THE USE OF BIG DATA PROCESSING IN A DIGITALIZED AGRO-INDUSTRY SYSTEM

R.A.Gulyaev[1]
[1]Dsc, Deputy for Investment, Innovation and Science in LLC "Agroklaster".
U.M. Ibragimov[2]
[2]PhD, Associate Professor in Bukhara engineering-technological Institute.
H.B. Ismoyilov[3]
[3]Head of the Department for Innovation and Development of Information and Communication Technologies.
https://doi.org/10.5281/zenodo.7115902

Abstract

Despite the fact that the 21st century is called the age of information technology and information society, agrarian technology and the cultivation of agricultural products necessary for the needs of people remain an urgent issue in all countries of the world. Only, as mentioned above, in the 21st century, the quality, accuracy and other characteristics of agricultural production can be affected by using modern information technologies and information society using its inventions and sciences. This article provides information on the application of big data analysis in a digitalized agro-industry system.

Keywords: BigData, API, system, contours, agro-industry, NDVI, interface, machine learning, classification, geofields, statistical analysis.

Bukhara region has long-standing experience in timely growing agricultural products of high quality. But at the same time, the difficult nature of the weather in this region, the complexity and scarcity of water supply, the salinity of the land and other effects make it difficult to grow these products. Due to this, the use of modern information technologies and related intellectual systems in this region is one of the new directions in the Republic along with the demand of the time. In the republic, Bukhara region remains a leading and experimental region in the application of new modern information technologies, software, automated systems and intellectual sciences in agriculture. Due to this, the quality and timely production of agricultural products in the existing state of the difficulties mentioned above, using a large amount of data in automated information systems used in agriculture and modern intelligent data processing systems, one of the main innovations is timely watering of agricultural products (or drip irrigation), medication (providing necessary vitamins or protection from pests), carrying out agochemical analysis, timely planting and other agotechnical works using intelligent systems to predict in advance, clustering images of growth according to the contours of the products based on the data classification using

machine learning method and pre-planning of agrotechnical work in which periods to increase the yield.

The above-mentioned issue is especially relevant today, considering the quality and timely delivery of products in the world.

In addition, in the Bukhara region, in particular, "Bukhara Agrocluster" LLC, in a large-scale cluster that unites more than 300,000 farmers, more than 40 regions and 6 districts, the use of the following automated systems as part of the hardware and software supplies for several years now provides all the opportunities to apply this issue in practice[1]. Taking into account the availability of basic systems, the practicality of using large volumes of data in this field with intellectual processing and the probability of its practical application increases several times:

- Control system of contours (on the distribution of crop cultivation) coordinated with maps (Google, UzKosmos and other providers) in the section of farms;
- A system of carrying out agochemical analyzes on contours coordinated with maps in the cross-section of farms;
- System of planning, verification and control of agro-technical works in terms of seasons (agricultural years) in the section of farms;
- A system that retrieves maps and related information from information systems providing space satellites;
- A system for storing and displaying on cloud servers the operation of equipment (tractors, earthmoving and other equipment) and the sensors installed on them in terms of seasons (agricultural years) and their contours in the section of farms;
- Collecting and systematically displaying weather information on farmers and their contours from physical weather stations and providing online service in the region;
- A system for taking photos of agricultural products in the contour section with mobile applications during the time period of the production process and displaying them in the system, as well as conducting analysis;
- System of keeping register of land, slope of slope, speed of slope and other data in section of contours;
- System of reports on the statistical analysis of the above-mentioned systems and their data, etc[2,3].

The use of big data analysis methods, in our case, determining the ripeness of the crop for timely harvesting and maintaining quality, classification methods

and algorithms help as an additional functionality for organizing big data processing and intelligent processing[4].

For this, in the Bukhara region, in particular, in the Bukhara Agrocluster LLC, in a large-scale cluster that unites more than 300,000 farmers, more than 40 regions and 6 districts, the following automated systems have been used for several years as part of the hardware and software. Taking into account the existence of basic systems, the practicality of using large volumes of data in this field with intellectual processing and the probability of its practical application increases several times:

In order to carry out this practical work, we may receive data from the following internal systems. Some data is generated through the API data exchange system, some data is obtained directly from the database (in particular, through the NoSQL MondoDB system), and data can be extracted in various formats and from other service systems. Obtaining this information is provided below. Screenshots and API documentation are omitted as they are implemented through authorization and authentication in the system[1,2].

