BIBLIOGRAPHY

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(1), 97–106. https://doi.org/10.1002/wics.51

Abedi, M., Norouzi, G. H., & Bahroudi, A. (2012). Support vector machine for multi-classification of mineral prospectivity areas. *Computers and Geosciences*, *46*, 272–283. https://doi.org/10.1016/j.cageo.2011.12.014

Afoakwa, E. O., Paterson, A., Fowler, M., & Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition*, *48*(9), 840–857. https://doi.org/10.1080/10408390701719272

Ajani, R., Oboh, G., Adefegha, S. A., Nwokocha, K. E., & Akindahunsi, A. A. (2020). Sensory attributes, nutritional qualities, and glycemic indices of bread blends produced from cocoa powder flavored yellow-fleshed cassava-wheat composite flours. *Journal of Food Processing and Preservation*, *44*(9), e14673. https://doi.org/10.1111/JFPP.14673

Albertini, B., Schoubben, A., Guarnaccia, D., Pinelli, F., della Vecchia, M., Ricci, M., di Renzo, G. C., & Blasi, P. (2015). Effect of Fermentation and Drying on Cocoa Polyphenols. *Journal of Agricultural and Food Chemistry*, *63*(45), 9948–9953. https://doi.org/10.1021/acs.jafc.5b01062

Aliaño-González, M. J., Ferreiro-González, M., Espada-Bellido, E., Palma, M., & Barbero, G. F. (2019). A screening method based on Visible-NIR spectroscopy for the identification and quantification of different adulterants in high-quality honey. *Talanta*, *203*, 235–241. https://doi.org/10.1016/j.talanta.2019.05.067

Arone, B. N., Padmanaban, V., Pandi, G., Shakila, R. J., Neethirajan, N., Saravanan, B., & Velsamy, K. (2019). Nutritional Value and Organoleptic Evaluation of Chocolate Biscuits Incorporated with Freeze Dried and Oven Dried Itoyori Surimi. *Https://Doi.Org/10.2112/SI86-008.1*, *86*(sp1), 56–60. https://doi.org/10.2112/SI86-008.1

Barea-Sepúlveda, M., Ferreiro-González, M., Calle, J. L. P., Barbero, G. F., Ayuso, J., & Palma, M. (2022). Comparison of different processing approaches by SVM and RF on HS-MS eNose and NIR Spectrometry data for the discrimination of gasoline samples. *Microchemical Journal*, *172*. https://doi.org/10.1016/j.microc.2021.106893

Benato, B. C., Telea, A. C., & Falcao, A. X. (2019). Semi-Supervised Learning with Interactive Label Propagation Guided by Feature Space Projections. *Proceedings - 31st Conference on Graphics, Patterns and Images, SIBGRAPI 2018*, 392–399. https://doi.org/10.1109/SIBGRAPI.2018.00057

Benković, M., Bosiljkov, T., Semić, A., Ježek, D., & Srečec, S. (2019). Influence of Carob Flour and Carob Bean Gum on Rheological Properties of Cocoa and Carob Pastry Fillings. *Foods 2019, Vol. 8, Page 66*, *8*(2), 66. https://doi.org/10.3390/FOODS8020066

Bonifazi, G., Capobianco, G., Gasbarrone, R., & Serranti, S. (2021). Contaminant detection in pistachio nuts by different classification methods applied to short-

wave infrared hyperspectral images. *Food Control*, *130*. https://doi.org/10.1016/j.foodcont.2021.108202

Bridges, C. C. (1966). *HIERARCHICAL CLUSTER ANALYSIS* (Vol. 18). @ Southern Universities Press.

Bunea, F., She, Y., Ombao, H., Gongvatana, A., Devlin, K., & Cohen, R. (2011). Penalized least squares regression methods and applications to neuroimaging. *NeuroImage*, *55*(4), 1519–1527. https://doi.org/10.1016/j.neuroimage.2010.12.028

Buttafuoco, G., & Conforti, M. (2021). Improving Mean Annual Precipitation Prediction Incorporating Elevation and Taking into Account Support Size. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/w13060830

