Acta Scientific NUTRITIONAL HEALTH (ISSN:2582-1423)

Volume 8 Issue 7 July 2024

Research Article

Predicting Peroxide Value of Peanut Oil using Machine Learning Models

Yuan Ting CHEN, Maurice HT LING*, Hong YANG, Martin Hui CAI, Rui Wen KOH, Rick YH TAN and Xuejia XUE*

School of Applied Science, Temasek Polytechnic, Singapore

***Corresponding Author:** Maurice HT Ling and Xuejia Xue, School of Applied Science, Temasek Polytechnic, Singapore.

DOI: 10.31080/ASNH.2024.08.1403

Received: June 04, 2024

Published: June 30, 2024

Abstract

Natural antioxidants (NATOs) derived from sources like rosemary, green tea, and oregano have acquired extensive attention for their efficacy in preserving edible oils, presenting a promising alternative to synthetic antioxidants (ATOs) due to their superior safety profile. However, integrating NATOs into the food industry faces challenges stemming from the variability in their chemical composition, necessitating prolonged stability tests based on peroxide values (PV). This study explores the predictability of PV in peanut oil using three chemical parameters (total phenolic content, total antioxidant content, and total carotenoid content), one physical parameter (partition coefficient), and storage duration. Six machine learning classifiers (logistic regression, multilayer perceptron, radial basis function, Gaussian Naïve Bayes classifier, support vector machine, and decision tree) were employed. The results have shown significant correlations between the chemical parameters and antioxidant activity. Our findings indicate that PV in peanut oil can be accurately predicted using these parameters and storage duration, with the multilayer perceptron demonstrating the highest predictive performance, achieving an accuracy of at least 89.8% in determining whether PV remains within acceptable limits post-storage.

Keywords: Peanut Oil; Shelf-Life; Peroxide Value; Antioxidants; Natural Antioxidants; Machine Learning Classifiers; Artificial Neural Networks; Multilayer Perceptron

Introduction

Oil deteriorates over storage time. Oxidation, including auto-oxidation, is the main cause of deterioration [1] and can be measured using peroxide value (PV) as indicator of freshness [2]. Antioxidants (ATOs) are widely used to extend shelf-life of oil [3]. Tocopherol (vitamin E); together with synthetic ATOs [4]; such as BHT (butylated hydroxytoluene), BHA (butylated hydroxyanisole), and TBHQ (tert-butylhydroquinone); are standard antioxidant regimen adopted in food industry [5] but long-term exposure to these synthetic ATOs has raised concerns, leading to the pursuit for natural alternatives [6, 7]. Natural ATOs (NATOs); such as rosemary, green tea and oregano; have been examined for their protective effects on foods [8-10]. Their comparable efficacy to synthetic ATOs and much safer health profile make them viable alternative food preservatives [11].

Mixtures of NATOs may be synergistic, leading to bringing down the cost [12]. However, finding the right blend of NATOs has been tedious and expensive, typically through trials and errors [13]. Machine learning (ML) has been shown to be able to pick up patterns from the input data to make reliable predictions for future/unknown events [14-17]; hence, may be useful in identifying optimal mixtures of NATOs or in predicting the usability of oil after storage via PV.

In this study, we examined the predictability of PV of peanut oil from 3 chemical parameters [(a) TPC (total phenolic content), (b) TAC (total antioxidant content), and (c) TCC (total carotenoid content)], 1 physical parameter (logP, also known as partition coefficient, which can be used to estimate the solubility of NATOs in peanut oil), and storage days; using 6 machine learning classifiers (MLCs) [(a) logistic regression (LR), (b) multilayer perceptron (MLP), (c) radial basis function (RBF), (d) Gaussian Naïve Bayes classifier (GNB), (e) support vector machine (SVM), and (f) decision tree (DT)]. TPC, TAC, and TCC have been suggested to be significantly correlated to antioxidant activity [18]. Our results suggest that PV of peanut oil can be predicted by TPC, TAC, TCC, logP, and number of days in storage; and MLP is the best performing MLC to predict whether PV will be within acceptable limits after storage with accuracy ≥ 89.8%.