- In the section of farms, the control system of contours (on the distribution of production) coordinated with maps (Google, UzKosmos and other providers);

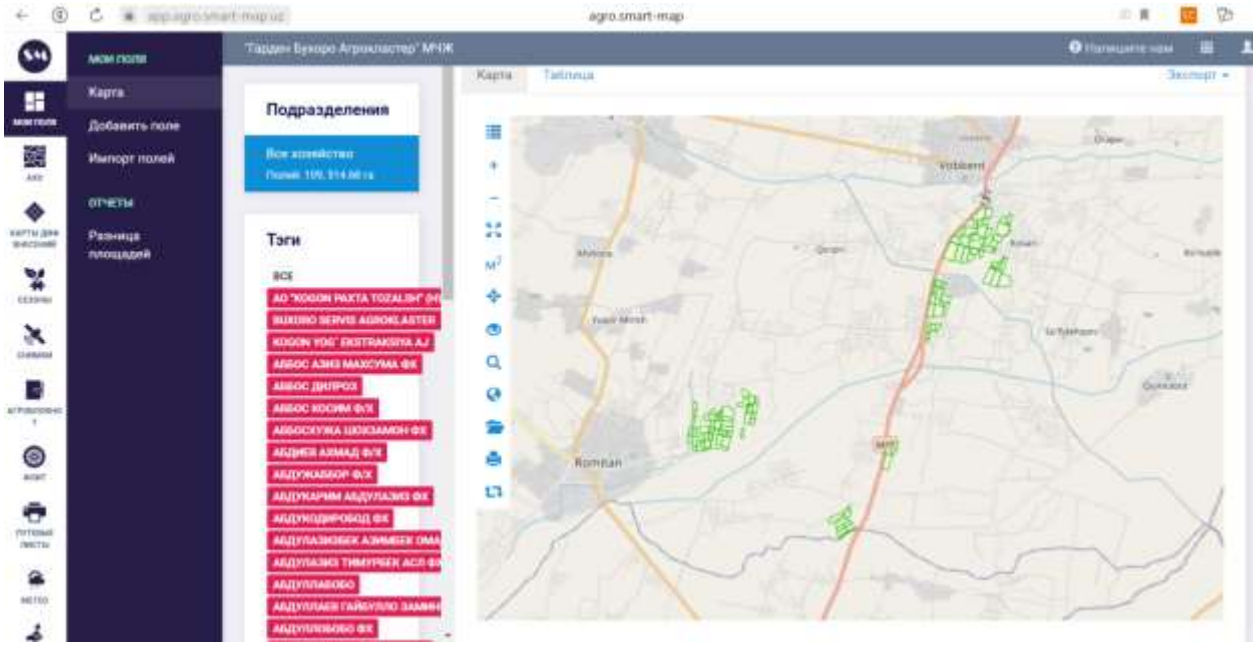

Fig 1. The section of farms, the control system of contours.

- A system for carrying out agrarian chemical analyzes in contours coordinated with maps in the section of farms;

