

# Trees' Condition Data Analysis Based on Drone Monitoring and Machine Learning Technology

Alina Dmytriv<sup>1</sup>, Victoria Vysotska<sup>2</sup>[0000-0001-6417-3689], Petro Kravets<sup>3</sup>[0000-0001-6124-3995],  
Ihor Karpov<sup>4</sup>[0000-0003-4885-5078], and Michael Emmerich<sup>5</sup>[0000-0002-7342-2090]

<sup>1-4</sup>Lviv Polytechnic National University, Lviv, Ukraine

<sup>5</sup>Leiden University, Leiden, The Netherlands

alinadmutriv@gmail.com<sup>1</sup>, Victoria.A.Vysotska@lpnu.ua<sup>2</sup>

**Abstract.** Goal is to create a system that will be analyze trees' condition using results from scanning and other obtained data and define options for damage detection. The project aim is the improvement of city management efficiency based on development of decision-making support systems according to the results of monitoring and analysis of urban environment parameters. This paper elaborates scientific and journalistic materials that include studies of tree monitoring, forest management and available statistics on the state of forests in Ukraine. A system of analogues (Znaydeno, LiDAR, and DJI drones) is searched and their strengths and weaknesses were described. In addition, there is a demonstration of implementation of the software and the application of machine learning. After that, an experimental approbation with data visualization is carried out and the analysis of the obtained research results was carried out, which performed the classification of trees according to their diseases.

**Keywords:** Tree Monitoring. Drones. Classification. Decision tree. Computer Science.

## 1 Introduction

Degradation of forests in Ukraine today is one of the biggest environmental problems. Deforestation and imperfect management of forestry activities lead to a catastrophic reduction in forest area. This significantly increases the risk of floods, deteriorating soil quality and fertility, reduces biodiversity, and destroys areas of rare species of flora and fauna, the loss of valuable natural ecosystems and the negative impact on the climate in Ukraine. To improve the condition of green areas, foresters still use the old methods of forest management. They inspect each tree for disease and damage, manually measure the diameter of the trunk and mark them. This work requires a lot of time and manpower. It is the latest technologies that will help speed up the production of drugs and increase the availability of up-to-date data on the state of the forest fund. Because this is an important prerequisite for effective management of forestry production. With the further development of technology, everyone will be able to determine the condition of

Copyright © 2020 for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

the tree. In addition, it is to evaluate treatment or felling to improve the sanitary condition of greenery. Obviously, this will increase the greening of the environment and prevent environmental disasters.

## **2 Analytical Review of Literary and Other Sources**

Nowadays, several studies have conducted on tree monitoring. Most studies focus on the development of new cost-effective methods of forest inventory using aerospace technologies. One of such studies is the use of a forest management database and satellite images of Sentinel-2 to decipher the species composition of forest stands. In this work, it used the technique of combining a set of satellite images obtained during one growing season in the form of a cloudless seasonal composite image and an algorithm, which provides for channel-by-channel analysis of Sentinel-2 satellite images. This study considers only the data of remote sensing of the Earth, but such data do not provide complete information about the state of greenery [1].

Currently, some analogue systems partially solve this problem. Such analogues include LiDAR and Znaydeno. Now scientists are working on a new technique: laser scanning of forests. This methodology identifies more accurate carbon stocks in forests, which is an important contribution to carbon monitoring and global climate policy. Laser measurements are used to accurately determine the size of the tree, its density and from this, you can find out the mass of wood, as the result is a three-dimensional spatial model. This technology is called LiDAR (Light Detection and Radar). The data obtained by this technology provide 90% accuracy. The technology is already actively used in architecture, mining and oil and gas industry, energy [2].

Another system in this industry is the startup Znaydeno, which can monitor deforestation and promptly notify. This startup already has the use of drones, which record the facts of crimes. Znaydeno combines the latest software development, remote sensing information and aerial photography of UAVs. This combination makes it possible to effectively monitor the state of natural resources and control their use. In this way, regulators, the public, activists and journalists will be able to see where the felling is taking place [3]. In addition, there are unmanned aerial vehicles for scanning the area and detecting damage, but this development is only for agricultural activities, not forestry. Some of these devices are DJI T16 and DJI Phantom 4 made in China, which is presented on the Ukrainian market [4].

## **3 The Choice of Design Methodology and Work Planning**

System analyzes trees' condition of researched object such as forest area, park territory, smallholding and others. System find options for improving trees' condition and do collection and analyzing data. The main processes that system does:

- Multi and hyperspectral, laser cameras scan the territory [5-10];
- Compares and analyses obtained data from result of scanning and client's previous researches and statistics [11-16];

- Defines trees' diseases, pests and others damages;
- Find options for improving trees' condition and damage prevention options for trees;
- Generate a report of the analysis and recommendations.

External entities that interact with the system are:

- An entrepreneur who has purchased a system to monitor the condition of his own facility, he can also be a user of the system;
- A user who uses the system and performs the relevant operations to study the object of the entrepreneur;
- A special commission that formulates recommendations for the system to improve the condition of green areas;
- An analyst who formulates rules for analysing the condition of green areas.

A use case diagram is constructed for a general idea of the functional purpose of the system (Fig. 1):

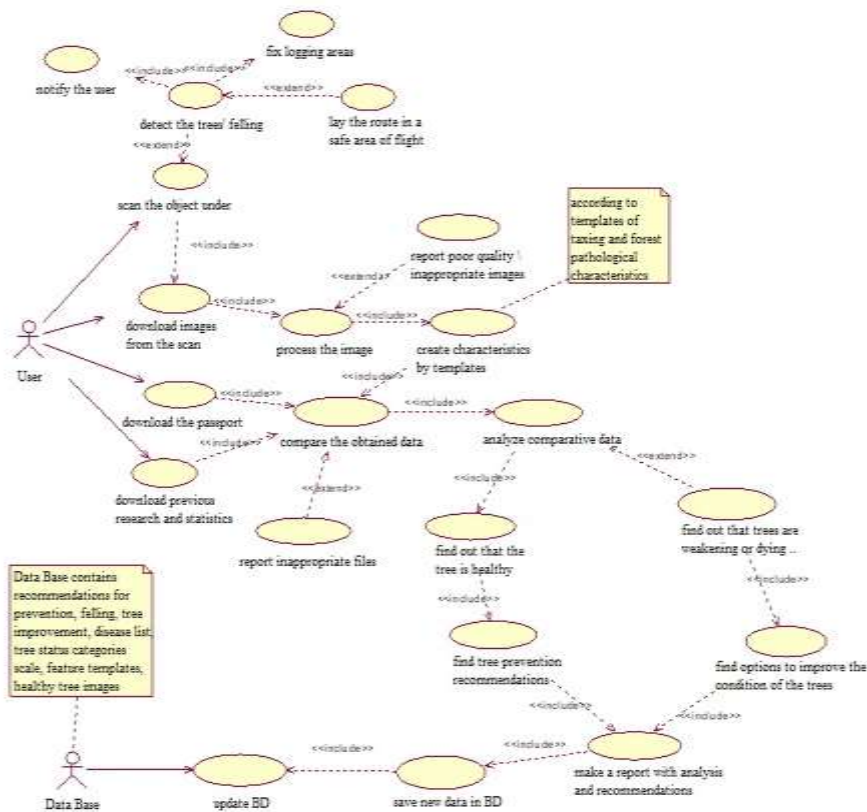


Fig. 1. Use Case diagram.

