agriculture

Article

The Use of Fluorescence Spectroscopic Data and Machine-Learning Algorithms to Discriminate Red Onion Cultivar and Breeding Line

Kadir Sabanci [1,*], Muhammet Fatih Aslan [1], Vanya Slavova [2] and Stefka Genova [2]

1 Department of Electrical and Electronics Engineering, Karamanoglu Mehmetbey University, Karaman 70100, Turkey
2 Department of Plant Breeding, Maritsa Vegetable Crops Research Institute, Agricultural Academy Bulgaria, 32, Brezovsko Shosse St., 4003 Plovdiv, Bulgaria
* Correspondence: kadirsabanci@kmu.edu.tr

Abstract: The objective of this study was to evaluate differences between the red onion cultivar and breeding line using models based on selected fluorescence spectroscopic data built using machine-learning algorithms from different groups of Trees, Functions, Bayes, Meta, Rules, and Lazy. The combination of fluorescence spectroscopy and machine learning is an original approach to the non-destructive and objective discrimination of red onion samples. The selected fluorescence spectroscopic data were used to build models using algorithms from the groups of Trees, Functions, Bayes, Meta, Rules, and Lazy. The most satisfactory results were obtained using J48 and LMT (Logistic Model Tree) from the group of Trees, Multilayer Perceptron, and QDA (Quadratic Discriminant Analysis) from Functions, Naive Bayes from Bayes, Logit Boost from Meta, JRip from Rules, and LWL (Locally Weighted Learning) from Lazy. The average accuracy of discrimination of onion bulbs belonging to 'Asenovgradska kaba' and a red breeding line equal to 100% was found in the case of models developed using the LMT, Multilayer Perceptron, Naive Bayes, Logit Boost, and LWL algorithms. The TPR (True Positive Rate), Precision, and F-Measure of 1.000 and FPR (False Positive Rate) of 0.000, as well as the Kappa statistic of 1.0, were determined. The results revealed the usefulness of the approach combining fluorescence spectroscopy and machine learning to distinguish red onion cultivars and breeding lines.

Keywords: onion bulb; onion cultivar; onion breeding line; fluorescence spectroscopy; machine-learning algorithms; discrimination

Citation: Sabanci, K.; Aslan, M.F.; Slavova, V.; Genova, S. The Use of Fluorescence Spectroscopic Data and Machine-Learning Algorithms to Discriminate Red Onion Cultivar and Breeding Line. *Agriculture* **2022**, *12*, 1652. https://doi.org/10.3390/agriculture12101652

Academic Editor: Roberto Alves Braga Júnior

Received: 14 July 2022
Accepted: 7 October 2022
Published: 9 October 2022

1. Introduction

The onion (*Allium cepa* L.) is an old and developed agricultural crop. Its production increased throughout the world. There are various onions, such as red, white, and yellow containing bioactive compounds, e.g., phenolics, anthocyanins, and flavonoids which are found in the greatest amount in red onions. Onion contains an average of 89.1% of water, 9.3% of carbohydrates, 1.1% of protein, and 0.1% of fat, as well as vitamins and minerals. As well as the flesh that is mostly used in the diet, the skin of the red onion is also a source of valuable bioactive compounds. The flavonoid content in the skin of the red onion can reach 2–10 g kg^{-1} compared to below 0.03 g kg^{-1} to above 1 g kg^{-1} in the edible parts [1]. Red onion skin may be a renewable raw material for the extraction of bioactive compounds useful as functional components in value-added products [2]. Anthocyanins may occur as pigments in red onions [3]. Anthocyanins are present in the epidermal cells of red onions. The inner epidermis is normally white, but anthocyanic red cells may be interspersed between anthocyanin-free white cells because small patches can sometimes redden if containing vacuolar anthocyanin [4]. Onions can be a source of additives inhibiting microbial spoiling and oxidative deterioration to retard the deterioration of foods. Due to the antioxidant and antimicrobial properties, flavonoids from onions may be considered as additives to increase food shelf-life. Red cultivars are characterized by the highest antioxidant activities [5]. Due to its properties, the onion is used as food and for medical purposes, e.g., for the prevention or treatment of cancer, obesity, coronary heart disease, type 2 diabetes, hypercholesterolemia, hypertension, disturbances of the gastrointestinal tract, and cataract [6].

The chemical composition of red onion, including, e.g., the content of flavonols, anthocyanins, total phenols, sulfur dioxide, and reducing sugars, can differ depending on, among others, the cultivar, line, and ecotype [7–10]. Therefore, distinguishing different onion types may be important. The approach involving fluorescence spectroscopy and machine learning can be considered useful for the discrimination of plant samples. It happens that after obtaining the data, people are not able to extract useful information from big datasets and make a decision. Machine learning enables processing and interpreting the data using machines learning by themselves from the data without being explicitly programmed. Machine learning uses various algorithms to solve problems and one universal algorithm does not exist. The choice of an appropriate algorithm depends on the number of variables and the kind of problem [11]. Machine learning creates new opportunities for multi-disciplinary agritechnologies. Applying machine learning to sensor data enables the evolvement of farm management systems in artificial intelligence-enabled programs. It provides recommendations and insights to support farmers in decision-making [12].