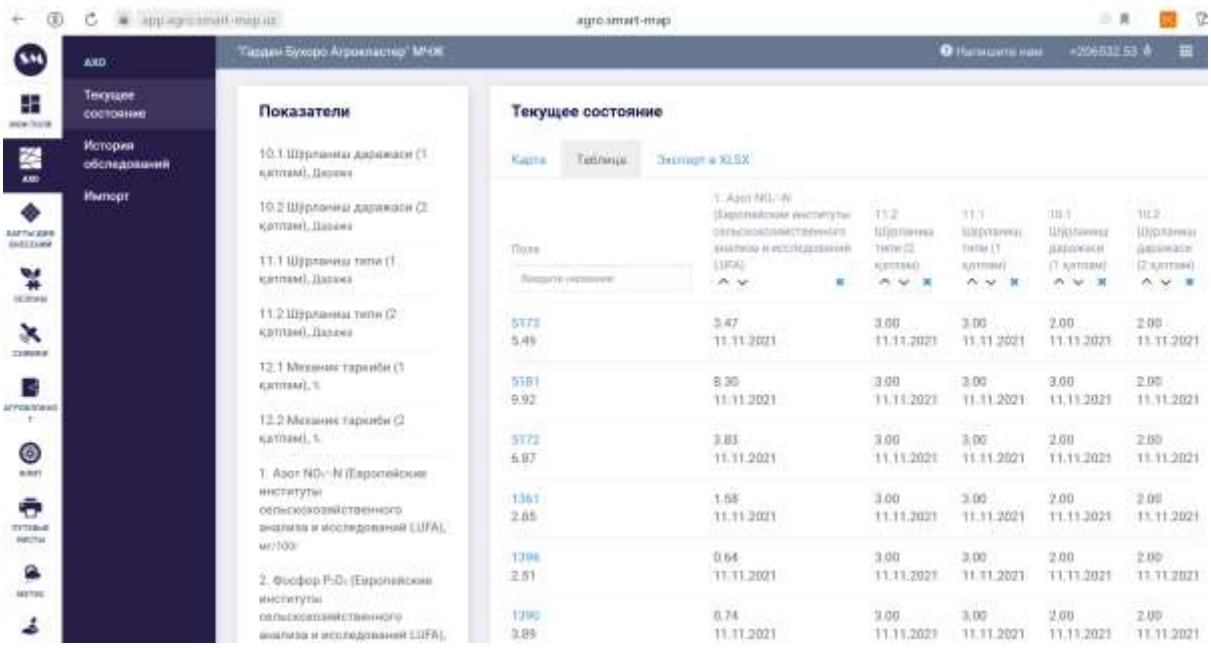

Fig 2. The system for carrying out agrarian chemical analyzes.

- A system of planning, verification and control of agrotechnical works by seasons (agricultural years) in the section of farms;

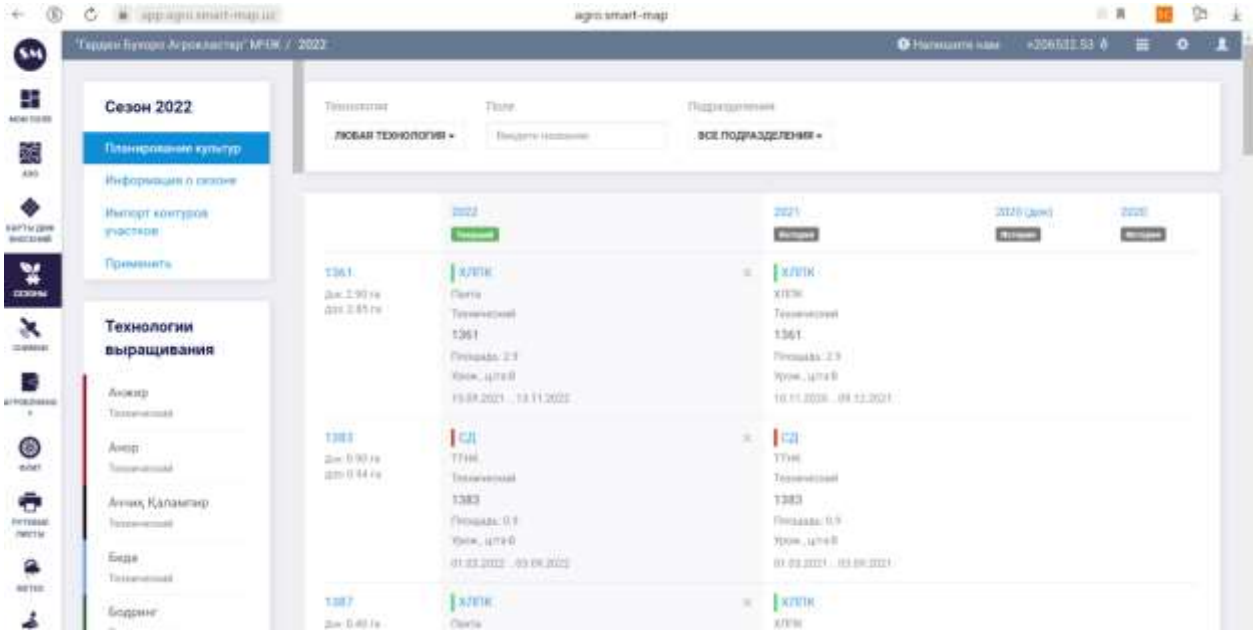

Fig 3. The system of planning, verification and control of agrotechnical works.