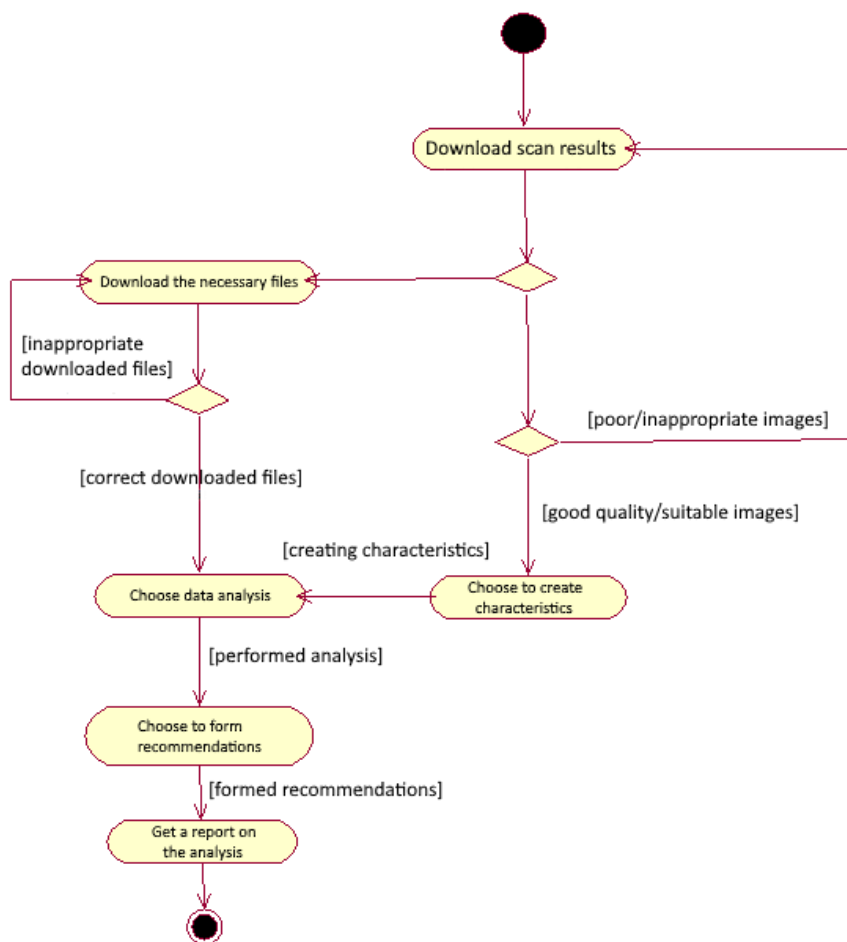
#### Actors:

- A PC user is an employee who will use the personal computer to analyze and formulate recommendations for the object under study, as well as to scan the object under study, using spectral cameras located on the drone;
- Database is exists before using the system, it contains the following data: a scale of spectral features and parameters for image processing, images of healthy trees, templates of taxation and forest pathological characteristics, a list of diseases, a scale of categories of tree conditions, recommendations for prevention and improvement of green plantings, rules of analysis. During the use of the system, the database is updated and supplemented with new collected and analysed information.

#### Use Cases:

- Various special cameras scan the object under study is the territory of the object and images of these objects are created.
- Detect tree felling is the drone detects tree felling.
- Notify user is a message is sent to the user that the wood is being felled.
- Record felling area is the area where the trees are cut down is recorded and the data is sent to the user.
- Pave the route to a safe flight area is in case of danger to the drone, for example, the drone is fired, the drone flies to a safe area.
- Download the scan result images is the received scan images are loaded into the program.
- Download object passport is uploads a file with the data of the object under study, namely its passport.
- Download previous researches and statistics is download files with object research and its statistics.
- Process the images is image processing is performed, object tree data such as height, elemental composition, chemical composition, etc. are determined.
- Report poor quality/inappropriate images is notify the user that the downloaded images have poor quality or are inappropriate.
- Create characteristics according to templates is taxation and forest pathological characteristics of trees are created according to templates, they have the data obtained from the previous stage.
- Compare the obtained data is the comparison of previous researches, statistical data and the received characteristics is carried out.
- Report inappropriate files is notify the user that the downloaded files are incorrect.
- Analyse comparable data is there is an analysis of comparable data in the previous stage according to the analysis rules for the detection of damage written by analysts.
- Identify signs of a healthy tree is the analysis in the previous step shows that this test object has no damage and is completely healthy.

- Detect weakening/dying/fresh/old dryness is if the analysis in the previous step shows that the trees in this object are damaged according to the scale of tree condition categories.
- Find recommendations for tree prevention is find recommendations from the recommendations of the special commission to prevent further tree damage.
- Find options to improve the trees' condition is search the options and recommended actions to restore the sanitary trees' condition according to the recommendations of the special commission.



**Fig. 2.** Activity diagram.

- Generate a report with the performed analysis and recommendations is form the documentation containing the processed images, the carried out analysis

and conclusions from it, recommendations on the revealed damage or prevention.

- Save data in the database is - save processed, compared and analysed in the database.
- Update database is the database is updated to obtain new data.

An activity diagram is constructed, which is a means of describing behavior in the form of a graph of activity (Fig. 2). Activities:

- Download required files is user uploads object passport, statistical observations and previous researches to the system.
- Download scan results is the user loads the creation image during multi-and hyperspectral, laser scanning.
- Choose to create characteristics is if the images are high quality and appropriate, then the user chooses to create characteristics and then receives the taxation and forest pathological characteristics. In addition, if the images are poor quality and inappropriate, then return to the activity "Download scan results".
- Select data analysis - if the data is loaded correctly and the characteristics are created; the user can analyse the data. If the downloaded data is not suitable, it returns to the "Download required data" activity.
- Choose to make recommendations - when the analysis is done, recommendations are made to improve the trees' condition or prevention.
- Receive an analysis report - if the recommendations are formed, then the user receives a report of all work, which includes processed images, creation of characteristics, compared and analysed data. Then the work ends.

#### **4 The Choice of Implementation Means and Design the System**

The idea is to create the proposed system as a computer program. It will need to process tabular data that will be created in files with the .csv extension. The pandas library was used to create a database for the system.

The program should analyze the trees' characteristics obtained from scanned images. These data contain general information about the tree, namely the date of the inspection, tree id, height, diameter, age, bonitet, kind and location of the tree, soil, relief and forest type. Then the data contains detailed information about specific parts of the tree, such as the trunk, crown, root, pine needles and cones. The classification method will be used here to find trees with diseases. Using this method, the program will split a given sample of data into subsets to separate healthy trees from sick ones. The program will select the appropriate parameters depending on the specific data set and the purpose of using the results. After that, there will be groups of similar data trees in the sample. The formed groups will be compared on signs and then the corresponding diseases will be defined. Thus, there are groups of trees with their diseases. Then their identifier will determine their location in the study area.

To implement the classification, we chose the decision tree method. This method is one of the predictive modelling approaches used in statistics, data mining and machine

learning. It uses a decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves).

For decision tree method it is important to decide which attribute/feature should be placed at the root node, which features will act as internal nodes or leaf nodes. To decide this, and how to split the tree, we use splitting measures like Gini Index and Information Gain. With more than one attribute taking part in the decision-making process, it is necessary to decide the relevance and importance of each of the attributes, thus placing the most relevant at the root node and further traversing down by splitting the nodes. Information Gain is used to determine which feature/attribute gives us the maximum information about a class. It is based on the concept of entropy, which is the degree of uncertainty, impurity or disorder. It aims to reduce the level of entropy starting from the root node to the leaf nodes. Gini index measures the degree or probability of a particular variable being wrongly classified when it is randomly chosen. If all the elements belong to a single class, then it can be called pure. The degree of Gini index varies between 0 and 1, where 0 denotes that all elements belong to a certain class or if there exists only one class, and 1 denotes that the elements are randomly distributed across various classes. A Gini Index of 0.5 denotes equally distributed elements into some classes [6].

