Heliyon 8 (2022) e10461

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Drying temperature-dependent profile of bioactive compounds and prediction of antioxidant capacity of cashew apple pomace using coupled Gaussian Process Regression and Support Vector Regression (GPR–SVR) model

Bobby Shekarau Luka [a,d,*], Taitiya Kenneth Yuguda [b], Meriem Adnouni [c], Zakka Riyang [d], Ibrahim Bako Abdulhamid [e], Bumbyerga Garboa Gargea [d]

[a] Department of Agricultural Engineering, Federal University Wukari, Taraba State, Nigeria
[b] College of Environment, Hohai University, Nanjing 210098, China
[c] Institution of Refrigeration and Cryogenics, Zhejiang University, Hangzhou, 310027, China
[d] Department of Food Science and Technology, Federal University Wukari, Taraba State, Nigeria
[e] Department of Computer Engineering, Federal University Wukari, Taraba State, Nigeria

ARTICLE INFO

Keywords:
Bioactive compounds quantification
Cashew apple pomace
Coupled GPR–SVR
Drying
Gaussian Process Regression modeling
Predicting antioxidant capacity
Support Vector Regression modeling

ABSTRACT

Crude extracts from cashew apple pomace (CAP) dried at different temperatures were used in High-Pressure Liquid Chromatography to quantify total alkaloids content (TAC), total flavanoids content (TFC), total saponin content (TSC) and total phenolics content (TPC). Diphenyl-1-picrylhydrazyl (DPPH) was used to determine the antioxidant capacity (AOC) of CAP. Fourier-Transformed Infrared Spectroscopy-Attenuated Total Reflectance (FTIR-ATR) was used to identify the functional groups present in the pomace. TAC, TFC, TSC and TPC were used as inputs to model AOC using Gaussian Process Regression (GPR), and Support Vector Regression (SVR) and a coupled model was developed using the residuals of GPR and SVR. It was found that increasing drying temperature decreased TAC, TFC, TPC and AOC but TSC increased. Both GPR and SVR predicted AOC with high accuracy. Drying CAP at lower temperature preserved more bioactive compounds hence high AOC; FTIR-ATR showed that CAP has good hydration capacity and contains majorly inorganic phosphates, aliphatic hydrocarbons and primary alcohols. Model coupling enhanced AOC prediction.

1. Introduction

The desire for medicinal value in foods has generated research interest on fruit residues and the influence of their processing methods on the profile and bioactivity of the secondary metabolites present in them. Cashew apple pomace has in recent times received renewed interest as substitute for cereals in food formulation such as cakes (Akubor et al., 2014; Adegunwa et al., 2020; Preethi et al., 2021), cookies (Ebere et al., 2015) and biscuits (Akubor, 2016) and had been used to improve the medicinal value of foods (Tańska et al., 2016). During the production of confectioneries, lipids (fat and oil) are added, and lipids tend to get rancid over time due to oxidation. As such, any endeavor that offers anti-oxidation effect through delay or inhibition is worthwhile. Owing to high moisture content in cashew apple pomace, prompt drying of the pomace to safe moisture content is necessary to inhibit microbial activity and biochemical metabolism within the matrix (Shekarau et al., 2020; Luka et al., 2018).

Heat sensitivity of bioactive compounds do not follow a definite order of response in biomaterials, bioactive compounds are preserved at higher temperature in some biomaterials (Nguyen and Le, 2018) and vice versa (Ahmad-Qasem et al., 2013; Yan and Kerr, 2012; Aldosari, 2014), this uncertainty has made every endeavor to understand the behavior of secondary metabolites present in fruits and vegetables worthwhile. The antioxidant capacity of fruits and vegetables is controlled by the quantity and types of bioactive compounds (Dabulici et al., 2015; Karasu et al., 2019). Determination of the antioxidant capacity of fruit and vegetable could be expensive and tedious, especially when the samples are bulky. As such, mathematical models could be in-

* Corresponding author at: Federal University Wukari, Nigeria.
E-mail address: lukabobby@fuwukari.edu.ng (B.S. Luka).

https://doi.org/10.1016/j.heliyon.2022.e10461
Received 13 November 2021; Received in revised form 22 January 2022; Accepted 22 August 2022

corporated into devices that can quantify bioactive compounds to at the same time project their antioxidant capacity with the aid of sensors.

Recently, the application of Gaussian Process Regression (GPR) and Support Vector Regression (SVR) in modeling has received special interest because GPR, being a new statistical machine learning algorithm in the context of Bayesian formulation, has the capability to use probabilistic regression to determine hyperparameters of multidimensional, small and nonlinear datasets (Lin et al., 2019), SVR has also derived eminence due to its capability to train small dataset size effectively and global minimization and maximization owing to its functionality base on structural risk minimization protocol.

Alfieri et al. (2019) used k-Nearest Neighbor to predict the antioxidant capacity of Sorghum using spectral data obtained from coupled Vis-NIR with different filter pretreatment, high coefficient of determination (R^2) value of 0.95 was obtained for the best performing model. Ahmadi et al. (2019) used Quantitative Structure-Activity Relationship (QSAR) to predict the antioxidant capacity of natural compounds, their model generated R^2 value for training and validation up to 0.86. Zhang et al. (2014) predicted the antioxidant capacity of Epimedium using chromatographic data with partial least square regression (PLSR). The model predicted the antioxidant capacity with R^2 of 0.9985 and RMSE of 22.8573. Martincic et al. (2015) developed Predictive Counter Propagative-Artificial Neural Network Models (CP-ANN) and SVR based on machine learning algorithms and statistical Multilinear regression (MLR) model to predict the antioxidant capacity of natural compounds using radiation data, the models generated R^2 values of 0.93, 0.86 and 0.89 for CP-ANN, SVR and MLR models, respectively. From pertinent literature, all the existing developed models used to predict the antioxidant capacity of biomaterials used data from spectroscopy, not from an actual profile of secondary metabolites. Models developed using spectral data are sometimes inconsistent due to the difference in the efficiency of pretreatment methods to accurately filter noise and the ability of variable selection techniques to select wavelengths that truly correlate with the response variable. Generally, studies on the prediction of antioxidant capacity of fruits and vegetables and their residues are scarce, this study uses quantitative data of metabolites to develop predictive models, and to the best of our knowledge is the first study to be carried out using this approach. The developed models were also coupled using a novel approach to enhance the predictive power, the first approach to predicting antioxidant capacity.

1.1. Gaussian Process Regression (GPR) modeling

GPR model is one of the sought after Bayesian machine learning modeling approaches due to its ability to work better with small datasets with no problem of underfitting or overfitting and is capable of giving information on the uncertainty of the output of the model (Wenqi et al., 2018). It perfectly learns nonlinear datasets using the probabilistic technique and estimates posterior degradation by limiting the prior distribution to fit the supplied dataset (Liu et al., 2009).