- A system that retrieves maps and related information from information systems that provide space satellite services;

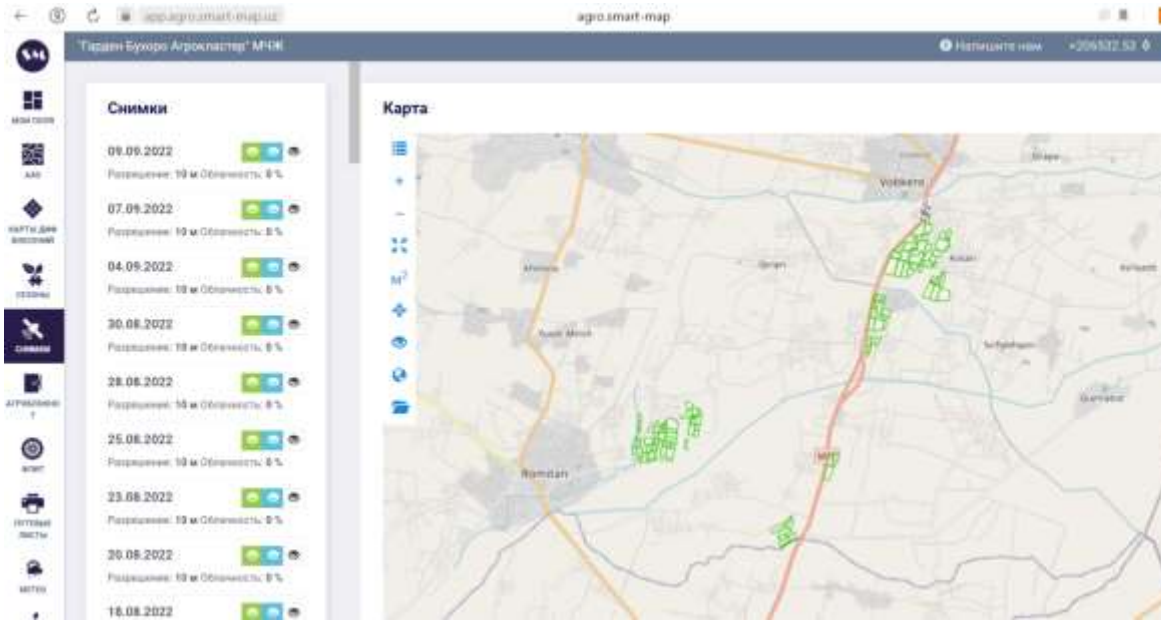

Fig 4. The system that retrieves maps and related information.

- A system for storing and displaying on cloud servers the performance of equipment (tractors, earthmoving and other equipment) and the time indicators of the sensors installed in them by seasons (agricultural years) and by their contours;