The Python programming language is chosen to implement the tree condition monitoring system. The following Python libraries are used: PyQt5, statistics, pandas, csv, sklearn.model\_selection, sklearn.tree, sklearn.metrics and graphviz. The PyQt5 library and the Qt Designer graphical designer are used to develop the graphical user interface, which is a GUI development tool on the Qt library.

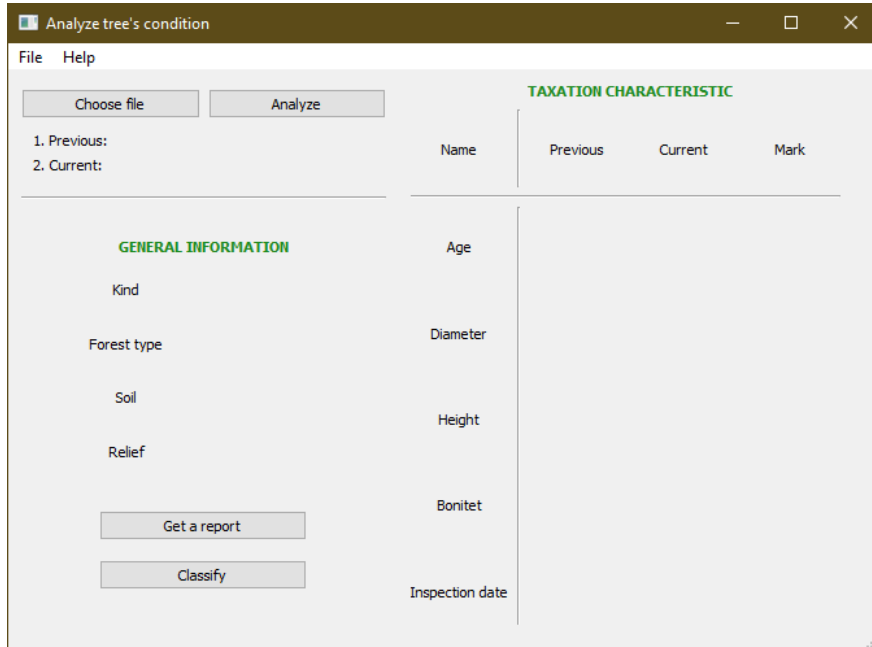
The input data of the program are downloaded taxation characteristics of the previous and current research. The initial data of the program is a report with an analysis and derived research indicators and their evaluation.

The created program is that the user downloads files with data about previous and current research. It is important to note that the user must first download the file with the data of the previous study, and then the file with the data of the current study. Files must be with a .csv extension. Then the data from the received files is analyzed and displayed with the evaluation, which is made according to the compared data. Then it is possible to view the report, which describes the detailed analysis of the data. You can also download this file on a personal computer. The program also includes instructions for use and details about the program. The following is each step of the program.

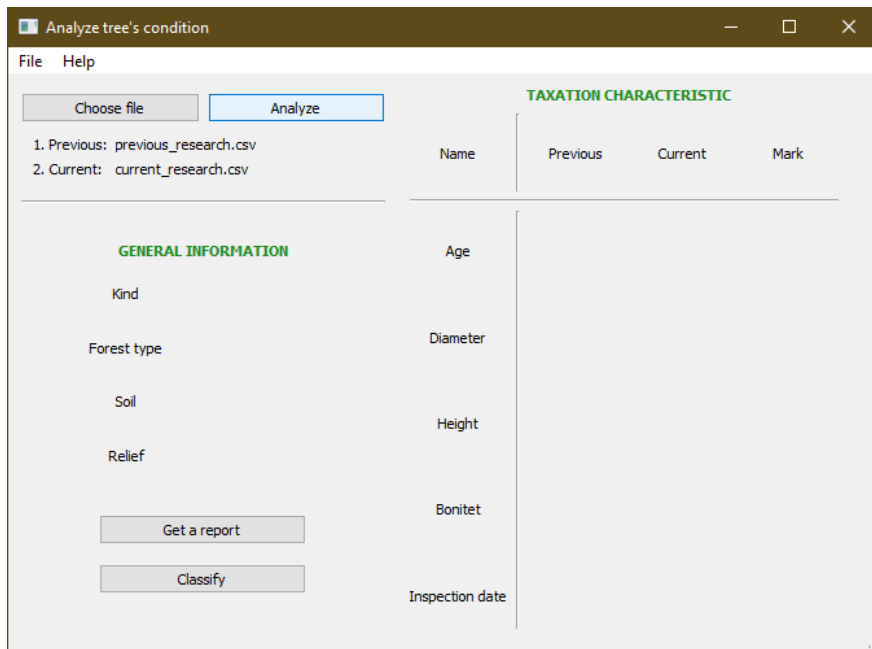
Two files were created for the program with the prepared data templates of the previous and current research. The file with the data of the previous study is presented in Fig. 3 a), while the file with the data of the current study is presented in Fig. 3 b).







**Fig. 4.** Main program window.



**Fig. 5.** Download the files.

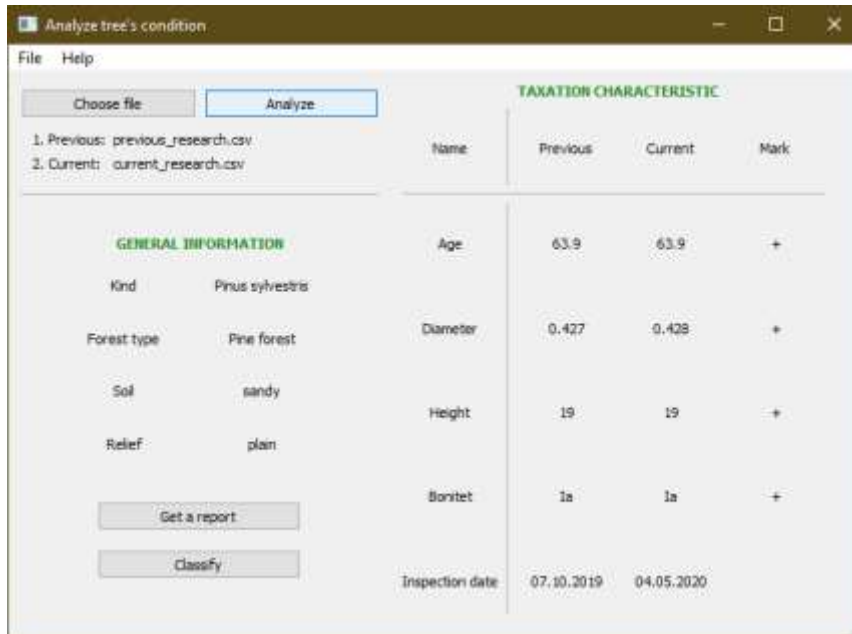


Fig. 6. Done analysis.