Calle, J. L. P., Barea-Sepúlveda, M., Ruiz-Rodríguez, A., Álvarez, J. Á., Ferreiro-González, M., & Palma, M. (2022). Rapid Detection and Quantification of Adulterants in Fruit Juices Using Machine Learning Tools and Spectroscopy Data. *Sensors*, *22*(10). https://doi.org/10.3390/s22103852

Calle, J. L. P., Ferreiro-González, M., Ruiz-Rodríguez, A., Barbero, G. F., Álvarez, J., Palma, M., & Ayuso, J. (2021). A methodology based on ft-ir data combined with random forest model to generate spectralprints for the characterization of high-quality vinegars. *Foods*, *10*(6). https://doi.org/10.3390/foods10061411

Calle, J. L. P., Ferreiro-González, M., Ruiz-Rodríguez, A., Fernández, D., & Palma, M. (2022a). Detection of Adulterations in Fruit Juices Using Machine Learning Methods over FT-IR Spectroscopic Data. *Agronomy*, *12*(3). https://doi.org/10.3390/agronomy12030683

Calle, J. L. P., Ferreiro-González, M., Ruiz-Rodríguez, A., Fernández, D., & Palma, M. (2022b). Detection of Adulterations in Fruit Juices Using Machine Learning Methods over FT-IR Spectroscopic Data. *Agronomy*, *12*(3). https://doi.org/10.3390/agronomy12030683

Calle, J. L. P., Ferreiro-González, M., Ruiz-Rodríguez, A., Fernández, D., & Palma, M. (2022c). Detection of Adulterations in Fruit Juices Using Machine Learning Methods over FT-IR Spectroscopic Data. *Agronomy*, *12*(3). https://doi.org/10.3390/agronomy12030683

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science and Technology*, *18*(2), 72–83. https://doi.org/10.1016/j.tifs.2006.09.003

Chen, C. C. M., Schwender, H., Keith, J., Nunkesser, R., Mengersen, K., & MacRossan, P. (2011). Methods for identifying SNP interactions: A review on variations of logic regression, Random Forest and bayesian logistic regression. In *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (Vol. 8, Issue 6, pp. 1580–1591). https://doi.org/10.1109/TCBB.2011.46

Chen, J., Ren, X., Zhang, Q., Diao, X., & Shen, Q. (2013). Determination of protein, total carbohydrates and crude fat contents of foxtail millet using effective wavelengths in NIR spectroscopy. *Journal of Cereal Science*, *58*(2), 241–247. https://doi.org/10.1016/j.jcs.2013.07.002

Cheng, J. H., & Sun, D. W. (2017). Partial Least Squares Regression (PLSR) Applied to NIR and HSI Spectral Data Modeling to Predict Chemical

Properties of Fish Muscle. In *Food Engineering Reviews* (Vol. 9, Issue 1, pp. 36–49). Springer New York LLC. https://doi.org/10.1007/s12393-016-9147-1

Couto, C. de C., Freitas-Silva, O., Oliveira, E. M. M., Sousa, C., & Casal, S. (2022). Near-infrared spectroscopy applied to the detection of multiple adulterants in roasted and ground arabica coffee. *Foods*, *11*(1). https://doi.org/10.3390/foods11010061

de Santana, F. B., Borges Neto, W., & Poppi, R. J. (2019). Random forest as one-class classifier and infrared spectroscopy for food adulteration detection. *Food Chemistry*, *293*, 323–332. https://doi.org/10.1016/j.foodchem.2019.04.073

Delgado-Ospina, J., Di Mattia, C. D., Paparella, A., Mastrocola, D., Martuscelli, M., & Chaves-Lopez, C. (2020). Effect of fermentation, drying and roasting on biogenic amines and other biocompounds in Colombian criollo cocoa beans and shells. *Foods*, *9*(4), 1–20. https://doi.org/10.3390/foods9040520

Deus, V. L., Bispo, E. S., Franca, A. S., & Gloria, M. B. A. (2021). Understanding amino acids and bioactive amines changes during on-farm cocoa fermentation. *Journal of Food Composition and Analysis*, *97*(December 2020). https://doi.org/10.1016/j.jfca.2020.103776

Esposito Vinzi, V., & Russolillo, G. (2013). Partial least squares algorithms and methods. In *Wiley Interdisciplinary Reviews: Computational Statistics* (Vol. 5, Issue 1, pp. 1–19). https://doi.org/10.1002/wics.1239