A Gaussian process (GP) is explained by its covariance function $C(x, x')$ and mean function $\mu(x)$ expressed as (Equations (1)–(5))

$$f(x) \sim GP(C(x, x'), \mu(x)) \tag{1}$$

$$\mu(x) \sim E[f(x)] \tag{2}$$

$$C(x, x') = E[(f(x) - \mu(x))(f(x') - \mu(x'))] \tag{3}$$

Where x and x′ are randomly chosen variables, E is the average expected output used in probabilistic predictions.
A noisy GPR model is expressed as

$$y_i = f\left(x_i\right) + \epsilon \tag{4}$$

Where ϵ is the Gaussian additive noise $N(0, \sigma^2)$.

The input, x^* are imported to establish Gaussian prior distribution of the training response, y and the test response, y^* as

$$\begin{bmatrix} y \\ y^* \end{bmatrix} = N\left(0, \begin{bmatrix} C(X, X^*)\sigma_n^2 I & C(X, x^*) \\ C(X, x^*) & C(x, x^*) \end{bmatrix}\right) \tag{5}$$

$C(X, X^*)$ is a positive definite covariance matrix of n-order symmetry, $C(X, x^*)$ is the covariance matrix $(n \times 1)$ of the test input x^* and the training input X and $C(x, x^*)$ is the covariance matrix of the test input dataset x^* under the stipulation of a given input x^* and the training dataset D. The Gaussian process calculates the test output variable, y^* following the posterior probability formula presented in Equations (6) and (7)

$$y^*[X]^*, D = N\mu \tag{6}$$

$$\mu_y = \frac{1}{C(x^*, X)\left(C(X, X) + \sigma_n^2 I\right) y} = \alpha C\left(x^i, x^*\right) \tag{7}$$

The four most commonly used covariance functions (squared exponential, rational quadratic, Matern class and exponential) in GPR modeling are expressed in Equations (8)–(11).
The squared exponential

$$C_{SE}(x^i x^j) = \sigma_f^2 \exp\left(-\frac{(x^i - x^j)^2}{2l^2}\right) \tag{8}$$

Rotational quadratic (Cauchy)

$$C_{RQ}(x^i x^j) = \sigma_f^2 \left(-\frac{(x^i - x^j)^2}{2\alpha l^2}\right)^{-\alpha} \tag{9}$$

The Matern class

$$C_M\left(x^i x^j\right) = \sigma_f^2 \left[1 + \sqrt{2M}\left(x^i - x^j\right) e^{\sqrt{3M}(x^i - x^j)}\right] \tag{10}$$

Exponential

$$C_E\left(x^i x^j\right) = e^{-\left(\frac{\left\|x^i - x^j\right\|^2}{2\sigma^2}\right)} \tag{11}$$

Where σ_f^2, α, l, M are hyperparameters, and the i and j constitute the i-th and the j-th vector in the input matrix X.

1.2. Support Vector Regression (SVR) modeling

SVR is one of the most reliable machine learning or statistical learning algorithms (Maroco et al., 2011) and can be used for both classification problems and regression problems. Support vector regression extends support vector classification that utilizes continuous responses; both applications use kernel function to achieve their target (Guenther and Schonlau, 2016). Unlike GPR, which minimizes squared error (ϵ) loss function and loss for responses, i-th, (quadratic loss), SVR minimizes ϵ-insensitive loss function such that any loss below the error margin is set to zero and above that constraint, the linear loss function is utilized as in Equation (12)–(14)

$$l_\epsilon = \begin{cases} 0 \\ |y_i - f(x_i)| \end{cases} \tag{12}$$

if $|y_i - f(x_i)| < \epsilon$, otherwise, the loss function for a linear function is presented as

$$f(x) = \beta_o - x_i^t \beta \tag{13}$$

And the loss function $\sum_{i=1}^{n} \max(y_i - x_i^t \beta - \beta_o - \epsilon, 0)$ (14)

Where ϵ represents the turning parameter and can be expressed as a constraint optimization formulation as presented in Equations (15)–(16)

$$\text{minimize } \frac{1}{2}\|\beta\|^2 \tag{15}$$

$$\text{subject to } \begin{cases} y_i - x_i^t \beta - \beta_o \le \epsilon, \\ -(y_i - x_i^t \beta - \beta_o) \le \epsilon \end{cases} \tag{16}$$

If all the measured variables fall outside the error bound, it will not generate any solution, hence slack variables, γ_i and γ_i^* are inserted to bring the observations within the regression line as expressed in Equation (17)

$$\begin{cases} y_i - x_i^t\beta - \beta_o \le \epsilon + \gamma \\ -\left(y_i - x_i^t\beta - \beta_o\right) \le \epsilon + \gamma_i^* \\ \gamma_i\gamma_i^* \ge 0 \end{cases} \tag{17}$$

The support vectors are those observations that fall outside the error bound. Kernel functions expedite the computation. These kernels create a window to influence the data and are presented in Equations (18)–(21).

$$\text{Gaussian Kernel (exponential)},\ K(x,y) = e^{-\left(\frac{\|x-y\|^2}{2\sigma^2}\right)} \tag{18}$$

$$\text{Gaussian Kernel radial basis function},\ K(x,y) = e^{-\left(\gamma\|x-y\|^2\right)} \tag{19}$$

$$\text{Sigmoid kernel},\ K(x,y) = \tanh\left(\gamma x^T y + r\right) \tag{20}$$

$$\text{Polynomial},\ K(x,y) = \tanh\left(\gamma x^T y + r\right)^d,\ \gamma > 0 \tag{21}$$

2. Materials and methods

2.1. Experimental methods

2.1.1. Drying of cashew apple pomace

Fresh yellow cashew apple fruits were obtained and then washed with distilled water. The cashew apple was then shredded into smaller sizes, fed into a fruit blender, and blended to obtain a paste. The paste was then tied in a filter material and pressed using a hydraulic press to express all the juice according to the method of Shekarau et al. (2020). The pomace was kept in a plastic container and stored in a freezer at −4 °C. The pomace was then divided into twenty-seven (27) equal samples and each sample was dried at different temperatures of 60, 65, 70, 75, 80, 85, 90, 95 and 100 °C in a hot air oven drier at a constant air velocity of 2 m/s until they reach equilibrium moisture content. The process was carried out in triplicate.

2.1.2. Microwave pretreatment of dried cashew pomace

The dried pomace samples were grounded into fine particles using a blender. The microwave was purged with nitrogen to rid off oxygen. The first pomace sample was immediately fed into it in a thin layer and pretreated at 100 W for 15 minutes. This was necessary due to high pectin crystals formed due to drying and the high fiber content of the pomace. The nitrogen was used to inhibit oxidation while microwave power increased the cell porosity of the sample. The process was repeated for all the samples at constant time and microwave power. The pretreated samples were then poured into a cartridge water filter case with one end covered with filter material. Carbon dioxide was used for flushing out oxygen from the container via the inlet and discharged at the outlet and then tightly sealed with coke stopper and was wrapped in an aluminum foil and stored in a freezer at −4 °C.