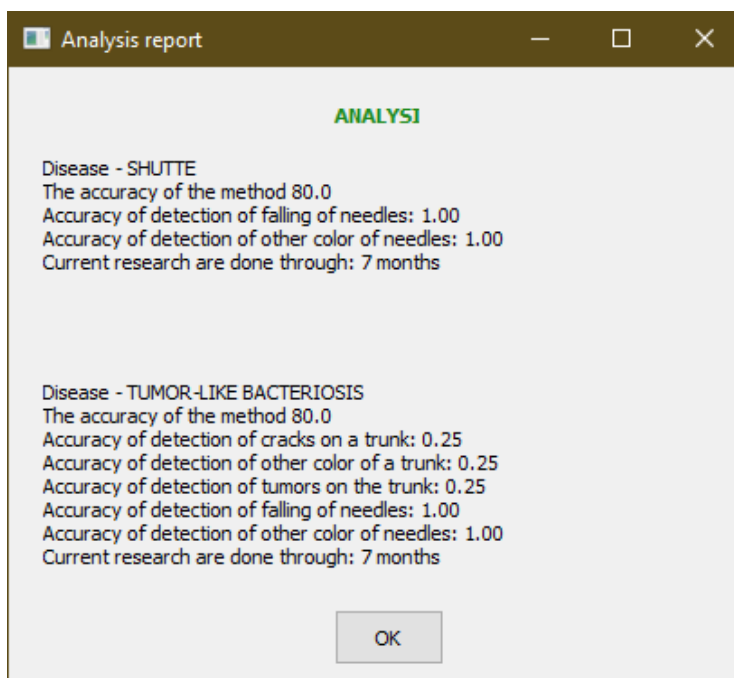


Fig. 7. The report of analysis.

Classification report

**CLASSIFICATION**

Disease - SHUTTE  
Accuracy : 80.0

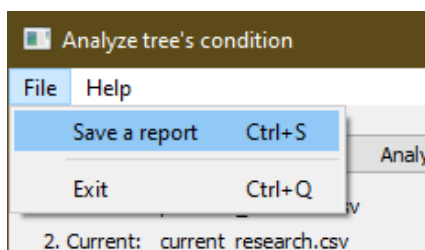
	precision	recall	f1-score	support
needle_falling	1.00	0.25	0.40	4
needle_color_difference	1.00	1.00	1.00	5
micro avg	1.00	0.67	0.80	9
macro avg	1.00	0.62	0.70	9
weighted avg	1.00	0.67	0.73	9
samples avg	0.33	0.23	0.27	9

Disease - TUMOR-LIKE BACTERIOSIS  
Accuracy : 80.0

	precision	recall	f1-score	support
trunk_crack	0.25	1.00	0.40	1
trunk_color_difference	0.25	1.00	0.40	1
needle_falling	1.00	0.25	0.40	4
needle_color_difference	1.00	1.00	1.00	5
trunk_tumor	0.25	1.00	0.40	1
micro avg	0.50	0.75	0.60	12
macro avg	0.55	0.85	0.52	12
weighted avg	0.81	0.75	0.65	12
samples avg	0.18	0.23	0.20	12

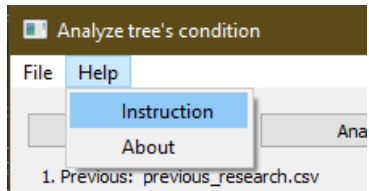
OK

**Fig. 8.** A report of the classification.

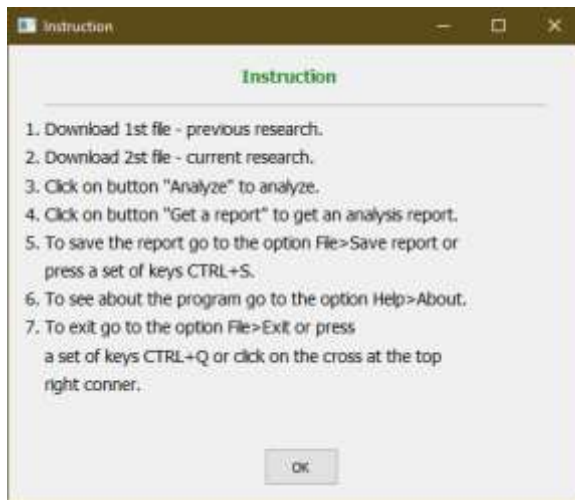


**Fig. 9.** Saving the report.

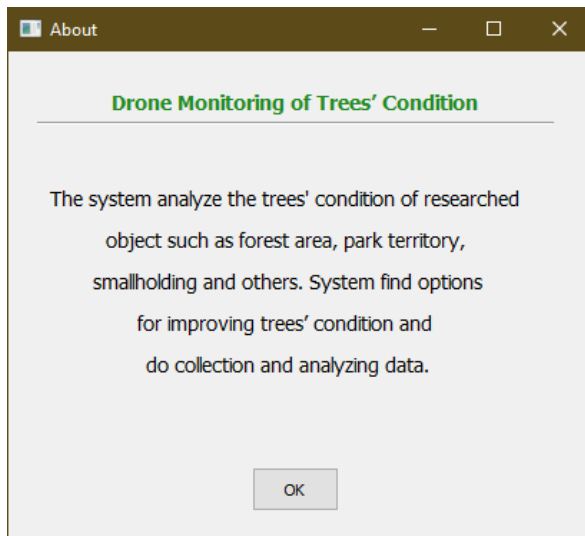
To find out information about the program and instructions for its use, the user goes to the option "Help> Instructions for use" and "Help> About the program" (Fig. 10), then a new window with instructions (Fig. 11), and for details - then a new window with information about the program appears (Fig. 12).



**Fig. 10.** Help tab.



**Fig. 11.** User instruction.



**Fig. 12.** About the program.

After implementing the software product, the expected result was obtained, as all the tasks set before it were performed, which included downloading data, performing analysis and obtaining an analysis report of downloaded data. This software product works without delays and errors. In addition, in the future, this program can be improved and there can introduce new features that could facilitate and increase data analysis, as well as apply more advanced machine learning for better analysis.

## 5 Machine Learning Application

The decision tree method for classifying trees according to their diseases in this article calculates the accuracy of detecting certain signs that belong to the corresponding pine disease. Since a table was created with the pines data of the certain area (Fig. 3 b). Some of the researched trees have diseases such as shutte and tumor-like bacteriosis. Target values for the decision tree classification are the sign of these diseases. The tabular data contains all the necessary properties for classification. Fig. 13 presents the characteristics of the tabular data.

```
Title: Taxation characteristic
Number of rows in the dataset: 50
Number of columns in the dataset: 42
Targets Distribution:
  20.0% trees have needle falling
  30.0% trees have needle color difference
  10.0% trees have trunk crack
  10.0% trees have trunk tumor
  10.0% trees have trunk color difference
  0.0% trees have mushrooms on the root
  0.0% trees have trunk angle inclination
```

**Fig. 13.** The characteristics of the tabular data.

Because of the method for each classification, we get the predicted values, the confusion matrix, the accuracy estimate and the result of the classification. The confusion matrix is used to understand the behavior of the prepared classifier over the test data set or data set validation. Accuracy assessment is used to calculate the accuracy of the prepared classifier, namely the accuracy of the forecasts made. The results of the classification contain the following calculations for each purpose:

- The precision of the classifier does not mark a negative sample as positive;
- recall - the ability of the classifier to find all positive samples;
- F-scope, which is the weighted average harmonic value of prediction and recall (reaches the best value in 1, and the worst - 0);
- Support that calculates the number of entries of each class in the test data set y.

As well as reported averages that include a macro average that calculates the score for each label and finds their unweighted average;

- Weighted average, which calculates the indicator for each label and finds their weighted average;
- The average value of the sample, which calculates the average value of each label;
- A micro-average that calculates true positive, false negative, and false positive indicators.

