Contract no. 651PED/2022
Project code: PN-III-P2-2.1-PED-2021-1095

Wine fingerprint recognition based on spectroscopic methods and Artificial Intelligence (WineRec – AI)

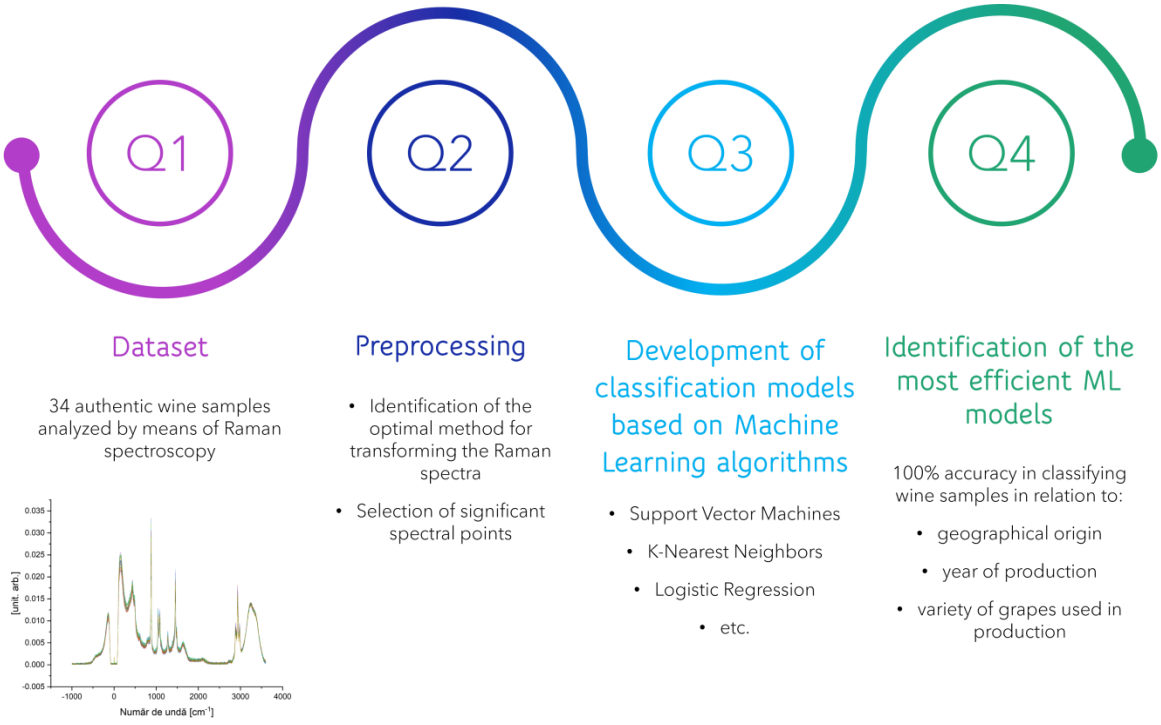
Phase 1 summary: Raman spectroscopy in association with Artificial Intelligence for the development of wine recognition models

Implementation period: 01.08.2022 – 31.12.2022

1. Summary of the activities completed during the implementation period

In this execution stage, the potential of Raman spectroscopy for differentiating wine based on three classification criteria – geographical origin, production year, and grape variety – was investigated. For this purpose, an experimental dataset obtained by analyzing 34 authentic wine samples produced in Romania over five consecutive production years (2012–2016) was used. An initial and crucial step in the development of reliable wine recognition models involved preprocessing the experimental data to improve the accuracy of the recognition models. Five techniques for transforming the Raman spectra (i.e. autoscale, variance (std) scaling, smoothing, 1st derivative, and Pareto) were applied, and their effectiveness was compared through the application of the Partial Least Squares Discriminant Analysis (PLS-DA) method. Additionally, a variable reduction method was applied by selecting those spectral points with the highest discrimination power for each classification criterion.

The application of Machine Learning algorithms enabled the development of differentiation models with a correct prediction capability of up to 100% for the geographical origin, cultivar, and production year classification in 10-fold cross-validation. It was observed that the reduction of the input dataset's dimensionality by selecting and subsequently using only the markers with highest discrimination capability leads to the construction of the most effective classification models. The Machine Learning methods that corresponded to the classifiers with the highest accuracy were: Support Vector Machines, K-Nearest Neighbors, Linear Discriminant Analysis, and Logistic Regression.



The results obtained in this stage of the project were disseminated as follows:

1. **1 ISI article** published in a journal in the first quartile (Q1 - Web of Science)
2. **2 presentations** at international conferences

2. Progress summary

	Planned Deliverables/Indicators	No.	Achieved Deliverables/Indicators	No.
1.	**ISI article submitted for publication**	**1**	**ISI article published**	**1**
2.	**Presentations at international conferences**	**2**	**Presentations at international conferences**	**2**
3.	**Prediction models**	**3**	**Prediction models**	**12**
4.	**Project web page**	**1**	**Project web page**	**1**

The dissemination of results is presented in detail as follows:

I. Publication of an article in a journal in the first quartile (Q1 – Web of Science)

1. Dehelean, A., Cristea, G., Feher, I., Hategan, A. R., Magdas, D. A. (2022). Differentiation of Transylvanian fruit distillates using supervised statistical tools based on isotopic and elemental fingerprint. *Journal of the Science of Food and Agriculture*, 103(3), 1454-1463. https://doi.org/10.1002/jsfa.12241

II. Participations at international conferences

1. A. R. Hategan, D. A. Magdas, A comparison among distinct data pre-processing methods for the improvement of wine recognition models. 10th International Conference "Agriculture & Food", 16-19 August 2022, Burgas, Bulgaria.
2. A. Pirnau, I. Feher, C. Sarbu, A.R. Hategan, F. Guyon, D. A. Magdas, Wine recognition model development through the association between ^{1}H-NMR spectroscopy and Fuzzy algorithms, RAFA, 6-9 September 2022, Prague, Czech Republic.

Project Director,
MAGDAS Dana-Alina