2.1.3. Extraction of bioactive compounds

Cashew apple pomace powder samples weighing (1.5 g) were dissolved in a solution of ethanol and water as solvents in a volumetric proportion of 1 : 1 of 15 mL and left under dark for 3 days at room temperature and homogenized periodically after six hours using a homogenizer (Bioben JRJ–300, China). The solutions were filtered using filter paper, and the crude extract was utilized for High-Pressure Liquid Chromatography (HPLC) analysis.

2.2. Analysis

2.2.1. Quantification of bioactive compounds using HPLC

Half gram of crude extract was dissolved in hexane in a 1.0 mL vial. Then, the vial contents were injected into a HPLC coupled with a fluorescence detector (Buck BLC10/11, USA) with excitation at 297 nm and emission at 325 nm. An analytical silica column (25 cm × 4.6 mm ID, stainless steel, 5 μm) was used to analyze the bioactive compounds present. The mobile phase employed was a mixture of hexane, tetrahydrofuran and isopropanol (1000 : 60 : 4) and fed at a volumetric flow rate of 1.0 mL/min. Standard samples for total alkaloid content (TAC), total flavanoids content (TFC), total saponin content (TSC) and total phenolic content (TPC) were also prepared using the same procedure. The concentrations of the bioactive compounds in the samples were calibrated using the standard sample, and quantification of the bioactive compounds was carried out by comparing peak areas and respective retention times with authentic standard compounds. The procedure was carried out in triplicate, and their standard deviation was calculated according to Luka et al. (2020). The concentration of each bioactive compound was calculated using Equation (22).

$$\text{T-BioComp} = \frac{\text{A Sample} \times \text{STD} \times \text{V HEX}}{\text{A STD} \times \text{Wt Sample}} \tag{22}$$

Where;

T-BioComp = Concentration of bioactive compounds (mg/g)
STD = Concentration of standard
A Sample = Area of sample
A STD = Area of standard
V HEX = Volume of hexane
Wt Sample = Weight of sample

2.2.2. Diphenyl-1-picrylhydrazyl (DPPH) scavenging activity (Antioxidant capacity) using UV-spectrophotometry

Forty milligrams of DPPH solution were dissolved in 100 mL of methanol to form a stock solution; the solution was kept in a freezer at −20 °C.

Three-hundred-fifty milliliters of the stock solution were mixed with 350 mL of methanol to obtain the absorbance of 0.70 ± 0.01 unit at 516 nm wavelength by using a UV-spectrophotometer (Jenway 6405, China).

2.2.3. Determination of functional groups using FTIR-ATR analysis

Powdered cashew apple pomace was placed in the attenuated total reflectance (ATR) crystal, and the spectra were obtained using Fourier-Transformed Infrared Spectroscopy (FTIR) (Nicolet, iS50 FTIR, Uppsala, Sweden) coupled ATR (platinum diamond) comprising of internal reflectant and diamond disc at 4000-400 cm^{-1}, air spectrum was used as the background spectra for calibration, the spectra from the samples were obtained at room temperature of approximately 32 °C and resolution of 4 cm^{-1}.

2.3. Modeling

2.3.1. GPR and SVR modeling

Both GPR and SVR models were trained using MATLAB 2019a (MathWorks Inc., Natick, MA, USA), the obtained quantities of TAC, TFC, TSC and TPC were used as the models' input variables. The GPR model used 5-fold cross-validation, isentropic kernel, and automatic kernel scaling and was set to standardize and optimize numeric parameters automatically. All the covariance functions in Equations (8)–(11) were tested, and their performance was evaluated.

The SVR model was trained using automatic box constraint and standardized data to select the support vector variables, all the kernel functions in Equations (18)–(21) were tested and evaluated. The best performing covariance and kernel functions for GPR and SVR were selected as the training kernel.

2.3.2. Model coupling

The two trained models (GPR and SVR) were coupled in parallel using a unique coupling procedure by utilizing the residual terms (R^2 and RMSE) using Equations (23)–(28) as presented in Fig. 1.

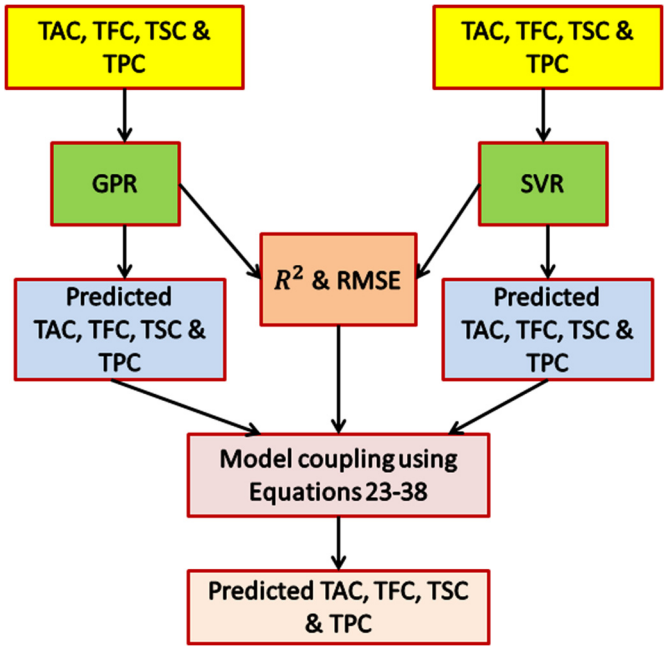

Fig. 1. GPR–SVR model coupling.

The generated residual terms were used as input weight against the predicted outputs of GPR and SVR models. Source code was first generated using MATLAB and involved allocating function to the generated model file 'model.c', the codes were then compiled (for each model) into a makefile. The individual makefiles were then combined into a single makefile. A combined simulation engine was then created via modification of the main program (rt_malloc_main,c); this helped the program initialize and call each model's model codes. The makefile was then run to compute the coupled model based on the supplied residual coupling formulation and the generated predicted output from GPR and SVR.

$$\varepsilon_m = \frac{\varepsilon_{SVR} + \varepsilon_{GPR}}{2} \tag{23}$$

$$R^2_{resid(SVR)} = 1 - R^2_{(SVR)} \tag{24}$$

$$R^2_{resid(GPR)} = 1 - R^2_{(GPR)} \tag{25}$$

$$\varepsilon_{adj(SVR)} = \frac{(\varepsilon_m \times R^2_{resid(SVR)})}{n} \tag{26}$$

$$\varepsilon_{adj(GPR)} = \frac{(\varepsilon_m \times R^2_{resid(GPR)})}{n} \tag{27}$$

$$y^*_{coupled} = \left[\frac{(y^*_{i(SVR)} + \varepsilon_{adj(GPR)}) + (y^*_{i(GPR)} + \varepsilon_{adj(SVR)})}{2}\right] \tag{28}$$

$i = 1, 2, 3 \ldots n$

Where;

ε_m = Mean error

ε_{adj} = Adjusted error

$R^2_{resid(SVR)}$ = Residual coefficient of determination of SVR model

$R^2_{resid(GPR)}$ = Residual coefficient of determination of GPR model

$R^2_{(SVR)}$ = coefficient of determination of GPR model

$R^2_{(GPR)}$ = coefficient of determination of GPR model

ε_{SVR} and ε_{GPR} are the mean square error of SVR and GPR models, respectively

$y^*_{i(SVR)}$ and $y^*_{i(GPR)}$ are predicted values of SVR and GPR models, respectively

3. Results and discussion

The heat sensitivity of bioactive compounds and their antioxidant capacity after drying are presented in Table 1 and discussed accordingly.