Firstly, we consider the demonstration of using this method to detect the disease of the shutte. The following signs recognize this disease: falling pine needles and discoloration of needles. These features will be used as targets for classification. In addition, it is important that this disease is detected in trees after 5 months and at a young age. To apply the method, we set goals - needle falling (needle\_falling) and needle color change (needle\_color\_difference), as well as transfer the entire database of the current study to the functions of the decision tree method. The obtained result is presented in Fig. 14 - 15.

```
Accuracy : 80.0
Report :
```

	precision	recall	f1-score	support
needle_falling	1.00	0.25	0.40	4
needle_color_difference	1.00	1.00	1.00	5
micro avg	1.00	0.67	0.80	9
macro avg	1.00	0.62	0.70	9
weighted avg	1.00	0.67	0.73	9
samples avg	0.33	0.23	0.27	9

**Fig. 14.** The method accuracy and the classification report for the shutte.

```
Predicted values:
[[0 0]
 [0 0]
 [0 0]
 [0 0]
 [0 1]
 [1 1]
 [0 0]
 [0 0]
 [0 0]
 [0 1]
 [0 1]
 [0 1]
 [0 0]
 [0 0]
 [0 1]
 [0 0]]
Confusion Matrix:
[[11  3]
 [ 0  1]]
```

**Fig. 15.** Predicted data and confusion matrix for the shutte.

Secondly, we consider the demonstration of using this method to detect the disease of tumor-like bacteriosis. The following signs recognize this disease: falling pine needles, discoloration of needles, cracks on the trunk (trunk\_crack), discoloration of the trunk (trunk\_color\_difference) and tumors on the trunk (trunk\_tumor). This disease already has more features that will be used as targets for classification. The obtained result is presented in Fig. 16 - 17.

```

Accuracy : 80.0
Report :

```

	precision	recall	f1-score	support
trunk_crack	0.25	1.00	0.40	1
trunk_color_difference	0.25	1.00	0.40	1
needle_falling	1.00	0.25	0.40	4
needle_color_difference	1.00	1.00	1.00	5
trunk_tumor	0.25	1.00	0.40	1
micro avg	0.50	0.75	0.60	12
macro avg	0.55	0.85	0.52	12
weighted avg	0.81	0.75	0.65	12
samples avg	0.18	0.23	0.20	12

**Fig. 16.** The method accuracy and the classification report for the tumor-like bacteriosis.

```

Predicted values:
[[0 0 0 0 0]
 [0 0 0 0 0]
 [0 0 0 0 0]
 [0 0 0 0 0]
 [1 1 0 1 1]
 [0 0 1 1 0]
 [0 0 0 0 0]
 [0 0 0 0 0]
 [0 0 0 0 0]
 [1 1 0 1 1]
 [1 1 0 1 1]
 [1 1 0 1 1]
 [0 0 0 0 0]
 [0 0 0 0 0]
 [1 1 0 1 1]
 [0 0 0 0 0]]
Confusion Matrix:
[[11  0]
 [ 3  1]]

```

**Fig. 17.** Predicted data and confusion matrix for the tumor-like bacteriosis.

In addition, we will use the decision tree method to detect root rot disease. The researched pines do not have this disease, but we check how accurately this machine learning method will detect it. This disease has the following signs: trunk angle (trunk\_angle\_inclination), the mushrooms near the root (root\_mushrooms) and cracks on the trunk. we pass these features as goals to the method and get the following results. The obtained result is presented in Fig. 18 - 19.

```

Accuracy : 86.66666666666667
Report :

```

	precision	recall	f1-score	support
trunk_angle_inclination	0.00	0.00	0.00	0
root_mushrooms	0.00	0.00	0.00	0
trunk_crack	0.33	1.00	0.50	1
micro avg	0.33	1.00	0.50	1
macro avg	0.11	0.33	0.17	1
weighted avg	0.33	1.00	0.50	1
samples avg	0.07	0.07	0.07	1

**Fig. 18.** The method accuracy and the classification report for the root rot.

```

Predicted values:
[[0 0 1]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 1]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 1]
 [0 0 0]
 [0 0 0]
 [0 0 0]
 [0 0 0]]
Confusion Matrix:
[[12  2]
 [ 0  1]]

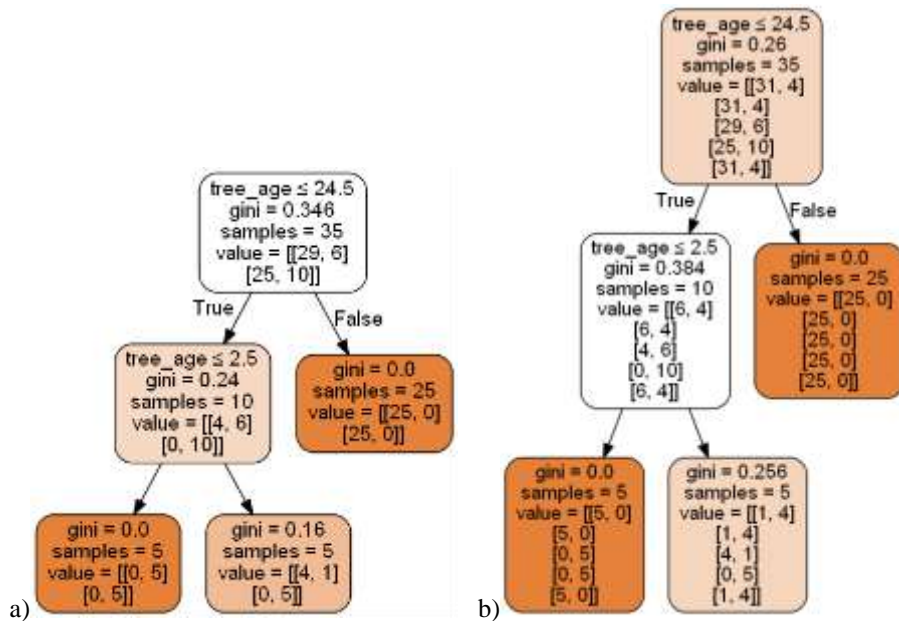
```

**Fig. 19.** Predicted data and confusion matrix for the root rot.

Since the chosen method of classification is a decision tree, so we visualized the results of the classification according to the previously described diseases in the form of a tree. This is a binary tree where each node represents a piece of data. Each node, which is not a leaf, divides its part of the data into two sections. The root node contains all the data (from the training set). Moreover, this classification tree predicts the probabilities of the class - the value of the node. The classes of this data set are [0, 1]. The resulting visualization of the tree contains the following features: the name of the column by which the data is divided;

- A Gini index equal to the degree or probability of misclassification of a variable;
- Variable samples (samples), which is equal to the number of samples on which the tree was trained; and the values, which provides a list of probabilities of belonging to a certain class of relevant goals.

We visualized and obtained a diagram in the form of a tree, which contains the result of classification of data according to diseases. Fig. 20a presents a solution tree for shutte. In addition, we visualized the confusion matrix (Fig. 21) and the classification report (Fig. 22).



**Fig. 20.** The visualization result of the classification according to a) the shutte and b) the tumor-like bacteriosis.



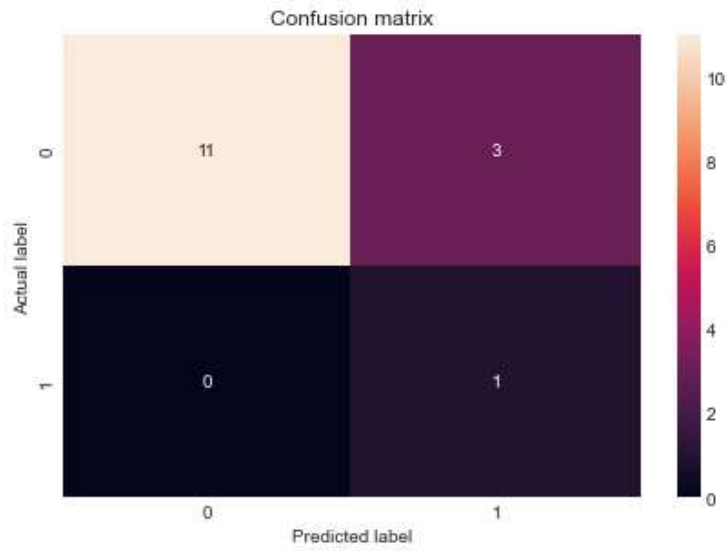


Fig. 21. Confusion matrix to detect shutte disease.