2. Related Work

In the case of onions, the combination of fluorescence spectroscopic data and machine learning proved to be useful for distinguishing different samples of onions subjected to normal watering and drought mode [13]. The evaluation of differentiation of varieties and breeding lines of potatoes using spectroscopic data and machine-learning algorithms was successfully applied by Slavova et al. [14]. The combination of machine learning and spectroscopy was also used, e.g., for the determination of fruit maturity [15,16]. Additionally, the machine-learning algorithms were applied in the previous research for the cultivar discrimination of different fruit and vegetable samples, e.g., whole fruit, flesh cross-sections, stones and seeds of peach [17], pits of sour cherry [18], flesh cross-sections and seeds of pepper [19].

This study aimed at evaluating the differences between red onion cultivars and breeding lines using models based on selected fluorescence spectroscopic data built using machine-learning algorithms from different groups of Trees, Functions, Bayes, Meta, Rules, and Lazy. The combination of fluorescence spectroscopy and machine learning is an original approach to the non-destructive and objective discrimination of red onion samples.

3. Materials and Methods

3.1. Materials

The materials are comprised of 'Asenovgradska kaba' cultivar and the red breeding line of onions. The bulbs of 'Asenovgradska kaba' are large, flat and round with a very characteristic conical back shape, flat to slightly concave on the neck, and strongly elongated on the bottom, with an index of 0.8–0.9 and a weight of 120–150 g. The sheath scales are colored purple-red. The fleshy inner scales are thick, loose, and purple-red as the outer ones are more intensely colored, and the inner ones are weaker and only on the surface. The red breeding line of onions was created by the method of the individual team. The bulbs are medium-sized and weigh 80–100 g. The outer shells are red. The fleshy inner scales are tightly arranged. The shape of the bulb is flat round with an index of 0.9. The sample images of onion bulbs of 'Asenovgradska kaba' cultivar and the red breeding line are presented in Figure 1.

(**a**)

(**b**)

Figure 1. The sample images of onion bulbs of 'Asenovgradska kaba' cultivar (**a**) and red breeding line (**b**).

The profile of the soil surface at 'Asenovgradska kaba' cultivar and the red breeding line of onion is the furrow–furrow with a high flat bed with five rows in the scheme 85 + 25 + 25 + 25 + 25 cm and the sowing rate is 0.6 kg/da. The seeds are sown to a depth of 3 cm. They are suitable for early spring sowing and have a faster rate of growth and development, are grown by direct sowing of seeds and for a shorter growing season of 95–100 days under irrigated conditions from large bulbs.

The harvest time of 'Asenovgradska kaba' cultivar and the red breeding line of onions is of great importance for the productivity, quality, and shelf life of the bulbs. They are removed when 25 to 75% of the false stems are lying down. Removal starts between the single and mass laying phases to complete the complete laying of the stems.

3.2. Fluorescence Spectroscopy

The fiber optic spectrometer allowing the generation of fluorescent emission signals in the range of 200–1200 nm was used in this study (Figure 2). The experimental setup consisted of a portable spectrometer model AvaSpec-ULS2048CL-EVO and a laser diode (optical power 16 mW, emission wavelength 285 nm, DC). The red onion samples were placed on a duralumin stand allowing the reception of an emission signal below 180° of U-shaped optical fiber that reduced aberrations and allowed the generation of a better-quality emission signal (Figure 2). The spectrometer resolution can range between 0.06–20 nm, and that of the setting of 0.09 nm was used for this study. The fluorescence can be often weak and, in all directions, in order not to saturate the receiver, the useful fluorescence signal can be measured in a direction that is below 180° to the excitation radiation. The application of a laser diode (LED) as a source in the circuit is preferable. Its spectral width is small. LED used in this study had a relatively wide spectral width of radiation (about 30–40 nm) and the angular distribution of its radiation was in a large angular range (±30°). The photodetector of the CMOS type model S9132 was chosen for the specific circuit because it can detect emission radiation from red onion bulbs with very high intensity. The sensitivity of the spectrometer ranged from 200 nm to 1200 nm. Its resolution was $\delta\lambda = 5$ nm.

