/2021/04/02/prioritizing-food-safety-in-ghana/. January 2023.

[29] D.L. Ortega, D.L. Tschirley, Demand for food safety in emerging and developing countries: a research agenda for Asia and Sub-Saharan Africa, J. Agribus. Dev. Emerg. Econ. 7 (1) (2017) 21–34.

[30] M.H. Kamano, M.W. Okoth, W. Kogi-Makau, P. Kuloba, Knowledge, Attitude and Practices (KAP) of Farmers on Postharvest Aflatoxin Contamination of Maize in Makueni and Baringo Counties, Kenya, 27nd International Sustainable Development Research Society Conference, Mid Sweden University, 2021.

[31] I. Adekoya, P. Njobeh, A. Obadina, C. Chilaka, S. Okoth, M. De Boevre, S. De Saeger, Awareness and prevalence of mycotoxin contamination in selected Nigerian fermented foods, Toxins 9 (11) (2017) 363.

[32] M. Aoun, W. Stafstrom, P. Priest, J. Fuchs, G.L. Windham, W.P. Williams, R.J. Nelson, Low-cost grain sorting technologies to reduce mycotoxin contamination in maize and groundnut, Food Control 118 (2020), 107363.

[33] M. Pascale, A.F. Logrieco, M. Graeber, M. Hirschberger, M. Reichel, V. Lippolis, K. Slettengren, Aflatoxin reduction in maize by industrial-scale cleaning solutions, Toxins 12 (5) (2020) 331.

[34] C. Tyroler, Gender considerations for researchers working in groundnuts, USAID Feed Futur 32 (2018), 1.