Faisal, M., Zamzami, E. M., & Sutarman. (2020). Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance. *Journal of Physics: Conference Series*, *1566*(1). https://doi.org/10.1088/1742-6596/1566/1/012112

Febrianto, N. A., & Zhu, F. (2020). Changes in the Composition of Methylxanthines, Polyphenols, and Volatiles and Sensory Profiles of Cocoa Beans from the sul 1 Genotype Affected by Fermentation. *Journal of Agricultural and Food Chemistry*, *68*(32), 8658–8675. https://doi.org/10.1021/acs.jafc.0c02909

Garza-Ulloa, J. (2018). Application of mathematical models in biomechatronics: artificial intelligence and time-frequency analysis. In *Applied Biomechatronics using Mathematical Models* (pp. 373–524). Elsevier. https://doi.org/10.1016/b978-0-12-812594-6.00006-8

Giglio, C., & Brown, S. D. (2018). Using elastic net regression to perform spectrally relevant variable selection. *Journal of Chemometrics*, *32*(8). https://doi.org/10.1002/cem.3034

Gomes, de araujo A., Azcarate, S. M., Diniz, P. H. G. D., de Sousa Fernandes, D. D., & Veras, G. (2022). Variable selection in the chemometric treatment of food data: A tutorial review. *Food Chemistry*, *370*(July 2021), 131072. https://doi.org/10.1016/j.foodchem.2021.131072

Gómez Hoyos, C., Mazo Márquez, P., Penagos Vélez, L., Serpa Guerra, A., Eceiza, A., Urbina, L., Velásquez-Cock, J., Gañán Rojo, P., Vélez Acosta, L., & Zuluaga, R. (2020). Cocoa shell: an industrial by-product for the preparation of suspensions of holocellulose nanofibers and fat. *Cellulose*, *27*(18), 10873–10884. https://doi.org/10.1007/s10570-020-03222-6

Greño, M., Plaza, M., Luisa Marina, M., & Castro Puyana, M. (2023). Untargeted HPLC-MS-based metabolomics approach to reveal cocoa powder adulterations. *Food Chemistry*, *402*. https://doi.org/10.1016/j.foodchem.2022.134209

Guindo, M. L., Kabir, M. H., Chen, R., & Liu, F. (2021). Potential of Vis-NIR to measure heavy metals in different varieties of organic-fertilizers using Boruta and deep belief network. *Ecotoxicology and Environmental Safety*, *228*. https://doi.org/10.1016/j.ecoenv.2021.112996

Hurst, W. J., Krake, S. H., Bergmeier, S. C., Payne, M. J., Miller, K. B., & Stuart, D. A. (2011). Impact of fermentation, drying, roasting and Dutch processing on flavan-3-ol stereochemistry in cacao beans and cocoa ingredients. *Chemistry Central Journal*, *5*(1), 1–10. https://doi.org/10.1186/1752-153X-5-53

Jardim, R., & Morgado-Dias, F. (2020). Savitzky–Golay filtering as image noise reduction with sharp color reset. *Microprocessors and Microsystems*, *74*. https://doi.org/10.1016/j.micpro.2020.103006

Jaskowiak, P. A., Campello, R. J. G. B., & Costa, I. G. (2014). On the selection of appropriate distances for gene expression data clustering. *BMC Bioinformatics*, *15*. https://doi.org/10.1186/1471-2105-15-S2-S2

Jha, S. N. (2016). *Rapid Detection of Food Adulterants and Contaminants, Theory and Practice*. 278.

Johnson, J. B. (2020). An overview of near-infrared spectroscopy (NIRS) for the detection of insect pests in stored grains. In *Journal of Stored Products Research* (Vol. 86). Elsevier Ltd. https://doi.org/10.1016/j.jspr.2019.101558

Jolić, S. M., Radojčic Redovnikovic, I., Marković, K., Ivanec Šipušić, D., & Delonga, K. (2011). Changes of phenolic compounds and antioxidant capacity in cocoa beans processing. *International Journal of Food Science and Technology*, *46*(9), 1793–1800. https://doi.org/10.1111/j.1365-2621.2011.02670.x