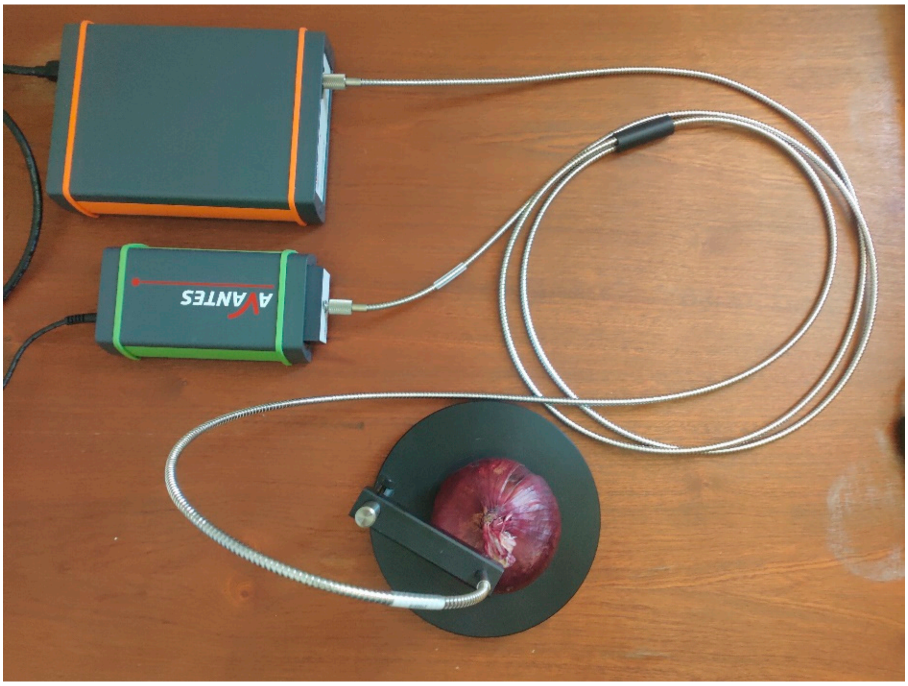

Figure 2. General view of the experimental installation used by fluorescence spectroscopy.

Recording the spectrum of excitation source and emission can be possible due to spectral installation based on fluorescent signals. The excitation spectrum meant the dependence of the emission intensity for one wavelength when scanning on the excitation wavelength. This spectrum was represented as a dependence of the wavelength of light on the light intensity incident on the photodetector in the spectrometer. The emission spectrum was the emission wave distribution measured for a constant excitation wavelength.

The laser radiation was removed from the source and fell on the red onion sample. Then, the emission signal fell on a U-shaped optical fiber with a core diameter of 200 µm with a numerical aperture of 0.22 and a step index of refractive index. Then, it was taken to the detector. In the spectrometer, the light signal was converted to electrical–digital using a USB 2.0 wire, downloaded to a computer with the use of AvaSoft8 software and exported to Excel. This allowed analysis, processing, and visualization of the results of the study. The fluorescence data were used for the development of discriminative models to distinguish of 'Asenovgradska kaba' cultivar and the red breeding line of red onion bulbs.

3.3. Statistical Analysis

The bulbs of the onion of 'Asenovgradska kaba' cultivar and the red breeding line were distinguished using the WEKA machine learning application (Machine Learning Group, University of Waikato, Hamilton, New Zealand) [20–22]. There are 5 graphs each from 'Asenovgradska kaba' and the red breeding line. A difference in the emission fluorescence signal of 'Asenovgradska kaba' and the red breeding line, as well as variety and selection line, is clearly observed. The spectral shift in wavelength and signal intensity level is due to a difference in the content of biologically active substances of a particular variety (Figure 3). The applied procedure of the red onion sample distinguishing is presented in Figure 4.

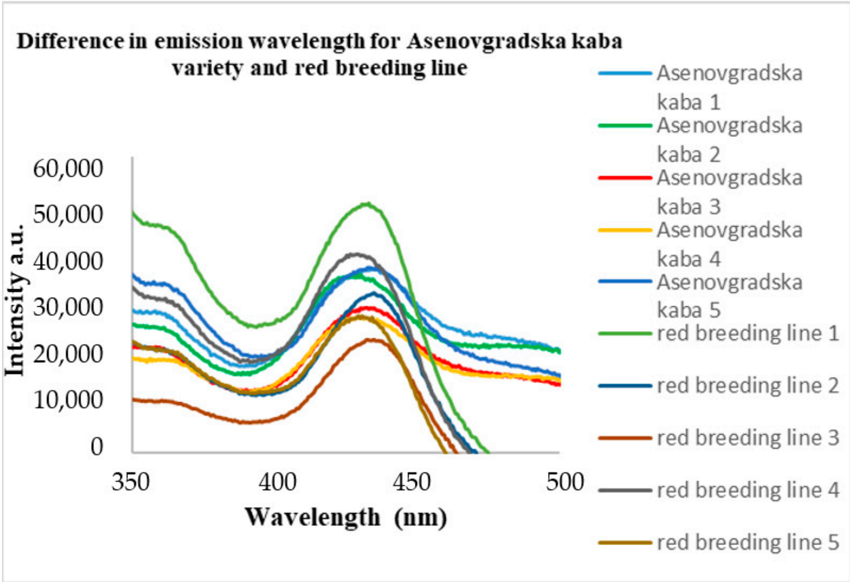

Figure 3. Difference in emission wavelength for 'Asenovgradska kaba' variety and red breeding line.

