The use of machine learning methods in the diagnosis of diseases of crops

Sergey Korchagin$^{1,*}$, Denis Serdechny$^2$, Roman Kim$^3$, Denis Terin$^4$, and Mihail Bey$^5$

$^1$Financial University under the Government of the Russian Federation, Department of Data Analysis, Decision Making and Financial Technology, 105187, st. Shcherbakovskaya, 38, Moscow, Russia

$^2$State University of Management, Department of Innovation Management, 109542, Ryazan Avenue, 99, Moscow, Russia,

$^3$Yuri Gagarin State Technical University of Saratov, Department of Information Security of Automated Systems, 410054, Polytechnic St., 77, Saratov, Russia

$^4$Saratov State University of Saratov, Department of Nano and Biomedical Technologies, 410012, st. Astrakhanskaya, 83, Saratov, Russia

$^5$Institute of Precision Mechanics and Control of the Russian Academy of Sciences, 410028, st. Working, 24, Saratov, Russia

Abstract. The approach to solving the problems of diagnosis and prognosis of diseases of agricultural crops using machine learning methods is described. To solve the problem of forecasting diseases of agricultural crops, it is proposed to use a genetic algorithm in the work. The analysis of the effectiveness of the proposed method is carried out depending on the convergence rate of such parameters as the mutation coefficient and population size. To solve the problem of diagnostics of agricultural crops, it is proposed to use a recurrent type of neural network. A software modelling complex has been developed that allows solving the problems of plant diseases diagnostics and making forecasts. The results obtained can reduce the costs of agricultural enterprises by reducing the cost of diagnosing agricultural diseases.

1 Introduction

Modern information technologies are actively used in many applied tasks in agriculture.

An analysis of the scientific literature [1-3] demonstrates the high potential of using machine learning methods in agricultural tasks related to the creation of new automated crop processing methods, the creation of robotic systems, computer vision systems, as well as diagnostic systems based on artificial intelligence.

Currently, specialists and scientists in the field of information technology specializing in the agricultural sector are using special programs and technical solutions that use machine learning methods and neural networks, for example, EuroMobile, AGRO-NET NG, AGRO-MAP PF, Alphabet’s X, etc.

An urgent task is the development of software systems and models for solving the problem of diagnosing diseases of agricultural crops and predicting their distribution.

* Corresponding author: SAKorchagin@fa.ru
Automation and increasing the speed of the process of searching for optimal solutions for these tasks is a promising interdisciplinary direction at the junction of agriculture and information technology, which allows to reduce the costs of agricultural enterprises.

2 Machine Learning Methods

In this work, the search for diseases of agricultural crops is carried out using the genetic algorithm [4]. Genetic algorithm is an optimization method based on imitation of natural selection processes [5]. The genetic algorithm uses a finite set of solutions, creating new data using the selection, mutation, and crossover operators. Output, external and internal parameters are a quantitative assessment of the parameters of the object. In fig. Figure 1 shows a general view of the genetic algorithm, which includes several main phases: generation, crossing, mutation, selection.

![Genetic Algorithm Diagram](image)

**Fig. 1.** Type of genetic algorithm

The development of the algorithm begins with the use of the smallest indivisible unit of a species subject to evolutionary transformation. Such is an individual, which is a leaf of potatoes, which is represented by a vector of internal parameters or a genotype. In the case of interaction of the individual with the external environment, a phenotype appears (manifestation of external signs), which is characterized by adaptation of the individual (infected leaf of potato). Thus, there is a criterion for evaluating an individual, in relation to others. A set of individuals with a similar genotype make up a population, the area of which is the set on which individuals are located.

At the initial stage, a starting population and an assessment of the fitness of an individual are generated; for this purpose, a target function is calculated. The next step is the crossing in the population and an increase in its number. Population growth depends on the choice of cross-breeding mechanism, so the population can double, triple or grow in a different proportion. The genotype of the descendants of individual individuals changes, i.e. a new generation goes through a mutation phase, which can take place either randomly or according to a given algorithm. The next step is the analysis of fitness (the calculation of the target function of the nanocomposite is carried out and it is determined how much the obtained characteristics are as close as possible to the desired).

At the next stage, an analysis of fitness is carried out for each individual, only the fittest individuals remain in the population, the rest are destroyed. If at this step all the affected
areas on the potato leaf are found, the process stops, otherwise the transition to the crossing step is performed.

The software package that implements the genetic algorithm is implemented in the C++ programming language and includes the classes described below. The Evolution class is used to model the evolutionary process. To use the properties of the materials that make up the composite, the potato class has been created, which allows one to parameters of disease (the database is loaded from the source, based on the data from [6]).

Before starting the evolution process, a Genotype class is created and configured for the current task. Setting up and managing the genotype is done using the GenotypeManager class. GenotypeLibrary contains a set of genotypes for various composites (amount of materials, geometric parameters, etc.). After that a population is generated (Person class) and the evolution process is started. For the work of the objective function, the Model class has been developed, which encapsulates the method for calculating the properties of the composite and the Algorithm class for implementing numerical methods for finding the roots of polynomial equations with complex variables.

To solve the problem of diagnosing late blight on potato leaves, neural networks were used. At the first stage of modelling, input data are prepared. As part of the task, to reduce the size of the training sample, as well as increase the convergence of the neural network. For the forecast problem, a recurrent type of neural network was used. Neural network training was based on data obtained experimentally in [7]. The size of the training sample was taken in the amount of 20,000 photos. As a machine vision technology, the YOLOv3 solution was used. The analysis of neural network accuracy comes down to determining the error function. In this study, we used the determination of the mean square error. The result of neural network prognosis is the determination of late blight disease on potato leaves.

![Potato leaves](image)

**Fig. 2.** Potato leaves

### 3 Conclusion

The paper presents the possibility of using the genetic algorithm in the problems of diagnosis of diseases of agricultural crops. As an example, the diagnosis of late blight on potato leaves is considered. To solve the problem of forecasting and diagnostics, it is proposed to use a genetic algorithm and a recurrent type of neural network, which allows
determining damaged areas on potato leaves. The results can reduce the costs of the agricultural sector in the cultivation of agricultural crops.

References

1. Anna Chlingaryan, Salah Sukkarieh, and Brett Whelan, Computers and electronics in agriculture, 151, 61-69 (2018)
2. Borja Espejo-Garcia, et al., Computers and electronics in agriculture, 150, 343-352 (2018)
3. Konstantinos G. Liakos, et al., Sensors 18 (8) 2674 (2018)
4. R. P. Kim, S. P. Romanchuk, D. V. Terin, S. A. Korchagin, Math. Mech. Inform., 19 (2), 217-225 (2019)
5. Chunyan Lai, et al., IEEE Transactions on Industry Applications, 53 (5), 4493-4503 (2017)
6. Dor Oppenheim, and Shani Guy, Advances in Animal Biosciences, 8 (2), 244-249 (2017)
7. Dor Oppenheim, et al., Phytopathology, 109 (6), 1083-1087 (2019)