Kamal, M., & Karoui, R. (2015). Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. In *Trends in Food Science and Technology* (Vol. 46, Issue 1, pp. 27–48). Elsevier Ltd. https://doi.org/10.1016/j.tifs.2015.07.007

Kene Ejeahalaka, K., & On, S. L. W. (2020). Effective detection and quantification of chemical adulterants in model fat-filled milk powders using NIRS and hierarchical modelling strategies. *Food Chemistry*, *309*. https://doi.org/10.1016/j.foodchem.2019.125785

Kim, D., & Heo, T. Y. (2022). Anomaly Detection with Feature Extraction Based on Machine Learning Using Hydraulic System IoT Sensor Data. *Sensors*, *22*(7). https://doi.org/10.3390/s22072479

Köhn, H., & Hubert, L. J. (2015). Hierarchical Cluster Analysis. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–13). Wiley. https://doi.org/10.1002/9781118445112.stat02449.pub2

Kong, D., Sun, D., Qiu, R., Zhang, W., Liu, Y., & He, Y. (2022). Rapid and nondestructive detection of marine fishmeal adulteration by hyperspectral

imaging and machine learning. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *273*. https://doi.org/10.1016/j.saa.2022.120990

Kongor, J. E., Hinneh, M., de Walle, D. Van, Afoakwa, E. O., Boeckx, P., & Dewettinck, K. (2016). Factors influencing quality variation in cocoa (Theobroma cacao) bean flavour profile - A review. *Food Research International*, *82*, 44–52. https://doi.org/10.1016/j.foodres.2016.01.012

Kumar, V., Chhabra, J. K., & Kumar, D. (2014). *Performance Evaluation of Distance Metrics in the Clustering Algorithms*.

Kursa, M. B., & Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. In *JSS Journal of Statistical Software* (Vol. 36). http://www.jstatsoft.org/

Laborde, A., Puig-Castellví, F., Jouan-Rimbaud Bouveresse, D., Eveleigh, L., Cordella, C., & Jaillais, B. (2021). Detection of chocolate powder adulteration with peanut using near-infrared hyperspectral imaging and Multivariate Curve Resolution. *Food Control*, *119*(April 2020). https://doi.org/10.1016/j.foodcont.2020.107454

Li, Y., Feng, Y., Zhu, S., Luo, C., Ma, J., & Zhong, F. (2012). The effect of alkalization on the bioactive and flavor related components in commercial cocoa powder. *Journal of Food Composition and Analysis*, *25*(1), 17–23. https://doi.org/10.1016/j.jfca.2011.04.010

Liu, W., & Li, Q. (2017). An efficient elastic net with regression coefficients method for variable selection of spectrum data. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0171122

Martínez, R., Torres, P., Meneses, M. A., Figueroa, J. G., Pérez-Álvarez, J. A., & Viuda-Martos, M. (2012). Chemical, technological and in vitro antioxidant properties of cocoa (Theobroma cacao L.) co-products. *Food Research International*, *49*(1), 39–45. https://doi.org/10.1016/j.foodres.2012.08.005

McDonald, G. C. (2009). Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 93–100. https://doi.org/10.1002/wics.14

Mendlein, A., Szkudlarek, C., & Goodpaster, J. V. (2013). Chemometrics. *Encyclopedia of Forensic Sciences: Second Edition*, 646–651. https://doi.org/10.1016/B978-0-12-382165-2.00259-2

Mishra, P., Biancolillo, A., Roger, J. M., Marini, F., & Rutledge, D. N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. In *TrAC - Trends in Analytical Chemistry* (Vol. 132). Elsevier B.V. https://doi.org/10.1016/j.trac.2020.116045

Monakhova, Y. B., & Diehl, B. W. K. (2015). Combining 1H NMR spectroscopy and multivariate regression techniques to quantitatively determine falsification of porcine heparin with bovine species. *Journal of Pharmaceutical and Biomedical Analysis*, *115*, 543–551. https://doi.org/10.1016/j.jpba.2015.08.017

Moore, J. C., Spink, J., & Lipp, M. (2012). Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010. *Journal of Food Science*, *77*(4). https://doi.org/10.1111/j.1750-3841.2012.02657.x

Nallan Chakravartula, S. S., Moscetti, R., Bedini, G., Nardella, M., & Massantini, R. (2022). Use of convolutional neural network (CNN) combined with FT-NIR spectroscopy to predict food adulteration: A case study on coffee. *Food Control*, *135*. https://doi.org/10.1016/j.foodcont.2022.108816