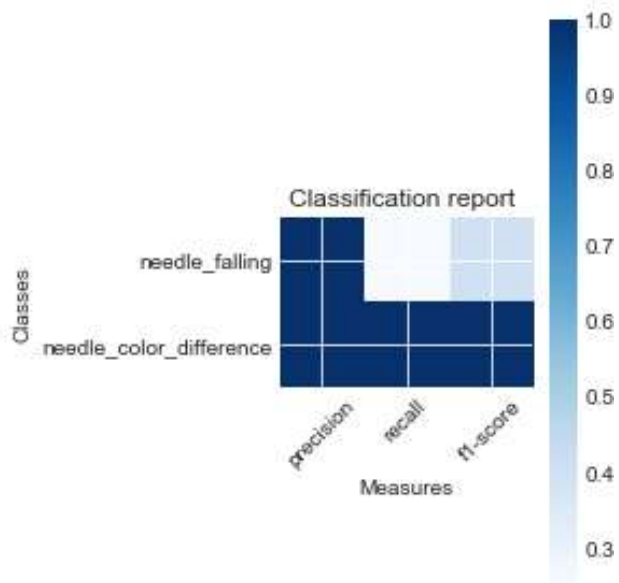
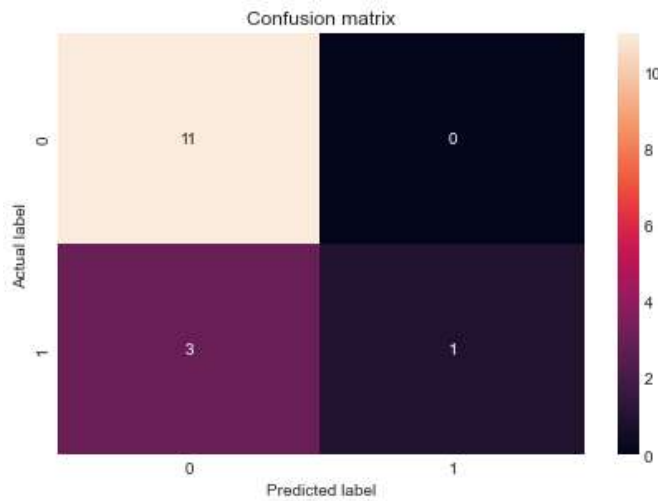
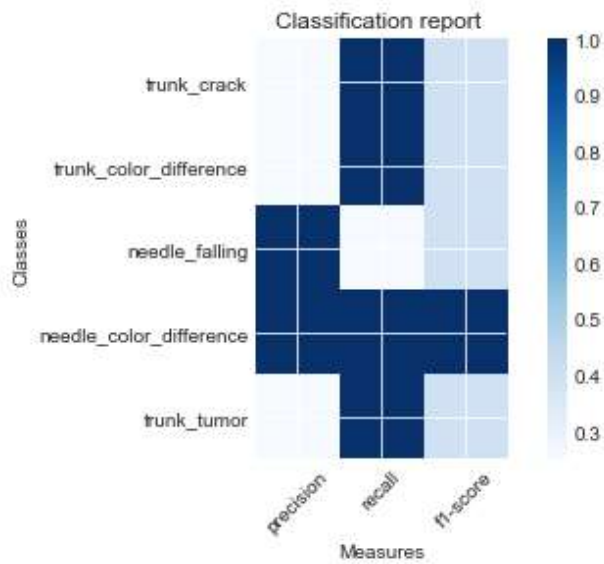


Fig. 22. Classification report to detect shutte disease.

The Fig. 20b presents a solution tree for the tumor-like bacteriosis. In addition, we visualized the confusion matrix (Fig. 23) and the classification report (Fig. 24). In addition, a sign of this disease is an increase in the diameter of the trunks, so we visualized information about the diameter of the trees in accordance previous and current research (Fig. 25).



**Fig. 23.** Confusion matrix to detect tumor-like bacteriosis disease.



**Fig. 24.** Classification report to detect tumor-like bacteriosis disease.

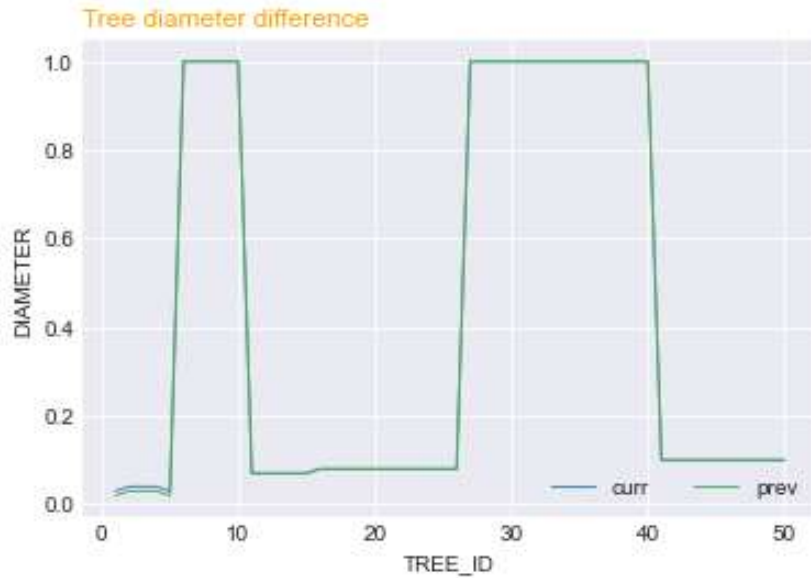


Fig. 25. The result of visualization of changes in tree diameter over 7 months.

The Fig. 26 presents a solution tree for root rot. In addition, we visualized the confusion matrix (Fig. 27) and the classification report (Fig. 28).

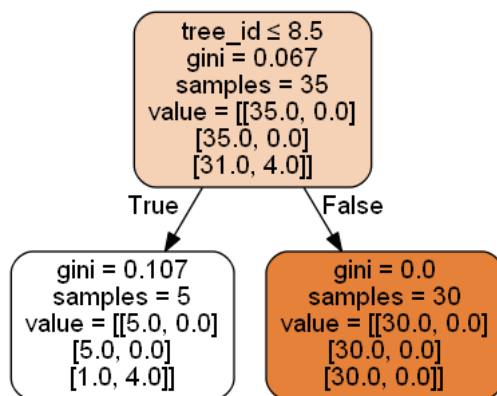


Fig. 26. The visualization result of the classification according to the root rot.