3.1. Influence of drying temperature on the profile of TSC

The profile of TSC as influenced by drying temperature in cashew apple pomace is presented in Table 1. It can be seen that drying cashew apple pomace at 60 °C gave the lowest TSC of 8.76 mg/g, as drying temperature increased from 60 to 65 °C, TSC reduced by 22.26%, however, as drying temperature was further increased to 70, 75, 80, 85, 90, 95, and 100 °C, TSC increased significantly ($p < 0.05, 0.019$) by 47.28, 20.93, 10.79, 17.18, 5.71, 12.97 and 13.87%, respectively.

High drying temperature favors the preservation of saponins. The highest value of 21.42 mg/g was obtained at 100 °C. The decrease in TSC at 65 °C can be due to distortion in chemical structure, which makes their extraction and determination inefficient (Qu and Ma, 2010); or the splitting of a transesterified bond and glycosylated bond, which, consequently, predispose saponins to hydrolase enzymes (Xu et al., 2007). Shekarau et al. (2020) attributed the transesterification of cashew apple pomace, during drying, to the high pectin content and gelatinization. When drying temperature increased from 65–100 °C, TSC also increased, showing a negative correlation (Fig. 3) with other metabolites, such response by TSC to increasing drying temperature could be attributed to the fact that at lower drying temperature, inactivation of oxidizing enzymes could not be achieved, thereby resulting to the rapid oxidation of saponins (Mrad et al., 2012; Youssef and Mokhtar, 2014). The sharp increase observed at 70 °C can be attributed to increased access to the compounds in the pomace matrix during extraction and shorter drying time, consequently inactivating saponin hydrolase enzyme, impeding the degradation of saponins. Generally, the preservation of TSC at elevated temperature could be ascribed to the presence of compounds from which saponins are formed by anabolic reaction (Que et al., 2008).

3.2. Influence of drying temperature on TAC, TFC and TPC in cashew apple pomace

The profile of TAC, TFC and TPC exhibited a similar response to increasing drying temperature (Table 1) and are discussed accordingly.

The highest value of TAC (36.32 mg/g) was obtained when the pomace was dried at 60 °C. As the drying temperature increased from 65–100 °C, the TAC significantly ($p < 0.05, 0.022$) declined by 1.34, 10.54, 9.82, 11.66, 6.85, 13.54, 11.91 and 10.21%, respectively.

Drying cashew apple pomace at 60 °C gave the highest TFC of 16.65 mg/g, however, as the drying temperature was increased from 60–90 °C, the quantity of TFC decreased significantly ($p < 0.05, 0.027$) by 14.75, 24.92, 4.89; at 100 °C, TFC increased by 3.30%.

The drying temperature-dependent profile of TFC follows the first-order kinetic with the highest value of 91.62 mg/g obtained at 60 °C drying temperature. From 60–100 °C, the quantity of TFC decreased significantly ($p < 0.05, 0.013$) by 1.08, 5.18, 4.94, 4.22, 2.87, 9.45, 8.57, 4.62%, respectively.

The drying temperature-dependent profile of TAC, TFC and TPC shows a negative relationship with drying temperature and practically implies that the compounds mentioned above in cashew apple pomace are heat sensitive. While many researchers attribute such decrease to thermal degradation and cell lysis leading to leakage and break down of bioactive compounds, Ionnou et al. (2012) added that degradation in TFC and TPC may not always necessarily be attributed to heat sensitivity. Still, the presence of oxygen and other bioactive compounds, the presence of oxygen highly degrades quercetin and rutin (bioflavanoid).

The trend of first-order kinetic that exists between drying temperature and TFC can also be due to the splitting of glycosylated bonds,

Table 1. Profile of bioactive compounds from cashew apple pomace and their antioxidant capacity at different drying temperatures.

Drying temperature (°C)	TAC (mg/g sample)	TFC (mg/g sample)	TSC (mg/g sample)	TPC (mg/g sample)	DPPH inhibition (%)
60	36.3 ± 0.1	16.61 ± 0.09	8.76 ± 0.52	91.63 ± 0.48	32.15 ± 0.43
65	35.83 ± 0.05	14.2 ± 1	6.81 ± 0.03	89 ± 1	31.8 ± 0.1
70	32.1 ± 0.3	10.6 ± 0.9	10.03 ± 0.03	87.11 ± 0.37	30 ± 1
75	28.9 ± 0.1	10.1 ± 0.6	12.13 ± 0.34	82 ± 2	29 ± 1
80	25.53 ± 0.05	9.6 ± 0.2	13.44 ± 0.76	78 ± 1	27 ± 1
85	23.78 ± 0.06	8.54 ± 0.11	15.75 ± 0.05	77 ± 1	26.7 ± 0.2
90	21 ± 1	7.6 ± 0.5	16.65 ± 0.86	72 ± 1	24.14 ± 0.19
95	18.1 ± 0.2	6.6 ± 0.7	18.81 ± 0.27	70.32 ± 0.48	22 ± 1
100	16.3 ± 0.2	6.9 ± 1	21.42 ± 0.45	66 ± 1	21 ± 1

Mean value for n number of test (n = 3) and standard deviation.