Nisar, N., Mustafa, F., Tahir, A., Qadri, R., Yang, Y., Khan, M. I., & Wang, F. (2020). Proximate composition, functional properties and quantitative analysis of benzoyl peroxide and benzoic acid in wheat flour samples: Effect on wheat flour quality. *PeerJ*, *2020*(3). https://doi.org/10.7717/peerj.8788

Nobari Moghaddam, H., Tamiji, Z., Akbari Lakeh, M., Khoshayand, M. R., & Haji Mahmoodi, M. (2022a). Multivariate analysis of food fraud: A review of NIR based instruments in tandem with chemometrics. In *Journal of Food Composition and Analysis* (Vol. 107). Academic Press Inc. https://doi.org/10.1016/j.jfca.2021.104343

Nobari Moghaddam, H., Tamiji, Z., Akbari Lakeh, M., Khoshayand, M. R., & Haji Mahmoodi, M. (2022b). Multivariate analysis of food fraud: A review of NIR based instruments in tandem with chemometrics. In *Journal of Food Composition and Analysis* (Vol. 107). Academic Press Inc. https://doi.org/10.1016/j.jfca.2021.104343

NSF. (2014). *Risk Modelling Of Food Fraud Motivation-"NSF Fraud Protection Model" Intelligent Risk Model Scoping Project FS 246004 NSF Safety and Quality UK Ltd*.

Pandiselvam, R., Mahanti, N. K., Manikantan, M. R., Kothakota, A., Chakraborty, S. K., Ramesh, S. v., & Beegum, P. P. S. (2022). Rapid detection of adulteration in desiccated coconut powder: vis-NIR spectroscopy and chemometric approach. *Food Control*, *133*. https://doi.org/10.1016/j.foodcont.2021.108588

Pavlou, M., Ambler, G., Seaman, S., de iorio, M., & Omar, R. Z. (2016). Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. *Statistics in Medicine*, *35*(7), 1159–1177. https://doi.org/10.1002/sim.6782

Pu, Y. Y., O'Donnell, C., Tobin, J. T., & O'Shea, N. (2020). Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing. In *International Dairy Journal* (Vol. 103). Elsevier Ltd. https://doi.org/10.1016/j.idairyj.2019.104623

Qin, X. W., Lai, J. X., Tan, L. H., Hao, C. Y., Li, F. P., He, S. Z., & Song, Y. H. (2017). Characterization of volatile compounds in Criollo, Forastero, and Trinitario cocoa seeds (Theobroma cacao L.) in China. *International Journal of Food Properties*, *20*(10), 2261–2275. https://doi.org/10.1080/10942912.2016.1236270

Quelal-Vásconez, M. A., Pérez-Esteve, É., Arnau-Bonachera, A., Barat, J. M., & Talens, P. (2018). Rapid fraud detection of cocoa powder with carob flour using near infrared spectroscopy. *Food Control*, *92*, 183–189. https://doi.org/10.1016/j.foodcont.2018.05.001

Rababah, T. M., Al-u'datt, M., Ereifej, K., Almajwal, A., Al-Mahasneh, M., Brewer, S., Alsheyab, F., & Yang, W. (2013). Chemical, Functional and

Sensory Properties of Carob Juice. *Journal of Food Quality*, *36*(4), 238–244. https://doi.org/10.1111/jfq.12033

Rahmani, N., & Mani-Varnosfaderani, A. (2022). Quality control, classification, and authentication of Iranian rice varieties using FT-IR spectroscopy and sparse chemometric methods. *Journal of Food Composition and Analysis*, *112*. https://doi.org/10.1016/j.jfca.2022.104650

Rajoub, B. (2020). Supervised and unsupervised learning. In *Biomedical Signal Processing and Artificial Intelligence in Healthcare* (pp. 51–89). Elsevier. https://doi.org/10.1016/b978-0-12-818946-7.00003-2

Ranstam, J., & Cook, J. A. (2018). LASSO regression. *British Journal of Surgery*, *105*(10), 1348. https://doi.org/10.1002/bjs.10895