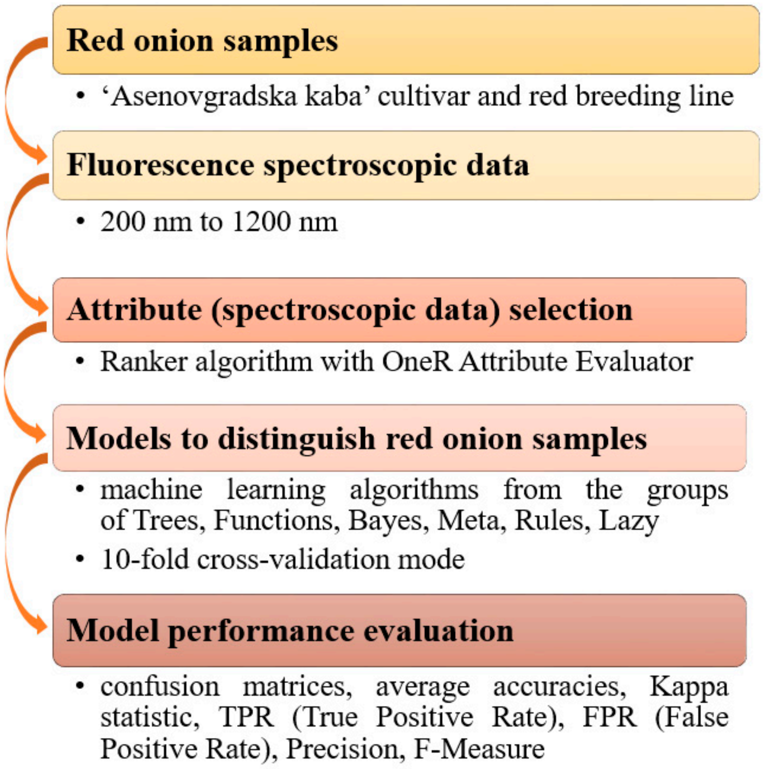

Figure 4. The procedure of distinguishing red onion samples using fluorescence spectroscopy data and machine learning.

The dataset consisted of 1353 attributes for each class ('Asenovgradska kaba' cultivar and red breeding line). These attributes were the measurements at different wavelengths in the considered range of 200–1200 nm. The attribute selection was carried out using the Ranker with the OneR Attribute Evaluator to choose the features with the highest discriminative power. The attributes were ranked by their individual evaluations. The worth of an attribute was evaluated using the OneR classifier at the following parameters: seed for the cross-validation, 1; the number of folds for the cross-validation, 10; the minimum number of objects in a bucket, 6; using the training data to evaluate attributes rather than cross-validation, False. Ten attributes were selected to maintain the ratio of attributes (10) to the number of cases (100) of 1:10. Then, the discriminative models were developed based on selected data using a 10-fold cross-validation mode. This mode randomly divides a set of selected data into 10 parts and each part is considered as the test set in turn and the remaining 9 parts are used as the training sets. Finally, the average of 10 estimates for learning performed 10 times using different training sets are computed. Different machine-learning algorithms from the groups of Trees, Functions, Bayes, Meta, Rules, and Lazy were tested to choose algorithms providing the most satisfactory discrimination performance metrics. The confusion matrices, average accuracies, as well as the values of TPR (True Positive Rate), FPR (False Positive Rate), Precision, F-Measure, and Kappa statistic were determined based on Equations (1)–(7).

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (1)$$

$$\text{TPR} = \text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (2)$$

$$\text{FPR} = \text{FP}/(\text{FP} + \text{TN}) \quad (3)$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (4)$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (5)$$

$$\text{F}-\text{Measure} = 2 \times ((\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall})) \quad (6)$$

$$\text{Kappa} = \frac{\frac{(\text{TP}+\text{FP})(\text{TP}+\text{FN})}{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})} + \frac{(\text{TN}+\text{FP})(\text{TN}+\text{FN})}{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}}{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})} \quad (7)$$

where TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative; TPR: True Positive Rate; FPR: False Positive Rate.

The criteria for the evaluation of the obtained results were the highest average accuracy and accuracies of discrimination for both 'Asenovgradska kaba' and the red breeding line and the highest values of TPR (True Positive Rate), Precision, F-Measure, and the Kappa statistic, as well as the lowest values of FPR (False Positive Rate). The most satisfactory results are the TPR, Precision, and F-Measure, and the Kappa statistic is equal to 1.000 and the FPR is equal to 0.000. Such results indicate the complete correctness of the sample discrimination. All cases belonging to one actual class are correctly included in this class and all cases from the second class are correctly classified as a second class.

4. Results

The confusion matrices and average accuracies of distinguishing 'Asenovgradska kaba' cultivar and the red breeding line of onion bulbs based on fluorescence spectroscopic data are presented in Table 1. The models were built based on selected fluorescence spectroscopic data using machine-learning algorithms from different groups. The results obtained using J48 from the group of Trees, LMT from Trees, Multilayer Perceptron from Functions, QDA from Functions, Naive Bayes from Bayes, Logit Boost from Meta, JRip from Rules, and LWL from Lazy were chosen to be presented. The accuracies and other metrics were the most satisfactory for these algorithms. The average accuracy of 100% was obtained for models built using LMT, Multilayer Perceptron, Naive Bayes, Logit Boost, and LWL. It meant that all examined onion bulbs were correctly classified. 100% of cases belonging to the actual class 'Asenovgradska kaba' were correctly included in the predicted class of 'Asenovgradska kaba' and 100% of cases of onion bulbs of the red breeding line were correctly classified as a red breeding line. For other machine-learning algorithms, also high accuracies were observed. An average accuracy of 95% was found for models built using J48 and JRip. In the case of both algorithms, 'Asenovgradska kaba' samples were correctly discriminated in 100%. Onion bulbs belonging to the red breeding line were correctly distinguished from onion bulbs of 'Asenovgradska kaba' in 90%. The remaining 10% of cases belonging to the red breeding line were incorrectly included in the predicted class of

'Asenovgradska kaba'. The models built using the QDA algorithm provided 90% average accuracy. Both classes of the red breeding line and 'Asenovgradska kaba' were correctly distinguished in 90% from each other. Additionally, 10% of bulbs of the red breeding line were incorrectly classified as 'Asenovgradska kaba' and 10% of bulbs belonging to 'Asenovgradska kaba' red onion were incorrectly included in the predicted class of the red breeding line. Such results are very promising and show that it is possible to distinguish a cultivar from a line of red onions with an accuracy of up to 100% using models built based on the selected fluorescence spectroscopic data using selected machine-learning algorithms.