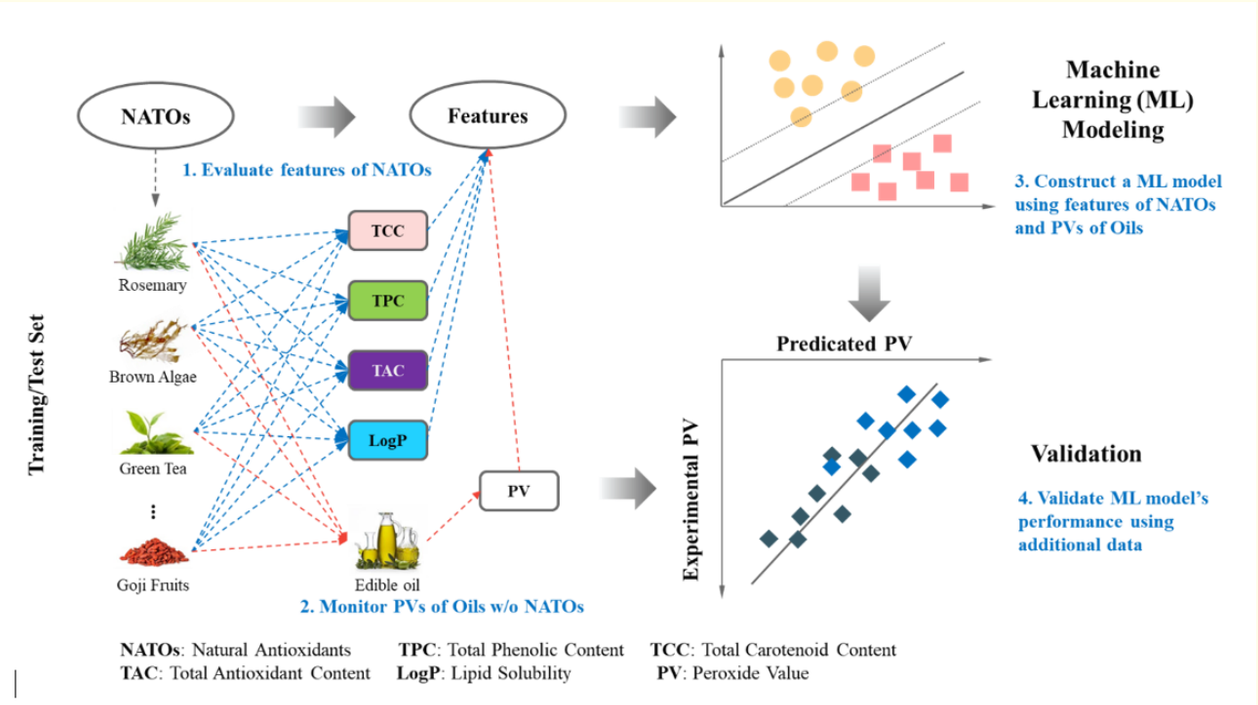

Figure 1: The sequential steps from the selection of NATOs and the machine learning modelling.