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. In *TrAC - Trends in Analytical Chemistry* (Vol. 28, Issue 10, pp. 1201–1222). https://doi.org/10.1016/j.trac.2009.07.007

Rodriguez-Campos, J., Escalona-Buendía, H. B., Contreras-Ramos, S. M., Orozco-Avila, I., Jaramillo-Flores, E., & Lugo-Cervantes, E. (2012). Effect of fermentation time and drying temperature on volatile compounds in cocoa. *Food Chemistry*, *132*(1), 277–288. https://doi.org/10.1016/j.foodchem.2011.10.078

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, *71*, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Sanguigni, V., Manco, M., Sorge, R., Gnessi, L., & Francomano, D. (2017). Natural antioxidant ice cream acutely reduces oxidative stress and improves vascular function and physical performance in healthy individuals. *Nutrition*, *33*, 225–233. https://doi.org/10.1016/J.NUT.2016.07.008

Santander Muñoz, M., Rodríguez Cortina, J., Vaillant, F. E., & Escobar Parra, S. (2020). An overview of the physical and biochemical transformation of cocoa seeds to beans and to chocolate: Flavor formation. In *Critical Reviews in Food Science and Nutrition* (Vol. 60, Issue 10, pp. 1593–1613). Taylor and Francis Inc. https://doi.org/10.1080/10408398.2019.1581726

Santos, I. A., Conceição, D. G., Viana, M. B., Silva, G. de J., Santos, L. S., & Ferrão, S. P. B. (2021). NIR and MIR spectroscopy for quick detection of the adulteration of cocoa content in chocolates. *Food Chemistry*, *349*(June 2020). https://doi.org/10.1016/j.foodchem.2021.129095

Sarker, S. D., & Nahar, L. (2015). Applications of High Performance Liquid Chromatography in the Analysis of Herbal Products. In *Evidence-Based Validation of Herbal Medicine*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-800874-4.00019-2

Schreiber-Gregory, D. N. (2018). Ridge Regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications*, *13*(4), 359–365. https://doi.org/10.3233/MAS-180446

Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture*, *183*. https://doi.org/10.1016/j.compag.2021.106036

Song, Q. C., Tang, C., & Wee, S. (2021). Making Sense of Model Generalizability: A Tutorial on Cross-Validation in R and Shiny. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920947067

Soni Madhulatha, T. (2012). *AN OVERVIEW ON CLUSTERING METHODS*. *2*(4), 719–725. www.iosrjen.org

Spizzirri, U. G., Ieri, F., Campo, M., Paolino, D., Restuccia, D., & Romani, A. (2019). Biogenic amines, phenolic, and aroma-related compounds of unroasted and roasted cocoa beans with different origin. *Foods*, *8*(8). https://doi.org/10.3390/foods8080306

Stanimirova, I., & Daszykowski, M. (2018). Exploratory Analysis of Metabolomic Data. In *Comprehensive Analytical Chemistry* (Vol. 82, pp. 227–264). Elsevier B.V. https://doi.org/10.1016/bs.coac.2018.08.005

Stankov, S., Dzhivoderova-Zarcheva, M., Dimitrova, E., Damyanova-Bakardzhieva, M., & Fidan, H. (2020). Rheological and sensory properties of glazes prepared with carob and cocoa powders. *Journal of Food Processing and Preservation*, *44*(8), e14580. https://doi.org/10.1111/JFPP.14580

Steinbach, A. (2013). *NIR Spectroscopy A guide to near-infrared spectroscopic analysis of industrial manufacturing processes Monograph NIR Spectroscopy*.

Strauss, T., & von Maltitz, M. J. (2017). Generalising ward's method for use with manhattan distances. *PLoS ONE*, *12*(1). https://doi.org/10.1371/journal.pone.0168288

Su, N., Weng, S., Wang, L., & Xu, T. (2021). Reflectance spectroscopy with multivariate methods for non-destructive discrimination of edible oil adulteration. *Biosensors*, *11*(12). https://doi.org/10.3390/bios11120492

Teye, E., Huang, X. yi, Lei, W., & Dai, H. (2014). Feasibility study on the use of Fourier transform near-infrared spectroscopy together with chemometrics to discriminate and quantify adulteration in cocoa beans. *Food Research International*, *55*, 288–293. https://doi.org/10.1016/j.foodres.2013.11.021