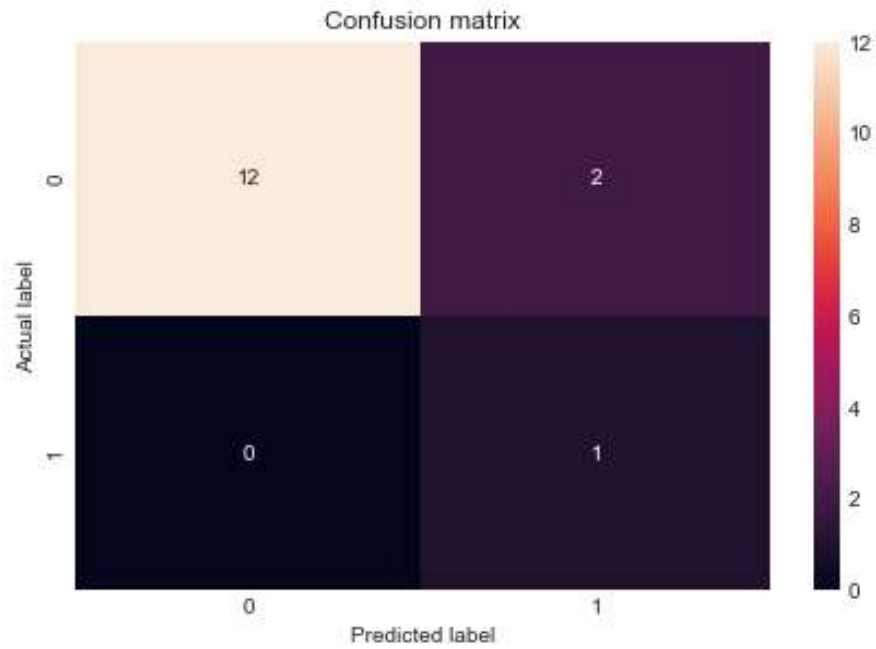


Fig. 27. Confusion matrix to detect root rot disease.

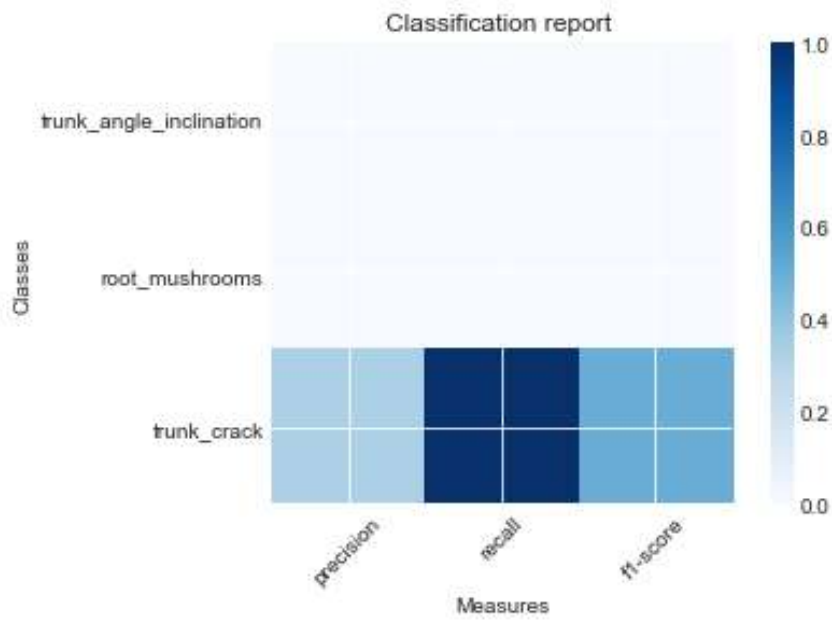


Fig. 28. Classification report to detect root rot disease.

After receiving all the classification results, we can analyze the study. We consider separately the results of the classification of each disease.

The obtained results of the disease are presented in Fig. 14 - 15, and its visualization is presented in Fig. 20 a. Firstly, we consider the results of the machine learning application. The matrix with predicted data contains 2 columns representing the whole classification and 15 rows. From the confusion matrix, we get the following information: 11 elements with leaf falling and discoloration of needles are defined as true positive, 3 elements - false positive, 1 element as false negative and no element as true negative. Thus, 13 elements out of 15 make the correct prediction. Therefore, the accuracy is 80%.

Now we analyze the classification results, which are presented in Fig. 14. The calculated indicators for the target "needle fall" are:

- 100% accuracy of the classifier, namely, that does not have any marked negative sample as positive; the ability of the classifier to find all positive samples is equal to 25%;
- The weighted average harmonic value of the F1-score is 40%; and there are 4 inputs of each class in the test dataset y.

While the calculated indicators for the target "discoloration of needles" are:

- Accuracy of the classifier, the ability of the classifier to find all positive samples and the weighted average harmonic value of F1-score are equal 100% and there are only 5 inputs of each class in the test dataset y.

Regarding the averages, the accuracy of the classifier for the average value is 33%, and for the other 3 is 100%. Thus, the average values differ from the other 3. The recall and F1-score are in accordance 23% and 27%. Recall and F1-score of micro-, macro and weighted average are in accordance 62-67% and 70-80%. The number of each class entries in the set of test data for all four indicators is 9. The calculated indicators are quite high, which shows a well-performed classification and detection of shutte in trees.

The visualization tree of the performed classification divides the data by their age (younger than 24.5 years and older trees, respectively), which is shown in Fig. 20a. The number of samples on which the tree is trained is 35 samples. The Gini index is 0.345, which means a fairly high level of correct classification. We consider the leaves of a tree. They represent 10 samples, in which the disease is detected. In 5 samples less than 2.5 years old, all the goals were accurately identified, namely the falling and discoloration of the needles. In 5 others samples older than 2.5 years, a sign of discoloration of the needle was detected, and only in 1 sample of them the needles falling. No one signs were detected in the other 25 samples.

Now let's move on to the next disease of tumor-like bacteriosis. The obtained results of the disease are presented in Fig. 16 - 17, and its visualization is presented in Fig. 20b. Firstly, we consider the results of the machine learning application. The matrix with predicted data contains 5 columns representing the whole classification, and 15 rows. We obtained the following information from the confusion matrix: 11 elements are defined as true positive, 3 elements - false negative, 1 element as true negative and

no element as false positive. Thus, 12 elements out of 15 make a correct prediction. From this, the accuracy is 80%.

Now we analyze the results of the classification, which are presented in Fig.16. The calculated indicators for the target of "needles falling" and "discoloration needles" are the same as for the previous disease. Regarding the target of "discoloration trunk", "crack in the trunk" and "tumors on the trunk", the calculated indicators for them are the same and are:

- 25% accuracy of the classifier, namely that has no marked negative sample as positive ;
- The ability of the classifier to find all positive samples is 100%; the weighted average harmonic value of the F1-scope is 40%;
- In addition, there is 1 inputs of each class in the test data set y.

Regarding the average indicators, the accuracy of the classifier for the average value is 18%, and for the other 3 is 50-81%. Yes, the average values are less than the other 3. The recall and F1 average are in accordance 23% and 20%. Recall and F1-score micro, macro and weighted average are in accordance 74-85% and 52-65%. The number of each class entries in the set of test data for all four indicators is 12. The calculated indicators are quite high, which shows a well-performed classification and detection of tumor-like bacteriosis in trees.

The visualization tree of the performed classification divides the data by their age (younger than 24.5 years and older trees, respectively), which is shown in Fig.20b. The number of samples on which the tree is trained is also 35 samples. The Gini index is 0.26, which means a high level of correct classification. We consider the leaves of a tree; they represent 10 samples, which revealed the disease tumor-like bacteriosis. In 5 samples less than 2.5 years old, only signs of falling and discoloration of the needles were detected. While 5 samples older than 2.5 years showed discoloration needles, only 1 sample of them showed a falling needles and 4 samples showed signs of discoloration of the trunk, the presence of cracks and tumors on the trunk. No other signs are detected in the other 25 samples.