Chemical Materials and Methods

- **Selection of NATOs:** Synthetic antioxidants tocopherol and ascorbyl palmitate were employed as the reference control in this study. In addition, 14 natural antioxidant extracts along with their paired combinations, resulting in a total of 136 permutations had been tested. These extracts (Table 1) included green tea extracts (extracted with ethanol and water), rosemary extracts (extracted with ethanol and water), oregano extracts (extracted with ethanol and water), kelp extracts, grapeseed extracts (extracted with ethanol), grapeseed capsules, blueberry extracts (extracted with ethanol), cranberry extracts (extracted with ethanol), goji berry (extracted with water), pine bark extracts (extracted with ethanol), and turmeric extracts (extracted with ethanol).
- **Chemical testing methods:** To examine their properties, each extract (or combination) will be dissolved in a solution of octanol and water in a 1:1 ratio (Figure 2). Subsequently, the layers of water and octanol will be separated for further analysis. The LogP value will be determined using the water layer, while the octanol layer will be utilized to measure TAC, TCC, and TPC. The logP will be directly determined using our own method. The total antioxidant activity will be assessed using the DPPH (1,1-diphenyl-2-picryl-hydrazyl-hydrate) assay, expressed as mmol Trolox equivalent per 100 grams of the samples. Folin-Ciocalteu's reagent will be employed to assess TPC, expressed as mg of gallic acid equivalent per gram of sample. TCC will be directly measured using a microplate reader, expressed as mg/kg.
- **Preparation of edible oil and peroxide value testing:** The peroxide value serves as an indicator reflecting the extent of lipid peroxidation within a sample, indicative of the formation of peroxides due to oxidation, which can result in rancidity and a decline in quality. Higher peroxide values signify heightened lipid peroxidation and diminished sample quality. In this study, freshly pressed peanut oil was acquired utilizing a tabletop oil press machine. Peanuts were placed in the upper container of the machine, where they were ground and pressed to extract their oil content, with residual peanut material collected aside. The freshly pressed peanut oil underwent centrifugation to eliminate any solid peanut particles. To evaluate the PV values of the oil samples, each extract (or combination) was dissolved in a 1:1 octanol/water solution. Subsequently, certain amounts of the octanol layer, with or without extracts, was blended with peanut oil. These treated peanut oil samples were subjected to an accelerated stability test in an oven set at 55°C to assess their stability under elevated temperatures. The changing peroxide values were continuously monitored over time, with assessments conducted every 3 to 4 days, offering a detailed and comprehensive overview of the oxidation process of edible oil. The ferric thiocyanate method was em-

ployed to determine the PV values (expressed in milliequivalents of peroxide per kilogram of sample).

Machine learning methods

- **Regression analysis:** Ordinary linear regression (OLR) [19] was used to evaluate the predictability of PV after a period of storage from TPC, TAC, TCC, logP, and storage days; using the following model: $PV = \beta_1 TPC + \beta_2 TAC + \beta_3 TCC + \beta_4 logP + \beta_5 Days + \beta_0$. OLR was also used to evaluate the predictability of each of TPC, TAC, TCC, and logP from NATOs using the following model: $\{ TPC \mid TAC \mid TCC \mid logP \} = \beta_1 AP + \beta_2 TO + \beta_3 GTW + \beta_4 GTE + \beta_5 RMW + \beta_6 RME + \beta_7 ORW + \beta_8 ORE + \beta_9 KP + \beta_{10} GSC + \beta_{11} GSE + \beta_{12} BBE + \beta_{13} CBE + \beta_{14} GJW + \beta_{15} PBE + \beta_{16} TME + \beta_0$, where the independent t variables are mass of NATOs. By merging the 2 sets of OLRs, the mass of NATOs and storage days were used to predict the resulting PV after a period of storage.
- **Machine learning classification of PV thresholds:** TPC, TAC, TCC, logP, and storage days were used to classify whether the PV will be acceptable after a number of days in storage. The numerical values of TPC, TAC, TCC, and logP were categorized to decile scores [20,21] where values lower than the 10th percentile were scored as 1. Hence, values from and including 90th percentile to less than 100th percentile were scored as 10, whereas values at 100th percentile were scored as 11. This allows for scoring of potential values above the experimental values obtained in this study as 11 or 100th percentile and above. PV values were binary coded to whether the value was below a specific percentile of PV values - "1" as below a specific percentile of PV value and "0" as at or above a specific percentile of PV value. The percentile thresholds were from 10% to 90%, at 10% increment; resulting in 9 thresholds (PV10 to PV90). The coded data were classified using 6 MLCs [(a) logistic regression (LR), (b) multilayer perceptron (MLP), (c) radial basis function (RBF), (d) Gaussian Naïve Bayes classifier (GNB), (e) support vector machine (SVM), and (f) decision tree (DT)]. LR [22], MLP [23], and RBF [24] from IBM SPSS were used. GNB [25], SVM [26], and DT [27] from BactClass [28] were used.