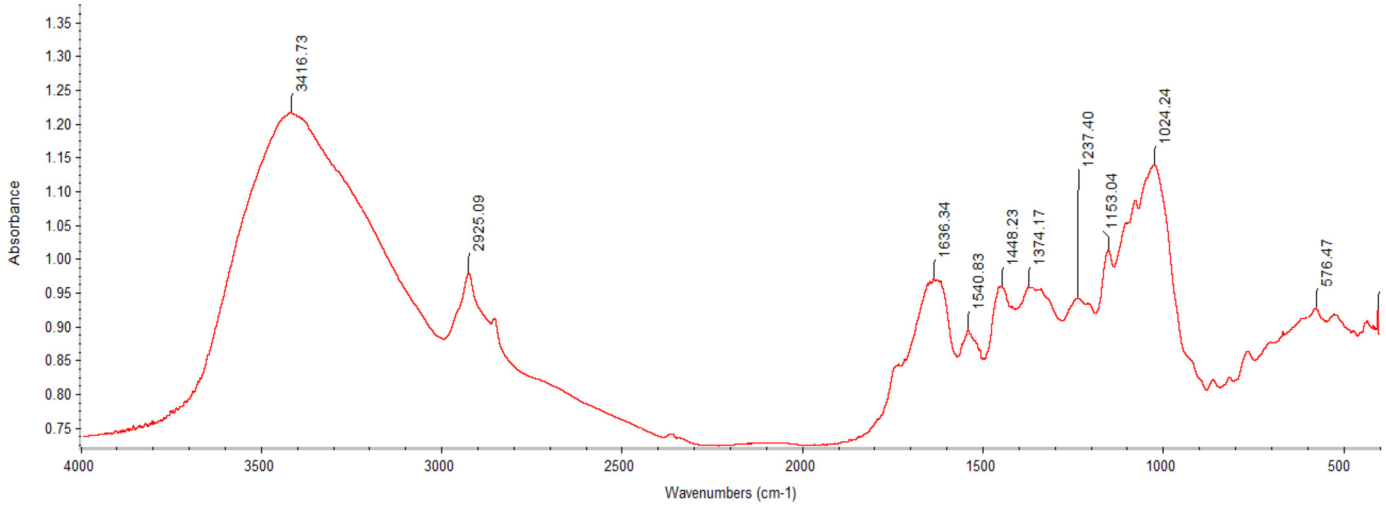

Fig. 2. FTIR-ATR light absorbance profile of cashew apple pomace at different wavenumber.

which predisposes phenolics to oxidase enzymes attack and degradation. It can also be ascribed to polyphenols interaction with protein (Xu et al., 2007; Preethi et al., 2021) or to the inability of the heat treatment to neutralize the activity of oxidase enzymes in the pomace matrix and the distortion in chemical structure, which make their extraction and quantification inefficient (Qu and Ma, 2010). De Ancos et al. (2000) added that, besides heat treatment, polyphenols and alkaloids might deteriorate due to organic acid content, sugar concentration and pH. The higher quantity of TPC relative to other bioactive compounds present in CAP has also been reported in Preethi et al. (2021). Similar temperature-dependent relationships were reported in a study on vacuum-belt drying of apple pomace powder by Yan and Kerr (2012), oven drying of olive pomace by Ahmad-Qasem et al. (2013), cabinet drying of apple pomace by Aldosari (2014) and hot air drying of maqui berries by (Rodríguez et al., 2016).

3.3. Influence of drying temperature on AOC of cashew apple pomace by DPPH assay

The DPPH antioxidant capacity (oxygen scavenging activity) of cashew apple pomace decreased with an increase in drying temperature (Table 1). This can be because the quantity of total antioxidants (TAC, TFC and TPC) continuously decreased as drying temperature was increased progressively. The highest oxygen scavenging activity of 32.15% inhibition was obtained at 60 °C drying temperature. It can be seen that despite the increase in TSC in the pomace with increasing drying temperature, the oxygen scavenging activity did not respond similarly. This can be attributed to heating to an elevated temperature. The heating impedes its ability to donate electrons (Sim et al., 2017); thus, TSC showed little or no contribution to the overall antioxidant capacity. This was corroborated by the outcome of sensitivity analysis, where only TPC significantly influenced the change in oxygen scavenging activity, this further substantiates the finding by Sumic et al. (2017) that TPC governs antioxidant capacity in vacuum dried mushroom. Similarly, in a study by Preethi et al. (2021), it was obvious that increasing the quantity of cashew apple pomace powder in cereal based extrudate steeply increases the TFC and TPC and their corresponding AOC.

3.4. Functional groups present in cashew apple pomace

The functional groups detected at different wavenumbers by FTIR-ATR from the cashew apple powder (Fig. 2) were majorly inorganic phosphates, aliphatic hydrocarbons and primary alcohols.

The peak absorption at 3416.73 cm^{-1} wavenumber characterized by both stretch in O–H and C–H bonds corresponds to alcohols functional group. The waveband at 2925.09 cm^{-1} corresponds to the assigned absorbance of the esterified carboxyl group characterized by O–H stretch; the esterified carboxyl group confirms the functional identity of pectin. Alkenes and phenol were identified where C=C stretch (conjugate) and O–H bending occur at a peak wavenumber of 1636.34 cm^{-1} and 1374.17 cm^{-1} respectively, in cellulosic fibers, at the wavenumber 1636.34 cm^{-1} adsorbed water form a major assignment of the wavenumber. Amides and Nitro groups were also identified at a peak value of 1540.83 cm^{-1} where N–H bend and NO_2 bend occur, respectively, the presence of amides characterized by the N–H bond can be due to amino acid breakdown. C–O stretch and C–C(O)–stretch were detected at a peak wavenumber of 1237.40 cm^{-1}, which corresponds to alcohols and esters functional groups. The esters can be formed due to the reaction between carboxylic acid with alcohol to release water. When the absorbance peaked at 1153.04 cm^{-1}, C–O stretch and C–N stretch occurred, corresponding to the spectroscopic infrared absorption of Anhydrides and Amines. Phosphines functional group was

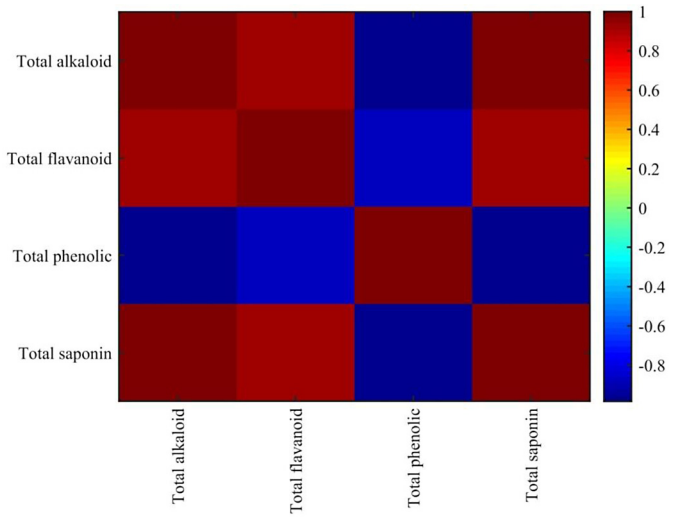

Fig. 3. Correlation matrix of bioactive compounds from dried cashew apple pomace.

also detected at a wavenumber of 1024.24 cm^{-1}, characterized by P–H bend in the molecular motion. Peaked absorbance obtained at 576.47 cm^{-1} falls within the bracket of wavenumber of alkyl halides functional group, where C–Br stretch, C–Cl stretch and C–I stretch characterized the molecular motion at that wavenumber.