Table 1. The accuracies of discrimination of onion bulbs belonging to 'Asenovgradska kaba' cultivar and red breeding line based on selected fluorescence spectroscopic data.

Classifier	Predicted Class (%)		Actual Class	Average Accuracy (%)
	'Asenovgradska kaba'	**Red Breeding Line**		
J48 (Trees)	100 0	0 90	'Asenovgradska kaba' red breeding line	95
LMT (Trees)	100 0	0 100	'Asenovgradska kaba' red breeding line	100
Multilayer Perceptron (Functions)	100 0	0 100	'Asenovgradska kaba' red breeding line	100
QDA (Functions)	90 0	0 90	'Asenovgradska kaba' red breeding line	90
Naive Bayes (Bayes)	100 0	0 100	'Asenovgradska kaba' red breeding line	100
Logit Boost (Meta)	100 0	0 100	'Asenovgradska kaba' red breeding line	100
JRip (Rules)	100 0	0 90	'Asenovgradska kaba' red breeding line	95
LWL (Lazy)	100 0	0 100	'Asenovgradska kaba' red breeding line	100

In addition to accuracies, other discrimination performance metrics of discrimination 'Asenovgradska kaba' cultivar and the red breeding line of red onion bulbs belonging to, for example, TPR (True Positive Rate), FPR (False Positive Rate), Precision, F-Measure, and Kappa statistic were computed, and the obtained results are shown in Table 2. In the case of models developed using LMT (Trees), Multilayer Perceptron (Functions), Naive Bayes (Bayes), Logit Boost (Meta), and LWL (Lazy) algorithms, the values of TPR Precision, and F-Measure were equal to 1.000 for both 'Asenovgradska kaba' and the red breeding line of red onion. The Kappa statistic reached the value of 1.0. The values of FPR (False Positive Rate) of 0.000 were determined for each class. The model developed using QDA was characterized by the lowest Kappa statistic of 0.8, TPR, Precision, and F-Measure of 0.900 for 'Asenovgradska kaba' cultivar and the red breeding line and FPR of 0.100 for both classes. Additionally, for the J48 and JRip algorithms, high values of TPR (1.000 for 'Asenovgradska kaba' and 0.900 for the red breeding line), Precision (0.909 for 'Asenovgradska kaba' and 1.000 for the red breeding line), F-Measure (0.952 for 'Asenovgradska kaba' and 0.947 for the red breeding line) and Kappa statistic (0.9), and low TPR (0.100 for 'Asenovgradska kaba' and 0.000 for the red breeding line) were obtained (Table 2).

Table 2. The performance metrics of discrimination of onion bulbs of 'Asenovgradska kaba' and red breeding line based on selected fluorescence spectroscopic data.

Classifier	Predicted Class	TPR	FPR	Precision	F-Measure	Kappa Statistic
J48 (Trees)	'Asenovgradska kaba' red breeding line	1.000 0.900	0.100 0.000	0.909 1.000	0.952 0.947	0.9
LMT (Trees)	'Asenovgradska kaba' red breeding line	1.000 1.000	0.000 0.000	1.000 1.000	1.000 1.000	1.0
Multilayer Perceptron (Functions)	'Asenovgradska kaba' red breeding line	1.000 1.000	0.000 0.000	1.000 1.000	1.000 1.000	1.0
QDA (Functions)	'Asenovgradska kaba' red breeding line	0.900 0.900	0.100 0.100	0.900 0.900	0.900 0.900	0.8
Naive Bayes (Bayes)	'Asenovgradska kaba' red breeding line	1.000 1.000	0.000 0.000	1.000 1.000	1.000 1.000	1.0
Logit Boost (Meta)	'Asenovgradska kaba' red breeding line	1.000 1.000	0.000 0.000	1.000 1.000	1.000 1.000	1.0
JRip (Rules)	'Asenovgradska kaba' red breeding line	1.000 0.900	0.100 0.000	0.909 1.000	0.952 0.947	0.9
LWL (Lazy)	'Asenovgradska kaba' red breeding line	1.000 1.000	0.000 0.000	1.000 1.000	1.000 1.000	1.0

TPR—True Positive Rate; FPR—False Positive Rate.

5. Discussion

The carried out research is the first approach to distinguishing red onion cultivars and lines using innovative models built based on features selected from very large data sets obtained by fluorescence spectroscopy and processing using machine-learning algorithms from various groups of Trees, Functions, Bayes, Meta, Rules, and Lazy to build models based on selected features. The developed procedures can support quality assessment and decision-making. By applying novel approaches to assessing the differentiation of samples, combining the different characteristics obtained using a non-destructive technique and artificial intelligence, distinguishing cultivars and lines within one species can be objective and efficient, and not damage the sample. Therefore, the obtained results can be very useful in practice. Such statements may have large application potential.