Results and Discussion

Natural antioxidants can predict peroxide value: OLR analysis shows PV of oil after a period of storage can be predicted by TPC, TAC, TCC, logP, and storage days ($r^2 = 0.774$; $F_{2,2658} = 1817.947$, p-value < 1e-100) as PV = (-0.0616TPC) + (-0.0428TAC) + (0.00366TCC) + (-1.758logP) + (2.806Days) + 4.230. This result is plausible as TPC [29] and TCC [30] are correlated to PV, and TAC has good correlation to TPC [31]. Similarly, TPC ($r^2 = 0.978$; $F_{16,2647} = 13294.231$, p-value < 1e-100), TAC ($r^2 = 0.936$; $F_{16,2647} = 2433.939$, p-value < 1e-100), TCC ($r^2 = 0.942$; $F_{16,2647} = 2697.836$, p-value < 1e-100), and logP ($r^2 = 0.790$; $F_{16,2647} = 623.039$, p-value < 1e-100) can be predicted independently from NATOs using the coefficients listed in Table 2.

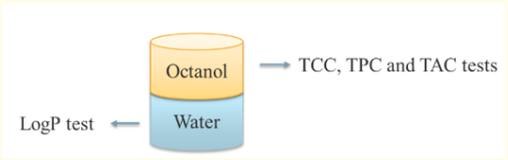

Figure 2: Experimental determination of LogP, TCC, TPC and TAC.

Given that all the regression models are highly significant (p-value < 1e-100), PV after a period of storage is predictable from mass of NATOs and storage days ($r^2 = 0.833$; $F_{17,2646} = 778.602$, p-value < 1e-100) with the following linear regression model: PV = (-1.680AP) + (-0.095TO) + (-0.797GTW) + (-1.955GTE) + (-0.301RMW) + (-1.622RME) + (0.087ORW) + (-0.079ORE) + (0.171KP) + (-0.072GSC) + (0.035GSE) + (-0.099BBE) + (-0.140CBE) + (-0.044GJW) + (-0.172PBE) + (0.405TME) + (2.809Days) + 10.799. The regression coefficients represent the efficiency of NATOs in reducing PV; hence, regression analysis suggests that GTE (green tea ethanol) is most effective in reducing PV while TME (turmeric ethanol) is least in reducing PV of peanut oil.

PV thresholds can be predicted using machine learning algorithms. Given that PV is often used as an indicator of oil freshness [2, 32], a practical question to ask will be whether peanut oil is within acceptable PV given the number of days in storage and NATO parameters (TPC, TCC, TAC, and logP). As the PV and NATO parameters are coded as decile scores, Table 3 shows the correspondence of value ranges to decile scores.

As PV values were coded as deciles from 1st decile (corresponding to 10th percentile) to 9th decile (corresponding to 90th percentile), the upper and lower boundaries of random prediction accuracy can be formed by assuming all PV values to be within range. For example, if all PV values of 8th decile (corresponding to 80th percentile) are assumed to be above the 8th decile, then the accuracy will be 80% since only 80% of the PV values are above 8th decile. Conversely, all PV values of 8th decile are assumed to be below the 8th decile, then the accuracy will be 20% since only 20% of the PV values are below 8th decile. Hence, this 80% / 20% accuracy form the upper and lower boundary of random prediction accuracy can be established for 8th decile of PV values. Thus, any classifier that perform above 80% accuracy will be deemed to perform better than random; therefore, able to predict PV thresholds given the number of days in storage and NATO parameters.

