Use of machine learning algorithms to estimate the phenolic compounds and antioxidant activity of honey based on colour parameters

A.K. Zaldivar-Ortega[1], J.C. Angeles-Hernandez[2], N. Esturau-Escofet[3], M. Jiménez Guarneros[4], A. M. Mier y Teran Lugo[3], P. A. Vázquez-Landaverde[5]

[1]Instituto de Ciencias Agropecuarias , Universidad Autónoma del Estado de Hidalgo, Tulancingo, Hidalgo, México.
[2]Departamento de Medicina y Zootecnia de Rumiantes, Facultad de Medicina Veterinaria y Zootecnia. UNAM México.
[3]Instituto de Química, UNAM 04510 CDMX, Mexico.
[4]Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM, 04510 CDMX, Mexico.
[5]Centro de Investigación en Ciencias Aplicadas y Tecnología Avanzada, IPN, 76090 Qro., Mexico. Corresponding author: juan_angeles@uaeh.edu.mx

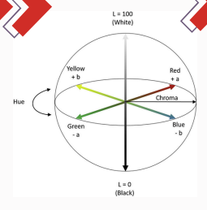
Introduction

The flavonoids and phenolic acids found in honey are derived from nectar, pollen and propolis, and contribute to its nutraceutical properties. These bioactive compounds are key to the antioxidant activity of honey, which helps mitigate oxidative reactions beneficial to human health. This study aims to estimate the phenolic compound content and antioxidant activity of honey based on colour using machine learning algorithms.

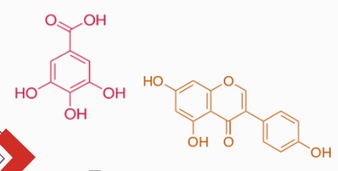
Materials and Methods

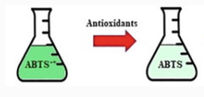
n=290

Predictor



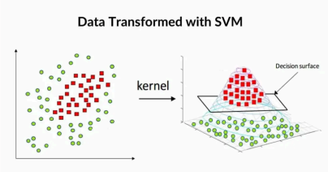
Response

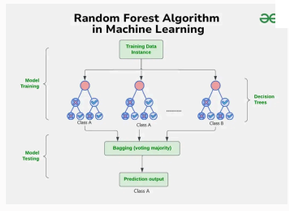


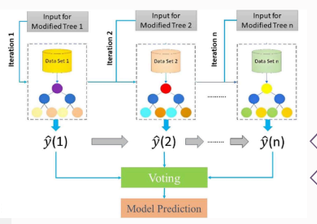


ABTS

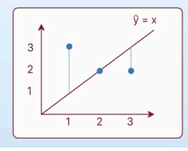
DPPH







Mean Squared Error



Symmetric Mean Absolute Percentage Error

$$\text{SMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{(|\hat{y}_i| + |y_i|)/2} \times 100$$

Results

Table 1. prediction of phenolic compounds in honey.

Method	Color variables(L, a, b)	
	SME	SMAPE (%)
SVR – RBF	0.0687	17.42
SVR – Polinomial	0.0048	16.02
XGBoost	0.0183	9.62
Random Forest	0.0087	5.87

SVR= Support Vector Regression. RBF=Radial Basis Function Kernel. XGBoost= eXtreme Gradient Boosting. SME= Square Mean Error. SMAPE= Symmetric Mean Absolute Percentage Error.

Conclusion

The results showed a strong relationship between honey colour, phenolic content, and antioxidant activity, allowing accurate estimations, especially using the Random Forest algorithm. This approach improves the accessibility of measuring these beneficial compounds at the farm level, benefiting both honey producers and consumers.

References