The presence of amide (a derivative in which the hydroxyl group of oxoacid is replaced by carboxylic acid) in cashew apple pomace indicates the presence of pectin (Sato et al., 2011). Pectin was reported by Boulos et al. (2000) to have high water holding capacity or absorption (hydration) property, and this is desirable because it facilitates stool bulking (Sahni and Shere, 2018) aided by the high fiber content of the pomace powder. However, this finding also infers that the hydrophilic nature of CAP may pose a threat to its shelf life and make it lose crispiness.

The identification of amines needs further characterization to rule out the presence of biogenic amines in high concentration since their accumulation in the body can cause toxicity.

3.5. Sensitivity analysis (variable selection)

To choose a reliable and accurate sensitivity analysis approach, Pearson correlation was carried out on the measured bioactive compounds, and this is necessary to ascertain the level or degree of linear statistical interaction between all the bioactive compounds at different drying temperatures and to check for possible collinearity between the variables, the correlation was checked base on Equation (29)

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}} \quad (29)$$

Where x_i is the value of each independent variable in the sample, $\overline{x}$ is the mean value of each independent variable in the sample.

It could be seen in Fig. 2 that there exist a very strong positive and negative correlation between the quantity of all bioactive compounds, the positive correlation ranges from 0.89 to 0.99 and the negative correlation ranges from −0.98 to −0.89, negative correlation indicates inverse relationship between the interacting compounds, such strong collinearity between the metabolites has strong influence on the integrity of the outcome of sensitivity analysis approach used, thus forward step variable selection (Sequential feature selection) which is not affected by multicollinearity was adopted and the selection procedure was carried out using MATLAB 2019a (MathWorks Inc., Natick, MA, USA). Sequential analysis base on defined criterion function was used to call cross-validation partition to evaluate the criterion, the evaluation returned TPC as the most informative variable with estimated dispersion of 0.504 hence, the only variable that significantly influences the antioxidant capacity of cashew apple pomace at $p < 0.05$.

3.6. Developed models and coupled model

Both GPR and SVR models proved adequate for simulating and predicting antioxidant capacity of dried cashew apple pomace at different hot air drying temperatures; this is corroborated by the higher coefficient of determination and low residual presented in Figs. 4 and 5. The rotational quadratic covariance function outperformed all other functions in GPR, while the quadratic kernel function proved more adequate among its counterpart kernel function in the SVR model.

Every modeling endeavor aims to obtain zero residual, which translates to perfect prediction of the response variables. The coupled model reduced the residual further, this can be seen by the closeness of the true and predicted responses to the line of perfect prediction and closeness of the residuals to the line of zero on Fig. 6 respectively, thus confirming our novel coupling approach as a promising strategy for enhancing machine learning prediction of antioxidant capacity of dried cashew apple pomace. Both GPR and SVR performed comparably with models developed by Alfieri et al. (2019), Ahmadi et al. (2019), Zhang et al. (2014) & Martincic et al. (2015), while the coupled model performed favorably with higher accuracy than the previous models. Our models have proved adequate for predicting AOC of CAP using but few data samples, and were developed solely using the quantity of the most abundant bioactive compounds in CAP with proven high oxygen scavenging activity.

4. Conclusion

A study on the identification of functional groups, the influence of drying temperature on the quantity of selected bioactive compounds and, in turn, bioactive compounds on the antioxidant capacity of cashew apple pomace and prediction of antioxidant capacity of cashew apple pomace has been carried out; the results draw the following conclusions.

- Increasing drying temperature decreases the quantity of total alkaloid, total flavanoid, and total phenolic while total saponin increases; thus, drying at a lower temperature is more beneficial.
- The antioxidant activity of cashew apple pomace is dominantly controlled by the quantity of total phenolics and decreases with increasing drying temperature.
- The functional groups found in cashew apple pomace are basically inorganic phosphates, aliphatic hydrocarbons and primary alcohol; the presence of pectin/amines also infers good water holding capacity of the fruit powder.
- Coupled GPR and SVR model using unique coupling terms was found to further enhance the prediction of antioxidant capacity of cashew apple pomace compared to individual models.

Declarations

Author contribution statement

Bobby Shekarau Luka: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Taitiya Kenneth Yuguda: Analyzed and interpreted the data.

Meriem Adnouni: Performed the experiments.

Zakka Riyang & Bumbyerga Garboa Gargea: Contributed reagents, materials, analysis tools or data.

Ibrahim Bako Abdulhamid: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

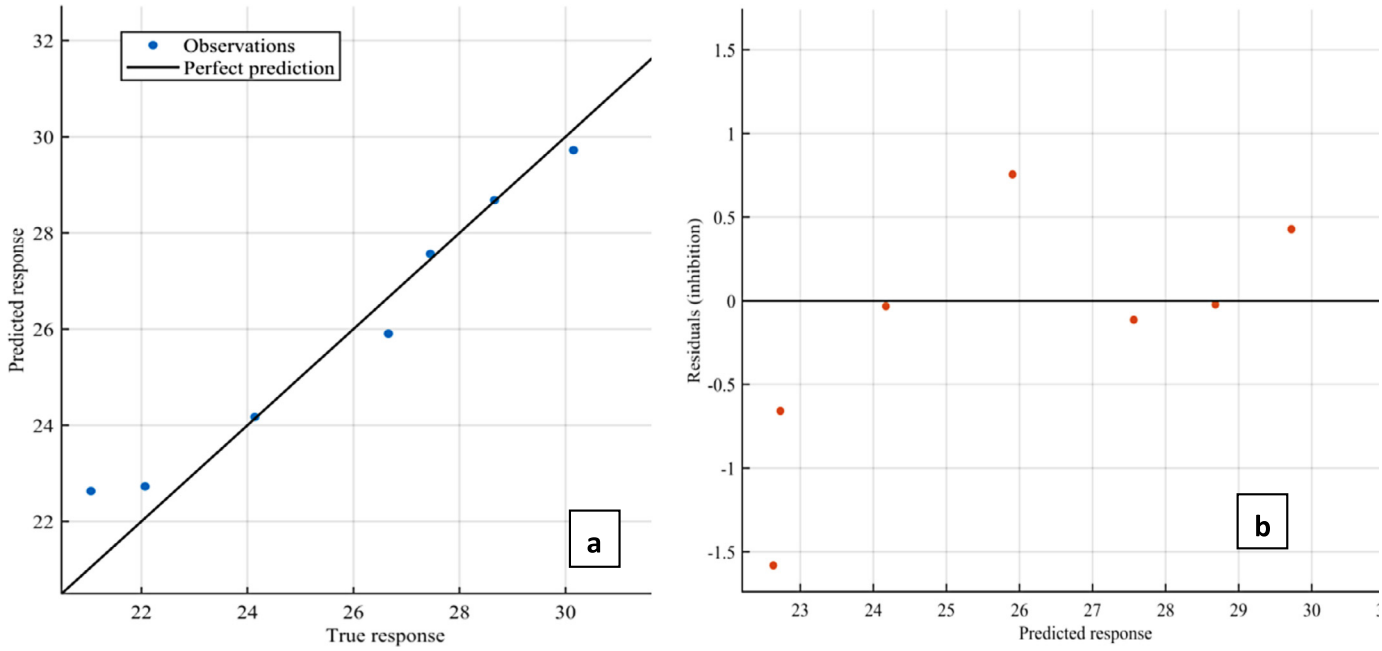

Fig. 4. Prediction performance of GPR model (a) Correlation (b) Residual.