No.	Extracts	Abbreviation	No.	Extracts	Abbreviation
(a)	Ascorbyl Palmitate	AP	(i)	Kelp	KP
(b)	Mixed To-copherol	TO	(j)	Grapeseed Capsule	GSC
(c)	Green Tea (Water)	GTW	(k)	Grapeseed (Ethanol)	GSE
(d)	Green Tea (Ethanol)	GTE	(l)	Blueberry (Ethanol)	BBE
(e)	Rosemary (Water)	RMW	(m)	Cranberry (Ethanol)	CBE
(f)	Rosemary (Ethanol)	RME	(n)	Goji (Water)	GJW
(g)	Oregano (Water)	ORW	(o)	Pine Bark (Ethanol)	PBE
(h)	Oregano (Ethanol)	ORE	(p)	Turmeric (Ethanol)	TME

Table 1: The abbreviation table for the project.

Independent Variables	Coefficients			
	TPC Model	TAC Model	TCC Model	Log P Model
Intercept (β_0)	103.253	69.789	398.517	-2.153
AP	-1.991	5.590	-10.322	0.146
TO	-3.269	7.387	-11.351	0.138
GTW	2.227	4.371	-12.183	0.019
GTE	12.262	12.501	-8.892	0.060
RMW	-3.082	-2.657	-11.647	0.015
RME	12.569	13.992	1.386	0.062
ORW	-2.894	-2.090	-11.855	0.018
ORE	-3.001	-2.178	-10.152	0.000
KP	-3.330	-2.185	-11.838	-0.036
GSC	-2.334	-1.572	-10.690	0.066
GSE	-2.501	-1.456	-9.547	-0.023
BBE	-3.004	-2.146	-10.255	-0.028
CBE	-2.862	-2.097	-9.888	-0.023
GJW	-3.218	-2.250	-11.987	-0.015
PBE	-2.893	-1.847	-8.390	-0.003
TME	-1.627	1.726	245.696	0.145

Table 2: Regression coefficients of models to predict TPC, TAC, TCC, and log P from NATOs.

Percentile Range	Decile Score	TPC Range	TCC Range	TAC Range	Log P Range	PV Range
[0, 10)	1	[0.00, 7.11)	[0.00, 0.65)	[0.00, 0.24)	< -3.16)	< -0.32
[10, 20)	2	[7.11, 10.51)	[0.65, 9.60)	[0.24, 2.82)	No range	[-0.32, 2.06)
[20, 30)	3	[10.51, 14.18)	[9.60, 22.21)	[2.82, 4.36)	[-3.16, -3.00)	[2.06, 5.91)
[30, 40)	4	[14.18, 21.51)	[22.21, 31.48)	[4.36, 10.92)	[-3.00, -1.92)	[5.91, 11.20)
[40, 50)	5	[21.51, 29.94)	[31.48, 39.49)	[10.92, 80.15)	[-1.92, -1.34)	[11.20, 20.01)
[50, 60)	6	[29.94, 36.09)	[39.49, 48.07)	[80.15, 113.72)	[-1.34, -1.07)	[20.01, 29.62)
[60, 70)	7	[36.09, 81.06)	[48.07, 57.69)	[113.72, 130.08)	[-1.07, -0.33)	[29.62, 43.03)
[70, 80)	8	[81.06, 228.65)	[57.69, 79.54)	[130.08, 165.51)	[-0.33, -0.24)	[43.03, 56.91)
[80, 90)	9	[228.65, 252.33)	[79.54, 224.45)	[165.52, 203.87)	[-0.24, -0.07)	[56.91, 71.79)
[90, 100)	10	[252.33, 538.88)	[224.45, 4424.31)	[203.87, 252.54)	[-0.07, 0.57)	[71.79, 88.81)
≥ 100	11	≥ 538.33	≥ 4596.69	≥ 544.27	≥ 3.00	≥ 150.59

Table 3: Binning of TPC, TCC, TAC, log P, and PV values into decile scores for classification. The range [x, y) represents that the range is between x and y, including x but excluding y; which can also be represented as x ≤ value < y.