The application of spectroscopy including, e.g., fluorescence spectroscopy or visible and infrared reflectance spectroscopy is considered as an innovative approach to plant quality evaluation. The advantages include objectiveness, non-destructiveness, no sample preparation, rapidness, accuracy, and effectiveness. Spectroscopy can be used, among others, to diagnose plant diseases [23] or for the direct estimation of sensory qualities of intact fruit and the determination of physicochemical features [24]. Fluorescence spectroscopy may have advantages over other techniques for assessing fruit quality. Fluorescence spectroscopy uses a larger area for analysis than colorimetric measurements which is advantageous when the sample is heterogeneous. In the case of visual classification, there are limitations related to inaccurate, subjective personal perceptions of sample appearance and the dependence of manual evaluation performed by a human on external conditions [25]. Often, the evaluation of fruit quality by spectrometry may require additional use of multivariate modeling or chemometric analysis [26]. Multivariate analysis techniques can be applied to extract significant spectral data. For the processing of spectral features, artificial intelligence can be used. For example, surface-enhanced Raman spectroscopy (SERS) combined with machine-learning algorithms applied to classify coffee beverages [27]. Combining reflection spectroscopy with machine learning allowed for the evaluation of the presence of acrylamide precursors in potato samples [28]. In the case of grained almonds, machine-learning algorithms were applied for the detection of aflatoxins B based on fluorescence spectroscopy data [29]. Raman spectroscopy, multivariate analysis and machine learning allowed for the successful rapid detection of foodborne pathogens [30]. The application of Raman spectroscopy coupled with machine learning was also used to evaluate edible oils [31]. Total synchronous fluorescence spectroscopy and deep learning were used for the rapid identification of the sesame oil authenticity [30]. Promising own results and literature data may prompt further research involving spectroscopy and machine learning for the quality evaluation of fruit and vegetables. Future research may also compare the usefulness of machine learning to build models based on spectroscopic data with models including data obtained using other non-destructive techniques, such as image analysis. Digital image analysis allows the determination of the morphological, textural, and optical parameters and their interpretation in a non-destructive, inexpensive, and fast way. Among the morphological parameters, image processing allows computing the geometric parameters including shape factors and linear dimensions. Textural image analysis can allow for quantitative analysis of textures to evaluate object quality. Image textures as numerical data can be different for objects even if they are characterized by the same number of pixels and color histograms but a dissimilar color distribution that can be difficult to relate to changes perceived visually. The features extracted from images may be useful for objective and reliable discrimination samples with the use of, for example, machine-learning algorithms. It can be of great practical importance for cultivar discrimination, detection of species, disease, or evaluation of the plant quality [17,32]. The exemplary results of previous studies available in the literature indicated the usefulness of both the textural and geometric features of images for cultivar discrimination. The results obtained for sweet cherries [33] revealed that there are features of endocarp images of sweet cherries that allow distinguishing cultivars with an accuracy of 100%. The complete discrimination (100%) of endocarp of 'Kordia' vs. 'Lapins' and 'Kordia' vs. 'Büttner's Red' was determined in the case of models built for several sets of features, including combined textures selected from all channels R, G, B, L, a, b, X, Y, Z, separate sets including textures selected for individual color spaces RGB, Lab, and XYZ, individual channels G, L and Y and for models combining selected textural and geometric features for all applied classifiers (Naive Bayes from Bayes, Logistic from Functions, Multi Class Classifier from Meta, PART from Rules, and LMT from Trees). All three cherry endocarp cultivars of 'Kordia', 'Lapins' and 'Büttner's Red' were distinguished with the correctness of up to 98% for the model built based on a set of combined selected texture and geometric features using the Logistic algorithm [34]. In the case of distinguishing pits of the sour cherry cultivars of 'Debreceni botermo', 'Kelleris', 'Łutówka', and 'Nefris', an accuracy of 96.25% was obtained for the model including the selected textures from images converted to all used color channels (R, G, B, L, a, b, X, Y, Z) built using Multilayer Perceptron from the group of Functions. In the case of the analysis carried out for pairs of cultivars, it was confirmed that the 'Łutówka' pits were completely distinguished (100%) from the other cultivars for each pair, each model (for selected textures from all channels, for individual color spaces and color channels) and each used machine learning algorithm [18]. With a sufficiently large amount of data, it may be a good idea to use deep learning as a subset of machine learning to process data. Deep learning can ensure the high accuracy of discrimination as a result of the training and inference phase using high computational and storage requirements. The training process is computationally intensive and space-consuming. The complexity of the data models makes training quite expensive. Additionally, the cost to the users can raise because of the need of using costly graphic user interfaces and machines [35]. Therefore, further experiments with more data can be carried out to compare the usefulness of traditional machine learning and deep learning to distinguish different red onion samples.

6. Conclusions

The selected fluorescence spectroscopic data and machine-learning algorithms were used to distinguish bulbs of red onion belonging to 'Asenovgradska kaba' cultivar and the red breeding line. The obtained accuracy reached 100%. The algorithms of LMT, Multilayer Perceptron, Naive Bayes, Logit Boost, and LWL proved to be the most effective. It confirmed the usefulness of the applied approach to evaluate the differentiation of different cultivars and lines within the same species. Thus, the obtained results are very promising. The combination of fluorescence spectroscopy and machine-learning algorithms can be used in practice to discriminate different red onion cultivars and breeding lines. Furthermore, future research may include more red onion cultivars and lines as well as other types of onion including white and yellow onions. The developed procedure can also be used to discriminate different samples of other vegetables and fruit.