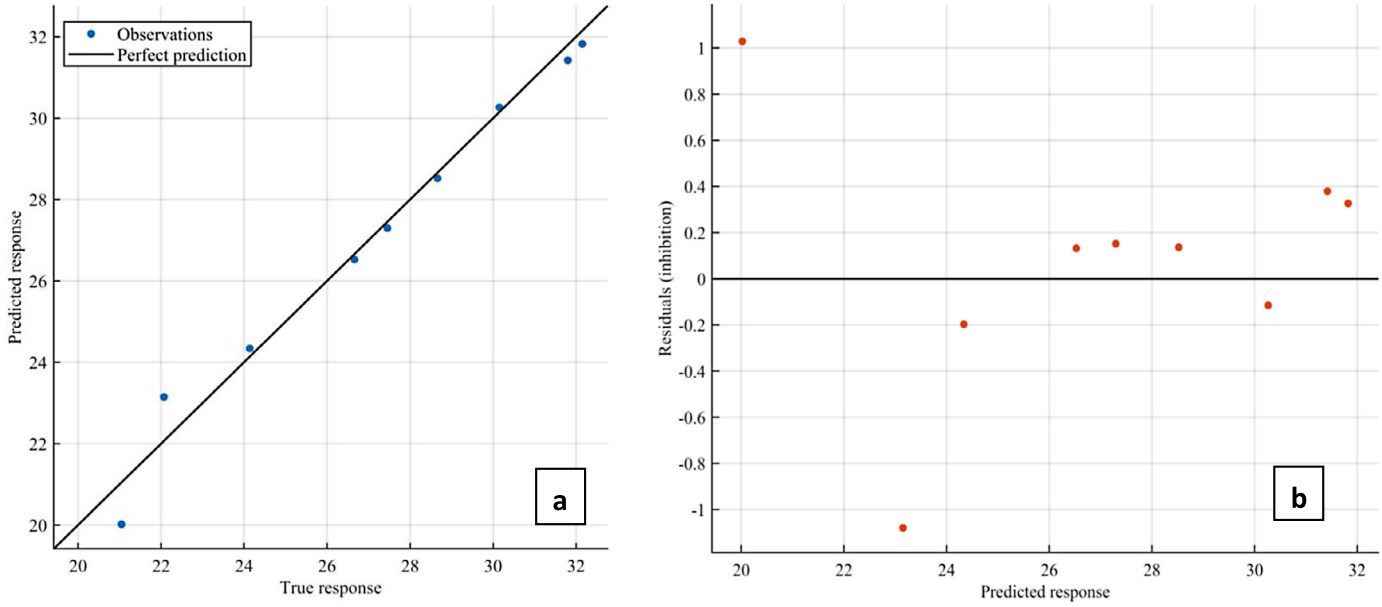

Fig. 5. Prediction performance of SVR model (a) Correlation (b) Residual.

Funding statement

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

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Adegunwa, M.O., Kayode, B.I., Kayode, R.M.O., Akeem, S.A., Adebowale, A.A., Bakare, H.A., 2020. Characterization of wheat flour enriched with cashew apple (Anacardium occidentale L.) fiber for cake production. J. Food Meas. Charact. 14, 1998–2009.

Ahmadi, S., Ghanbari, H., Lotfi, S., Azimi, N., 2019. Predictive QSAR modeling for the antioxidant activity of natural compounds derivatives based on Monte Carlo method. Mol. Divers. 25, 87–97.

Ahmad-Qasem, M.H., Barrajon-Catalan, E., Micol, V., Carcel, J.A., Garcia-Perez, J.V., 2013. Influence of air temperature on drying kinetics and antioxidant capacity of olive pomace. J. Food Eng. 119, 516–524.

Akubor, P.I., 2016. Chemical composition, physical and sensory properties of biscuits supplemented with cashew pomace flour. NSUK J. Sci. Technol. 6 (2), 138–141.

Akubor, P.I., Egbekun, M.K., Obiegbuna, J.E., 2014. Quality assessment of cakes supplemented with cashew pomace and soybean flour. Discovery 9 (20), 8–13.

Aldosari, M.T., 2014. Effects of different drying methods on the total phenolics, antioxidant properties, and functional properties of apple pomace. Unpublished thesis. Department of Food Science, Michigan State University, USA.

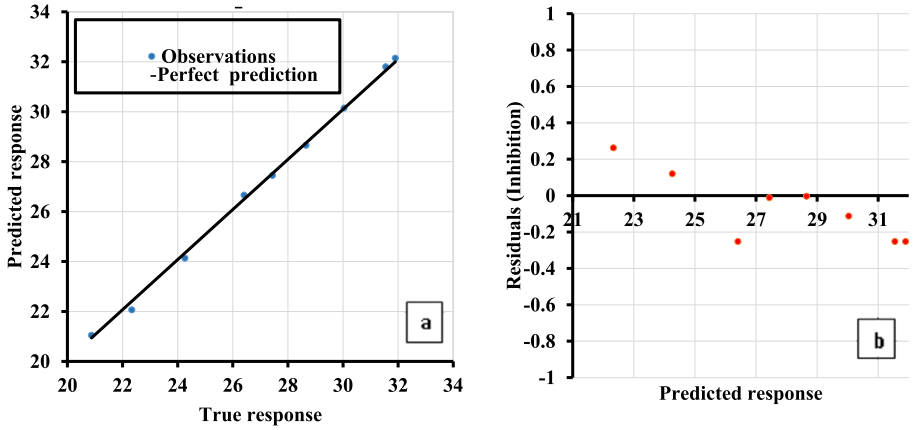

Fig. 6. Prediction performance of the coupled model (a) Correlation (b) Residual.

Alfieri, M., Cabassi, G., Habyarimana, E., Quaranta, F., Balconi, C., Redaelli, R., 2019. Discrimination and prediction of polyphenolic compounds and total antioxidant capacity in sorghum grains. J. Near Infrared Spectrosc., 1–8.

Boulos, N.N., Grenfield, H., Wills, R.B.H., 2000. Water holding capacity of some selected soluble and insoluble dietary fibre. Int. J. Food Prop. 3 (2), 217–231.

Dabulici, C.M., Sârbu, I., Vamanu, E., 2015. The bioactive potential of functional products and bioavailability of phenolic compounds. Foods 9, 953.