Threshold	DT	GNB	SVM	LR	MLP	RBF
PV10	55.8 (8.13)	90.0 (0.04)	90.0 (0.04)	92.0	93.7	92.0
PV20	51.0 (8.75)	78.5 (1.62)	80.0 (0.04)	90.8	95.8	89.9
PV30	46.7 (7.58)	65.8 (2.48)	70.0 (0.02)	91.5	96.1	87.6
PV40	54.5 (4.15)	65.1 (3.24)	66.0 (3.06)	89.8	94.2	85.9
PV50	47.0 (6.46)	62.1 (4.28)	65.7 (3.90)	90.4	93.2	87.8
PV60	54.8 (5.74)	62.6 (4.13)	63.6 (3.13)	90.3	94.7	88.7
PV70	58.9 (6.11)	66.4 (3.08)	70.0 (0.02)	90.0	93.8	86.8
PV80	70.6 (4.11)	72.4 (2.71)	80.0 (0.04)	92.2	94.8	85.7
PV90	73.3 (7.20)	90.0 (0.04)	90.0 (0.04)	92.8	95.6	89.4

Table 4: Mean accuracy of classifiers on PV thresholds.Values in backets are standard errors from 5-fold cross-validation, if given by the system. Green coloured cells represent accuracies above upper boundary of random prediction. Orange coloured cells represent accuracies below lower boundary of random prediction.

Based on this argument, our classification results show that LR, MLP, and RBF perform better than random for all nine PV thresholds while GNB, and SVM perform better than random for three PV thresholds (PV40, PV50, and PV60); LR consistently outperforms DT, GNB, and SVM; and DT performs worse than random for PV50 (Table 4). Our results are consistent with a study suggesting that LR can outperform other MLCs, including SVM and artificial neural networks (ANNs), in classifying diabetes patients [33]. MLP and RBF [34] are different types of ANNs but our results show that MLP outperforms LR average of (MLP - LR) = 3.6% with a standard error of 0.38%; paired t-test p-value = 0.0000125 but not RBF average of (LR - RBF) = 2.9% with a standard error of 0.64%; paired t-test p-value = 0.00197; hence, suggesting that LR may outperform only specific ANNs.

Our results also show that ANN outperforms SVM but this cannot be generalized as one study [35] suggests that ANN outperforms SVM while another study [36] suggests that SVM outperforms ANN. Similarly, contradictory results from studies comparing LR and ANN can be found [37, 38]. Naïve Bayes classifiers, of which GNB is an example, have also been expected to perform better than DT [39] in certain cases [40]. Similarly, an earlier study also suggests that LR can outperform DT [41] but a recent study suggests that an ensemble of DTs may outperform LR [42]. Taken together, this suggests that the performance of various MLCs cannot be generalized and is highly dependent on situation. As such, methods to combine different MLCs have emerged [43-45]. Nevertheless, our results suggest that PV thresholds can be predicted using MLCs, and MLP is the best performing MLC to predict whether PV will be within acceptable limits after storage with accuracy ≥ 89.8%.

Conclusion

PV values of peanut oil can be predicted by four chemical and physical parameters of NATOs (TPC, TAC, TCC, and logP), and number of days in storage, and MLP is the best performing MLC to predict whether PV will be within acceptable limits after storage with accuracy ≥ 89.8%.

Supplementary Materials

Data file for this work can be downloaded at https://bit.ly/NATO_Oil.

Acknowledgement

This work is supported by Ministry of Education, Singapore (Grant No. MOE2021-TIF-0012).

Conflict of Interest

The authors declare no conflict of interest.

Bibliography

1. Sanmartin C., *et al*. "The Effects of Packaging and Storage Temperature on the Shelf-Life of Extra Virgin Olive Oil". *Heliyon* 4.11 (2018): e00888.

2. Gilbraith WE., *et al*. "Improving Prediction of Peroxide Value of Edible Oils Using Regularized Regression Models". *Molecules* 26.23 (2021): 7281.

3. Petcu CD., *et al*. "Harnessing Natural Antioxidants for Enhancing Food Shelf Life: Exploring Sources and Applications in the Food Industry". *Foods* 12.17 (2023): 3176.

4. Fadda A., *et al*. "Innovative and Sustainable Technologies to Enhance the Oxidative Stability of Vegetable Oils". *Sustainability* 14.2 (2022): 849.