Tian, F., Yang, L., Lv, F., & Zhou, P. (2009). Predicting liquid chromatographic retention times of peptides from the Drosophila melanogaster proteome by machine learning approaches. *Analytica Chimica Acta*, *644*(1–2), 10–16. https://doi.org/10.1016/j.aca.2009.04.010

Veselá, A., Barros, A. S., Synytsya, A., Delgadillo, I., Čopíková, J., & Coimbra, M. A. (2007). Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder. *Analytica Chimica Acta*, *601*(1), 77–86. https://doi.org/10.1016/j.aca.2007.08.039

Wang, Y. Y., Li, J. Q., Liu, H. G., & Wang, Y. Z. (2019). Attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) combined with chemometrics methods for the classification of Lingzhi species. *Molecules*, *24*(12). https://doi.org/10.3390/molecules24122210

Wang, Z., Wang, Y., Zeng, R., Srinivasan, R. S., & Ahrentzen, S. (2018). Random Forest based hourly building energy prediction. *Energy and Buildings*, *171*, 11–25. https://doi.org/10.1016/j.enbuild.2018.04.008

Wei, J., Chu, X., Sun, X. Y., Xu, K., Deng, H. X., Chen, J., Wei, Z., & Lei, M. (2019). Machine learning in materials science. In *InfoMat* (Vol. 1, Issue 3, pp. 338–358). Blackwell Publishing Ltd. https://doi.org/10.1002/inf2.12028

Weng, S., Guo, B., Tang, P., Yin, X., Pan, F., Zhao, J., Huang, L., & Zhang, D. (2020). Rapid detection of adulteration of minced beef using Vis/NIR reflectance spectroscopy with multivariate methods. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *230*. https://doi.org/10.1016/j.saa.2019.118005

Wilde, A. S., Haughey, S. A., Galvin-King, P., & Elliott, C. T. (2019). The feasibility of applying NIR and FT-IR fingerprinting to detect adulteration in black pepper. *Food Control*, *100*, 1–7. https://doi.org/10.1016/j.foodcont.2018.12.039

World Health Organization. (2015). *Food safety: What you should know*.

Yang, W. li, Hu, M. hua, Chen, S. wei, Wang, Q., Zhu, S., Dai, J., & Li, X. zhong. (2015). Identification of Adulterated Cocoa Powder Using Chromatographic Fingerprints of Polysaccharides Coupled with Principal Component Analysis. *Food Analytical Methods*, *8*(9), 2360–2367. https://doi.org/10.1007/s12161-015-0126-6

Yao, H., & Lewis, D. (2010). Spectral Preprocessing and Calibration Techniques. In *Hyperspectral Imaging for Food Quality Analysis and Control* (pp. 45–78). Elsevier Inc. https://doi.org/10.1016/B978-0-12-374753-2.10002-4

Yim, O., & Ramdeen, K. T. (2015). *Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data* (Vol. 11, Issue 1).

Yong, C. H., Muhammad, S. A., Aziz, F. A., Ng, J. S., Nasir, F. I., Adenan, M. N. H., Moosa, S., Othman, Z., Abdullah, S. N. A., Sharif, Z., Ismail, F. F., Kelly, S. D., Cannavan, A., & Seow, E. K. (2022). Detection of adulteration activities in edible bird's nest using untargeted 1H-NMR metabolomics with chemometrics. *Food Control*, *132*(August 2021), 108542. https://doi.org/10.1016/j.foodcont.2021.108542

Zhu, Y., Zou, L., & Tan, T. L. (2022). A clustering group lasso method for quantification of adulteration in black cumin seed oil using Fourier transform infrared spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, *220*. https://doi.org/10.1016/j.chemolab.2021.104471

Zhu, Z., Chen, S., Wu, X., Xing, C., & Yuan, J. (2018). Determination of soybean routine quality parameters using near-infrared spectroscopy. *Food Science and Nutrition*, *6*(4), 1109–1118. https://doi.org/10.1002/fsn3.652

Zuo, R., & Carranza, E. J. M. (2011). Support vector machine: A tool for mapping mineral prospectivity. *Computers and Geosciences*, *37*(12), 1967–1975. https://doi.org/10.1016/j.cageo.2010.09.014