De Ancos, B., Ibanez, E., Regiero, G., Cano, M.P., 2000. Frozen storage effects on anthocyanins and volatile compounds of raspberry fruit. J. Agric. Food Chem. 48, 873–879.

Ebere, C.O., Emelike, N.J.T., Kiin-Kabar, D.B., 2015. Physico-chemical and sensory properties of cookies prepared from wheat flour and cashew-apple residue as a source of fibre. Asian J. Agric. Food Sci. 3 (2), 213–218.

Guenther, N., Schonlau, M., 2016. Support vector machines. Stata J. 16 (4), 917.

Ionnou, I., Hafsa, S., Hamdi, S., Charbonnel, C., Ghoul, M., 2012. Review of the effect of food processing and formulation on flavonol and anthocyanin bahaviour. J. Food Eng. 111, 208–217.

Karasu, S., Akman, S., Perihan, A., Goktas, H., Doymaz, I., Sagdıc, O., 2019. Effect of different drying methods on total bioactive compounds, phenolic profile, *in vitro* bioaccessibility of phenolic and HMF formation of persimmon. LWT - Food Sci. Technol.

Lin, C., Li, T., Chen, S., Liu, X., Lin, C., Liang, C., 2019. Gaussian process regression-based forecasting model of dam deformation. Neural Comput. Appl. 31, 8503–8518.

Liu, K., Liu, B., Xu, C., 2009. Intelligent analysis model of slope nonlinear displacement time series based on genetic-Gaussian process regression algorithm of combined kernel function. Chin. J. Rock Mech. Eng. 28 (10), 2128–2134.

Luka, B.S., Japhet, J.A., Dauda, P.A., 2018. Single layer drying characteristics of hospital too far leaves (jatropha tanjorensis) under open sun and in solar dryer. Acta Tech. Corviniensis, Bull. Eng. 4, 83–88.

Luka, B.S., Ejilah, R.I., Chisa, S.O., Japhet, J.A., Ibrahim, T.K., Udom, P.U., 2020. Effect of diesel fuel blend on flame and emission characteristics of used engine oil as heating fuel using swirl waste oil burner. Environ. Clim. Technol. 24 (1), 545–561.

Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., de Mendonca, A., 2011. Data mining methods in the prediction of Dementia: a real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. BMC Res. Notes 4, 299.

Martincic, R., Kuzmanovski, I., Wagner, A., Novič, M., 2015. Development of models for prediction of the antioxidant activity of derivatives of natural compounds. Anal. Chim. Acta.

Mrad, N.D.N., Bhoudhrioua, N., Kechaou, N., Courtois, F., Bonazzi, C., 2012. Influence of air drying temperature on kinetics, physicochemical properties, total phenolic content and ascorbic acid of pears. Food Bioprod. Process. 90, 433–441.

Nguyen, V.T., Le, D.M., 2018. Influence of various drying conditions on phytochemical compounds and antioxidant capacity of carrot peel. Beverages 4 (80), 1–12.

Preethi, P., Mangalassery, S., Shradha, K., Pandiselvam, R., Manikantan, M.R., Reddy, S.V.R., Devi, S.R., Nayak, M.G., 2021. Cashew apple pomace powder enriched the proximate, mineral, functional and structural properties of cereal based extrudates. LWT - Food Sci. Technol. 139, 110536.

Qu, W.Z., Ma, H., 2010. Extraction, modeling and activities of antioxidants from pomegranate marc. J. Food Eng. 99 (1), 16–23.

Que, F., Mao, L., Fang, X., Wu, T., 2008. Comparison of hot air drying and freeze drying on the physicochemical properties and antioxidant activities of pumpkin (*cucurbita moschata Duch.*) flours. Int. J. Food Sci. Technol. 43 (7), 195–1201.

Rodríguez, K., Ah-Hen, K.S., Galvez, A.V., Vasquez, V., Quispe-Fuentes, I., Rojas, P., Lemus-Mondaca, R., 2016. Changes in bioactive components and antioxidant capacity of maqui, Aristotelia chilensis [Mol] Stuntz, berries during drying. LWT - Food Sci. Technol. 65, 537–542.

Sahni, P., Shere, D.M., 2018. Utilization of fruit and vegetable pomace as functional ingredient in bakery product. Asian J. Diary & Food Res. 37 (3), 202–211.

Sato, M.F., Rigoni, D.C., Canteri, M.H.G., Petkowicz, G.L.O., Nogueira, A., Wosiacki, G., 2011. Chemical and instrumental characterization of pectin from dried pomace of eleven apple cultivars. Acta Sci. Agron. 33 (3), 383–389.

Shekarau, B.L., Zakka, R., Tsokwa, T., Yuguda, T.K., Okon, P.U., 2020. Mathematical modelling of thin layer drying kinetics of cashew apple pomace in hot air oven dryer. Acta Period. Technol. 51, 119–136.

Sim, K.Y., Liew, J.Y., Ding, X.Y., Chddng, W.S., Intan, S., 2017. Effect of vacuum and oven drying on the radical scavenging activity and nutritional contents of submerged fermented Maitake (Grifola frondosa) mycelia. Food Sci. Technol. 37 (1), 131–135.

Sumic, Z., Tepic, A., Vidovic, S., Vakula, A., Vladic, J., Pavlić, B., 2017. Process optimization of chanterelle (Cantharellus cibarius) mushrooms vacuum drying. J. Food Process. Preserv. 5 (5), 989–996.

Tańska, M., Roszkowska, B., Czaplicki, S., Borowska, E.J., Bojarska, J., Dąbrowska, A., 2016. Effect of fruit pomace addition on shortbread cookies to improve their physical and nutritional values. Plant Foods Hum. Nutr. 71, 307–313.

Wenqi, F., Shitia, Z., Hui, H., Shaobo, D., Zhejun, H., Wenfei, L., Zheng, W., Tianfu, S., Huiyun, L., 2018. Learn to make decision with small data for autonomous driving: deep gaussian process and feedback control. J. Adv. Transp., 8495264.

Xu, G., Ye, X., Chen, J., Liu, D., 2007. Effect of heat treatment on the phenolic compounds and antioxidant capacity of citrus peel extract. J. Agric. Food Chem. 55 (2), 330–335.

Yan, H., Kerr, W.L., 2012. Total phenolic content, anthocyanins, and dietary fibre content of apple pomace powders produced by vacuum-belt drying. J. Sci. Food Agric. 93 (6), 1499–1505.

Youssef, K.M., Mokhtar, S.M., 2014. Effect of drying methods on the antioxidant capacity, color and phytochemicals of portulaca oleracea l. leaves. J. Nutr. Food Sci. 4, 322.

Zhang, L., Zhang, Z., Luo, O., Lu, H., Liang, Y., 2014. Evaluation and prediction of the antioxidant activity of epimedium from multi-wavelength chromatographic fingerprints and chemometrics. Anal. Methods 6, 10–36.