5. Choi H., *et al*. "NPCARE: Database of Natural Products and Fractional Extracts for Cancer Regulation". *Journal of Cheminformatics* 9.1 (2017): 2.

6. Arabsorkhi B., *et al*. "Investigating the Effect of Replacing the Antioxidants Ascorbyl Palmitate and Tocopherol Instead of TBHQ on the Shelf Life of Sunflower Oil Using Temperature Accelerated Method". *Food Chemistry Advances* 2 (2023): 100246.

7. Shebis Y., *et al*. "Natural Antioxidants: Function and Sources". *Food and Nutrition Sciences* 04.06 (2013): 643-649.

8. Awad AM., *et al*. "Green Extraction of Bioactive Compounds from Plant Biomass and Their Application in Meat as Natural Antioxidant". *Antioxidants* 10.9 (2021): 1465.

9. Kumar Y., *et al*. "Recent Trends in the Use of Natural Antioxidants for Meat and Meat Products". *Comprehensive Reviews in Food Science and Food Safety* 14.6 (2015): 796-812.

10. Reddy DM., *et al*. "Application of Natural Antioxidants in Meat and Meat Products - A Review". *Food and Nutrition Journal* 7.3 (2018).

11. Karpińska M., *et al*. "The Use of Natural Antioxidants in Ready-to-Serve Food". *Food Chemistry* 72.1 (2001): 5-9.

12. Colon M and Nerín C. "Synergistic, Antagonistic and Additive Interactions of Green Tea Polyphenols". *European Food Research and Technology* 242.2 (2016): 211-220.

13. Pokorný J. "Natural Antioxidants for Food Use". *Trends in Food Science and Technology* 2 (1991): 223-227.

14. Qi D and Majda AJ. "Using Machine Learning to Predict Extreme Events in Complex Systems". *Proceedings of the National Academy of Sciences* 117.1 (2020): 52-59.

15. Saldana DA., *et al*. "Prediction of Flash Points for Fuel Mixtures Using Machine Learning and a Novel Equation". *Energy and Fuels* 27.7 (2013): 3811-3820.

16. Steinberg F., *et al*. "A Novel Machine Learning Model for Predicting Late Supplier Deliveries of Low-Volume-High-Variety Products with Application in a German Machinery Industry". *Supply Chain Analytics* 1 (2023): 100003.

17. Yang Q., *et al*. "Holistic Prediction of the pKa in Diverse Solvents Based on a Machine-Learning Approach". *Angewandte Chemie International Edition* 59.43 (2020): 19282-19291.

18. So V., *et al*. "Bioactive Compounds, Antioxidant Activities, and HPLC Analysis of Nine Edible Sprouts in Cambodia". *Molecules* 28.6 (2023): 2874.

19. Galton F. "Regression Towards Mediocrity in Hereditary Stature". *The Journal of the Anthropological Institute of Great Britain and Ireland* 15 (1886): 246-263.

20. Fonseca A., *et al*. "Antibody Selection Strategies and Their Impact in Predicting Clinical Malaria Based on Multi-Sera Data". *BioData Mining* 17.1 (2024): 2.

21. Zhang R., *et al*. "The Association of Retinopathy and Plasma Glucose and HbA1c: A Validation of Diabetes Diagnostic Criteria in a Chinese Population". *Journal of Diabetes Research* (2016): 4034129.

22. Cox DR. "The Regression Analysis of Binary Sequences". *Journal of the Royal Statistical Society: Series B (Methodological)* 20.2 (1958): 215-232.

23. Haykin S. "Neural Networks: A Comprehensive Foundation (Prentice Hall PTR)" (1998).

24. Broomhead D and Lowe D. "Multivariable Functional Interpolation and Adaptive Networks". *Complex Systems* 2 (1988): 321-355.

25. Anand MV., *et al*. "Gaussian Naïve Bayes Algorithm: A Reliable Technique Involved in the Assortment of the Segregation in Cancer". *Mobile Information Systems* (2022): 1-7.