Author Contributions: Conceptualization, K.S., M.F.A., V.S. and S.G.; methodology, K.S., V.S. and S.G.; software, K.S.; validation, K.S. and M.F.A.; formal analysis, K.S. and V.S.; investigation, K.S., V.S. and S.G.; resources, V.S. and S.G.; data curation, K.S., V.S. and S.G.; writing—original draft preparation, K.S. and V.S.; writing—review and editing, K.S., M.F.A. and V.S.; visualization, K.S., M.F.A., V.S. and S.G.; supervision, K.S. and V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chadorshabi, S.; Hallaj-Nezhadi, S.; Ghasempour, Z. Red onion skin active ingredients, extraction and biological properties for functional food applications. *Food Chem.* **2022**, *386*, 132737. [CrossRef] [PubMed]
2. Stoica, F.; Condurache, N.N.; Aprodu, I.; Andronoiu, D.G.; Enachi, E.; Stănciuc, N.; Bahrim, G.E.; Croitoru, C.; Râpeanu, G. Value-added salad dressing enriched with red onion skin anthocyanins entrapped in different biopolymers. *Food Chem. X* **2022**, *15*, 100374. [CrossRef]
3. Fossen, T.; Andersen, Ø.M. Anthocyanins from red onion, *Allium cepa*, with novel aglycone. *Phytochemistry* **2003**, *62*, 1217–1220. [CrossRef]
4. Collings, D.A. Anthocyanin in the Vacuole of Red Onion Epidermal Cells Quenches Other Fluorescent Molecules. *Plants* **2019**, *8*, 596. [CrossRef] [PubMed]
5. Dairi, S.; Dahmoune, F.; Belbahi, A.; Remini, H.; Kadri, N.; Aoun, O.; Bouaoudia, N.; Madani, K. Optimization of microwave extraction method of phenolic compounds from red onion using response surface methodology and inhibition of lipoprotein low-density oxidation. *J. Appl. Res. Med. Aromat. Plants* **2021**, *22*, 100301. [CrossRef]
6. Corea, G.; Fattorusso, E.; Lanzotti, V.; Capasso, R.; Izzo, A.A. Antispasmodic saponins from bulbs of red onion, *Allium cepa* L. var. *Tropea*. *J. Agric. Food Chem.* **2005**, *53*, 935–940. [CrossRef] [PubMed]
7. Pérez-Gregorio, R.M.; García-Falcón, M.S.; Simal-Gándara, J.; Rodrigues, A.S.; Almeida, D.P.F. Identification and quantification of flavonoids in traditional cultivars of red and white onions at harvest. *J. Food Compos. Anal.* **2010**, *23*, 592–598. [CrossRef]
8. Rodrigues, A.S.; Pérez-Gregorio, M.R.; García-Falcón, M.S.; Simal-Gándara, J.; Almeida, D.P. Effect of post-harvest practices on flavonoid content of red and white onion cultivars. *Food Control* **2010**, *21*, 878–884. [CrossRef]
9. Rodrigues, A.S.; Pérez-Gregorio, M.R.; García-Falcón, M.S.; Simal-Gándara, J.; Almeida, D.P.F. Effect of meteorological conditions on antioxidant flavonoids in Portuguese cultivars of white and red onions. *Food Chem.* **2011**, *124*, 303–308. [CrossRef]
10. Brahimi, A.; El Ouardi, M.; Kaouachi, A.; Boudboud, A.; Hajji, L.; Hajjaj, H.; Mazouz, H. Characterization of the Biochemical Potential of Moroccan Onions (*Allium cepa* L.). *Int. J. Food Sci.* **2022**, *2022*, 2103151. [CrossRef]
11. Mahesh, B. Machine Learning Algorithms—A Review. *Int. J. Sci. Res. IJSR* **2020**, *9*, 381–386.
12. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [CrossRef]
13. Ropelewska, E.; Slavova, V.; Sabanci, K.; Aslan, M.F.; Cai, X.; Genova, S. Discrimination of onion subjected to drought and normal watering mode based on fluorescence spectroscopic data. *Comput. Electron. Agric.* **2022**, *196*, 106916. [CrossRef]
14. Slavova, V.; Ropelewska, E.; Sabanci, K.; Aslan, M.F.; Nacheva, E. A comparative evaluation of Bayes, functions, trees, meta, rules and lazy machine learning algorithms for the discrimination of different breeding lines and varieties of potato based on spectroscopic data. *Eur. Food Res. Technol.* **2022**, *248*, 1765–1775. [CrossRef]
15. Pourdarbani, R.; Sabzi, S.; Kalantari, D.; Karimzadeh, R.; Ilbeygi, E.; Arribas, J.I. Automatic non-destructive video estimation of maturation levels in Fuji apple (*Malus Malus pumila*) fruit in orchard based on colour (Vis) and spectral (NIR) data. *Biosyst. Eng.* **2020**, *195*, 136–151. [CrossRef]