Finally, we consider the disease of root rot, which is not in the researched pines. The obtained results of the disease are presented in Fig. 18 - 19, and its visualization is presented in Fig. 26. Firstly, we consider the results of the machine learning application. The matrix with predicted data contains 3 columns representing the whole classification and 15 rows. From the confusion matrix, we obtained the following information: 12 elements were defined as true positive, 2 elements - false positive, 1 element as true negative and no element as false negative. Thus, 13 elements out of 15 make the correct prediction. The method accuracy is 86.67%.

Now we analyze the results of the classification, which are presented in Fig.18. All indicators for the targets of "trunk angle inclination" and "mushrooms near the root" are calculated as 0%. Regarding the target of "crack on the trunk", the calculated indicators for it are:

- 33% accuracy of the classifier, namely, that has no marked negative pattern as positive; the ability of the classifier to find all positive samples is 100%;

- The weighted average harmonic value of the F1-score is 50%;
- In addition, there is 1 inputs of each class in the test data set y.

This indicates that these signs are present in trees and play an important role in the previous disease, and does not belong to this disease. As for the averages, for the average value they are all the same and are 7%. While the precision, recall and F1-score of micro, macro and weighted average are in accordance 11-33%, 33-100% and 17-50%. The number of each class entries in the set of test data for all four indicators is equal to 1. The calculated indicators are quite low, which shows a well-performed classification and no detection of root rot disease in trees.

The visualization tree of the performed classification divides the data by their identifier (less than 8.5 and correspondingly larger), which is shown in Fig.26. The number of samples on which the tree is trained is also 35 samples. The Gini index is 0.067, which means a very high level of correct classification. We consider the leaves of a tree. They represent 5 samples, which showed a sign of "cracks in the trunk", and 30 samples, which did not show any signs of disease.

## 6 Conclusion

Thus, the article presents a system for drones monitoring of trees' condition. To this end, research sources and comparable existing analogue systems, including LiDAR, Znaydeno, DJI T16 and DJI Phantom 4, are developed and their strengths and weaknesses were described. This proves that this topic is relevant and needs to be developed and researched more and more.

The system of drones monitoring of trees' condition is designed using IBM Rational Rose software. A Use Case diagram was constructed, depicting a large number of usage options and 2 actors - the Database and the PC User, and an activity diagram showing the behavior of the system on the user side.

The article also describes the implementation algorithms, methods, and implementation of the software. In the future, the program can be improved and enhanced.

The last section presented a demonstration of using the machine learning to classify 3 types of pine diseases: shutte, tumor-like bacteriosis and root rot. The necessary features for these diseases were described, which are used then as targets in the implementation of the decision tree method. The results of the performed classification method are shown in the figures and their derived indicators were described. In addition, the results of experimental testing of the implemented machine learning with using visual data are also provided. Namely, the creation of a decision tree diagram for each implemented classification. The diagram contains certain indicators, so their value is also described. After that, it was analyzed the obtained research results, namely the confusion matrix, the accuracy of the classifier, the classification report and the visualization. The following conclusions can be drawn from the analysis: the decision tree classification method is excellent for classifying trees according to different diseases. The method revealed shutte and tumor-like bacteriosis in 10 trees, which is the correct result, and did not reveal root rot, which was not present in the researched pines. This also confirms the accuracy of the classifier of this method, which is 80% and above.

## References

1. Georgian, M., Myroniuk, V.: Classification of tree species composition of forest stands using sentinel-2 satellite images. In: Forestry and horticulture, 11 (2017), [http://nbuv.gov.ua/UJRN/licgoc\\_2017\\_11\\_6](http://nbuv.gov.ua/UJRN/licgoc_2017_11_6)
2. Scanning trees will help them weigh and detect carbon content, <https://ukurier.gov.ua/uk/news/skanuvannya-derev-dopomozhe-yih-zvazhuvati-ta-viya/>
3. How big data and drones can save Ukrainian forests, [http://texty.org.ua/pg/blog/nar-text/read/76201/Jak\\_big\\_data\\_i\\_drony\\_mozhut\\_vratuvaty](http://texty.org.ua/pg/blog/nar-text/read/76201/Jak_big_data_i_drony_mozhut_vratuvaty)
4. Farmers in Ukraine are reluctant to buy industrial agro drones – expert, <https://superagronom.com/news/8229-agrariyi-v-ukrayini-ne-ohoche-kupuyut-promislovi-agrodroni--ekspert>
5. DJI presents a new agrodron for PPE, <https://superagronom.com/news/8254-dji-prezentuye-noviy-agrodron-dlya-vnesennya-zzr>
6. Gini Index For Decision Trees, <https://blog.quantinsti.com/gini-index/>
7. Peleshko D., Rak T., Lytvyn V., Vysotska V., Noennig J.: Drone monitoring system DROMOS of urban environmental dynamics. In: CEUR Workshop Proceedings, Vol-2565, 178–19. (2020)
8. Lytvyn, V., Kowalska-Styczen, A., Peleshko, D., Rak, T., Voloshyn, V., Noennig, J. R., Vysotska, V., Nykolyshyn, L., Pryshchepa, H.: Aviation Aircraft Planning System Project Development. In: Advances in Intelligent Systems and Computing IV, Springer, Cham, 1080, 315-348. (2020)
9. Lytvyn, V., Dmytriv, A., Berko, A., Alieksieiev, V., Basyuk, T., Noennig, J., Peleshko, D., Rak, T., Voloshyn, V.: Conceptual Model of Information System for Drone Monitoring of Trees' Condition. In: Computational Linguistics and Intelligent Systems, COLINS, CEUR workshop proceedings, Vol-2604, 695-714. (2020).
10. Alieksieiev, V., Markovych, B.: Implementation of UAV for environment monitoring of a Smart City with an airspace regulation by AIXM-format data streaming. In: Industry 4.0” – Scientific-Technical Union of Mechanical Engineering “Industry 4.0”, Sofia, Bulgaria, 5(2/2020), 90–93. (2020), <https://stumejournals.com/journals/i4/2020/2/90>
11. Lozynskyy, A.B., Romanyshyn, I.M., Rusyn B.P.: Intensity Estimation of Noise-Like Signal in Presence of Uncorrelated Pulse Interferences. In: Radioelectronics and Communications Systems, 62(5), 214-222. (2019)
12. Kosarevych, R.Y., Lutsyk, O.A., Kapshii, O.V., Rusyn B.P.: Random point patterns and bags of visual words for remote sensing imagery. In: Journal of Applied Remote Sensing, 13(3), 034521. (2019)
13. Rusyn, B., Lutsyk, O., Kosarevych, R., Varetsky, Y.: Automated Recognition of Numeric Display Based on Deep Learning. In: Proceedings of 2019 3rd International Conference on Advanced Information and Communications Technologies, AICT, 244-247. (2019)
14. Rusyn, B.P., Lutsyk, A.A., Kosarevych, R.Ya.: Modified architecture of lossless image compression based on FPGA for on - Board devices with linear CCD. In: Journal of Automation and Information Sciences, 51(2), 41-49. (2019)
15. Rusyn, B., Korniy, V., Lutsyk, O., Kosarevych, R.: Deep Learning for Atmospheric Cloud Image Segmentation. In: 11th International Scientific and Technical Conference on Electronics and Information Technologies, ELIT 2019-Proccedings, 188-191. (2019)
16. Kosarevych, R.J., Rusyn, B.P., Korniy, V.V., Kerod, T.I.: Image Segmentation Based on the Evaluation of the Tendency of Image Elements to form Clusters with the Help of Point Field Characteristics. In: Cybernetics and Systems Analysis, 51(5), 704-713. (2015)