26. Cortes C and Vapnik V. "Support-Vector Networks". *Machine Learning* 20.3 (1995): 273-297.

27. Wu X., *et al*. "Top 10 Algorithms in Data Mining". *Knowledge and Information Systems* 14.1 (2008): 1-37.

28. Liu TT and Ling MH. "BactClass: Simplifying the Use of Machine Learning in Biology and Medicine". *Acta Scientific Medical Sciences* 4.11 (2020): 43-47.

29. Grossi M., *et al*. "Field-Deployable Determinations of Peroxide Index and Total Phenolic Content in Olive Oil Using a Promising Portable Sensor System". *Sensors* 23.11 (2023): 5002.

30. Liang J., *et al*. "Reduction of Chlorophyll in Cold-Pressed Hemp (Cannabis sativa) Seed Oil by Ultrasonic Bleaching and Enhancement of Oxidative Stability". *European Journal of Lipid Science and Technology* 120.4 (2018): 1700349.

31. Matei PL., *et al*. "Ultrasound-Assisted Extraction of Blackberry Seed Oil: Optimization and Oil Characterization". *Molecules* 28.6 (2023): 2486.

32. Zhang N., *et al*. "Analytical Methods for Determining the Peroxide Value of Edible Oils: A Mini-Review". *Food Chemistry* 358 (2021): 129834.

33. Lynam AL., *et al*. "Logistic Regression has Similar Performance to Optimised Machine Learning Algorithms in a Clinical Setting: Application to the Discrimination between Type 1 and Type 2 Diabetes in Young Adults". *Diagnostic and Prognostic Research* 4 (2020): 6.

34. Schwenker F., *et al*. "Three Learning Phases for Radial-Basis-Function Networks". *Neural Networks* 14.4-5 (2001): 439-458.

35. Ren J. "ANN vs. SVM: Which One Performs Better in Classification of MCCs in Mammogram Imaging". *Knowledge-Based Systems* 26 (2012): 144-153.

36. Bisgin H., *et al*. "Comparing SVM and ANN based Machine Learning Methods for Species Identification of Food Contaminating Beetles". *Scientific Reports* 8.1 (2018): 6532.

37. Eftekhar B., *et al*. "Comparison of Artificial Neural Network and Logistic Regression Models for Prediction of Mortality in Head Trauma Based on Initial Clinical Data". *BMC Medical Informatics and Decision Making* 5.1 (2005): 3.

38. Kirişci M. "Comparison of Artificial Neural Network and Logistic Regression Model for Factors Affecting Birth Weight". *SN Applied Sciences* 1.4 (2019): 378.

39. Liang H and Yan Y. "Learning Naïve Bayes Tree for Conditional Probability Estimation". *Advances in Artificial Intelligence*, Lecture Notes in Computer Science., eds Tawfik AY, Goodwin SD (Springer Berlin Heidelberg, Berlin, Heidelberg) 3060 (2006): 455-466.

40. Ashari A., *et al*. "Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool". *International Journal of Advanced Computer Science and Applications* 4.11 (2013).

41. Long WJ., *et al*. "A Comparison of Logistic Regression to Decision-Tree Induction in a Medical Domain". *Computers and Biomedical Research* 26.1 (1993): 74-97.

42. Seto H., *et al*. "Gradient Boosting Decision Tree Becomes More Reliable than Logistic Regression in Predicting Probability for Diabetes with Big Data". *Scientific Reports* 12.1 (2022): 15889.

43. Atacak İ. "An Ensemble Approach Based on Fuzzy Logic Using Machine Learning Classifiers for Android Malware Detection". *Applied Sciences* 13.3 (2023): 1484.

44. Feng W., *et al*. "An Ensemble Machine Learning Approach for Classification Tasks Using Feature Generation". *Connection Science* 35.1 (2023): 2231168.

45. Tzougas G and Kutzkov K. "Enhancing Logistic Regression Using Neural Networks for Classification in Actuarial Learning". *Algorithms* 16.2 (2023): 99.