16. Zhang, M.; Zhang, B.; Li, H.; Shen, M.; Tian, S.; Zhang, H.; Ren, X.; Xing, L.; Zhao, J. Determination of bagged 'Fuji' apple maturity by visible and near-infrared spectroscopy combined with a machine learning algorithm. *Infrared Phys. Technol.* **2020**, *111*, 103529. [CrossRef]
17. Ropelewska, E.; Rutkowski, K.P. Differentiation of peach cultivars by image analysis based on the skin, flesh, stone and seed textures. *Eur. Food Res. Technol.* **2021**, *247*, 2371–2377. [CrossRef]
18. Ropelewska, E. Classification of the pits of different sour cherry cultivars based on the surface textural features. *J. Saudi Soc. Agric. Sci.* **2021**, *20*, 52–57. [CrossRef]
19. Ropelewska, E.; Szwejda-Grzybowska, J. A comparative analysis of the discrimination of pepper (*Capsicum annuum* L.) based on the cross-section and seed textures determined using image processing. *J. Food Process Eng.* **2021**, *44*, e13694. [CrossRef]
20. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Elsevier: San Francisco, CA, USA, 2005; 558p.
21. Bouckaert, R.R.; Frank, E.; Hall, M.; Kirkby, R.; Reutemann, P.; Seewald, A.; Scuse, D. *WEKA Manual for Version 3-9-1*; The University of Waikato: Hamilton, New Zealand, 2016.
22. Frank, E.; Hall, M.A.; Witten, I.H. The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Fourth Edition 2016. Available online: https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf (accessed on 3 July 2022).
23. Cordon, G.; Andrade, C.; Barbara, L.; Romero, A.M. Early detection of tomato bacterial canker by reflectance indices. *Inf. Process. Agric.* **2022**, *9*, 184–194. [CrossRef]
24. Li, X.; Tsuta, M.; Hayakawa, F.; Nakano, Y.; Kazami, Y.; Ikehata, A. Estimating the sensory qualities of tomatoes using visible and near-infrared spectroscopy and interpretation based on gas chromatography–mass spectrometry metabolomics. *Food Chem.* **2021**, *343*, 128470. [CrossRef] [PubMed]
25. Kim, D.S.; Lee, D.U.; Choi, J.H.; Kim, S.; Lim, J.H. Prediction of carotenoid content in tomato fruit using a fluorescence screening method. *Postharvest Biol. Technol.* **2019**, *156*, 110917. [CrossRef]
26. Qin, J.; Chao, K.; Kim, M.S. Nondestructive evaluation of internal maturity of tomatoes using spatially offset Raman spectroscopy. *Postharvest Biol. Technol.* **2012**, *71*, 21–31. [CrossRef]
27. Hu, Q.; Sellers, C.; Kwon, J.S.I.; Wu, H.J. Integration of surface-enhanced Raman spectroscopy (SERS) and machine learning tools for coffee beverage classification. *Digit. Chem. Eng.* **2022**, *3*, 100020. [CrossRef]
28. Smeesters, L.; Magnus, I.; Virte, M.; Thienpont, H.; Meulebroeck, W. Potato quality assessment by monitoring the acrylamide precursors using reflection spectroscopy and machine learning. *J. Food Eng.* **2021**, *311*, 110699. [CrossRef]
29. Bertani, F.R.; Businaro, L.; Gambacorta, L.; Mencattini, A.; Brenda, D.; Di Giuseppe, D.; De Ninno, A.; Solfrizzo, M.; Martinelli, E.; Gerardino, A. Optical detection of aflatoxins B in grained almonds using fluorescence spectroscopy and machine learning algorithms. *Food Control* **2020**, *112*, 107073. [CrossRef]
30. Yan, S.; Wang, S.; Qiu, J.; Li, M.; Li, D.; Xu, D.; Li, D.; Liu, Q. Raman spectroscopy combined with machine learning for rapid detection of food-borne pathogens at the single-cell level. *Talanta* **2021**, *226*, 122195. [CrossRef] [PubMed]
31. Berghian-Grosan, C.; Magdas, D.A. Raman spectroscopy and machine-learning for edible oils evaluation. *Talanta* **2020**, *218*, 121176. [CrossRef] [PubMed]
32. Wu, X.; Zhao, Z.; Tian, R.; Niu, Y.; Gao, S.; Liu, H. Total synchronous fluorescence spectroscopy coupled with deep learning to rapidly identify the authenticity of sesame oil. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *244*, 118841. [CrossRef] [PubMed]
33. Ropelewska, E.; Sabanci, K.; Aslan, M.F. Discriminative Power of Geometric Parameters of Different Cultivars of Sour Cherry Pits Determined Using Machine Learning. *Agriculture* **2021**, *11*, 1212. [CrossRef]
34. Ropelewska, E. The Application of Machine Learning for Cultivar Discrimination of Sweet Cherry Endocarp. *Agriculture* **2021**, *11*, 6. [CrossRef]
35. Mamat, N.; Othman, M.F.; Abdoulghafor, R.; Belhaouari, S.B.; Mamat, N.; Mohd Hussein, S.F. Advanced Technology in Agriculture Industry by Implementing Image Annotation Technique and Deep Learning Approach: A Review. *Agriculture* **2022**, *12*, 1033. [CrossRef]