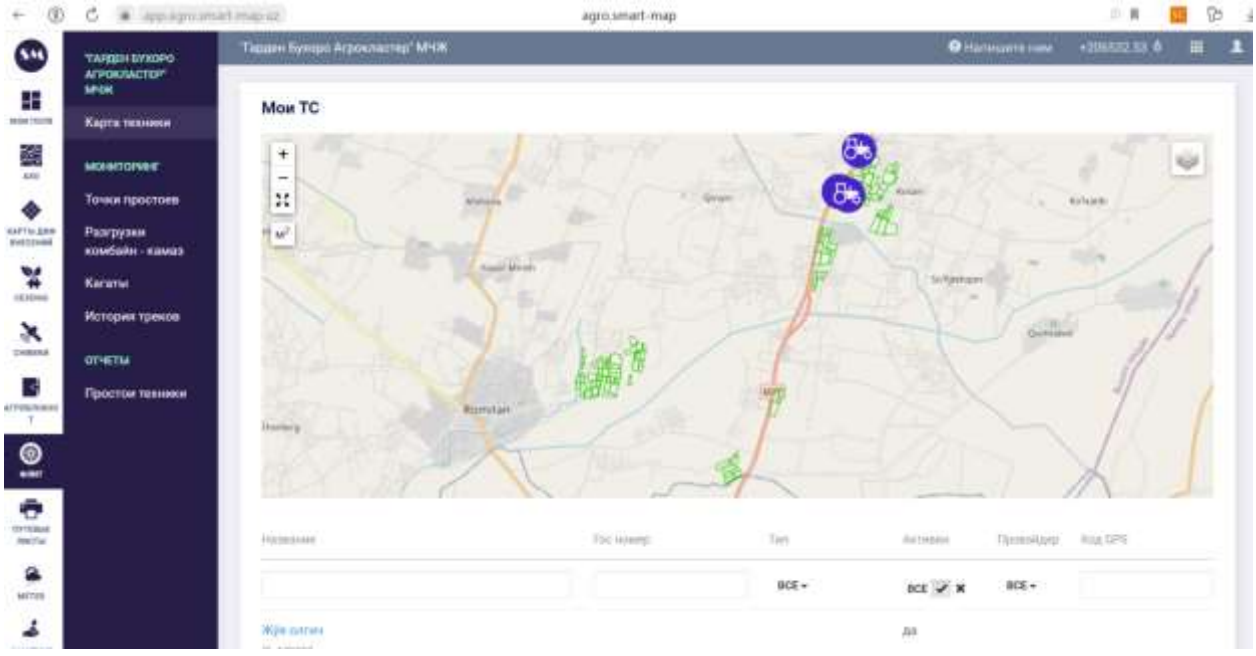

Fig 5. The system for storing and displaying on cloud servers the performance of equipment.

- Data collection and systematic display of meteorological data on farmers and their contours from physical meteorological stations in the region and providing online service;

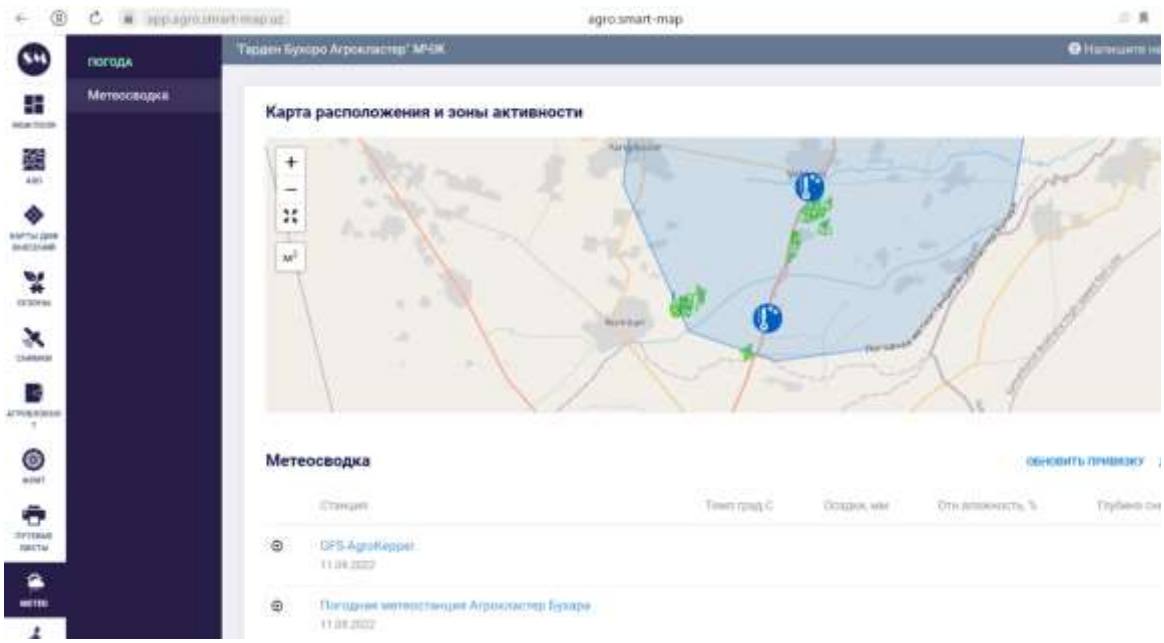

Fig 6. Data collection and systematic display of meteorological data on farmers and their contours from physical meteorological stations.

- A system for taking photos of agricultural products in the contour section with mobile applications during the period of time, reflecting them in the system, and conducting analyzes;

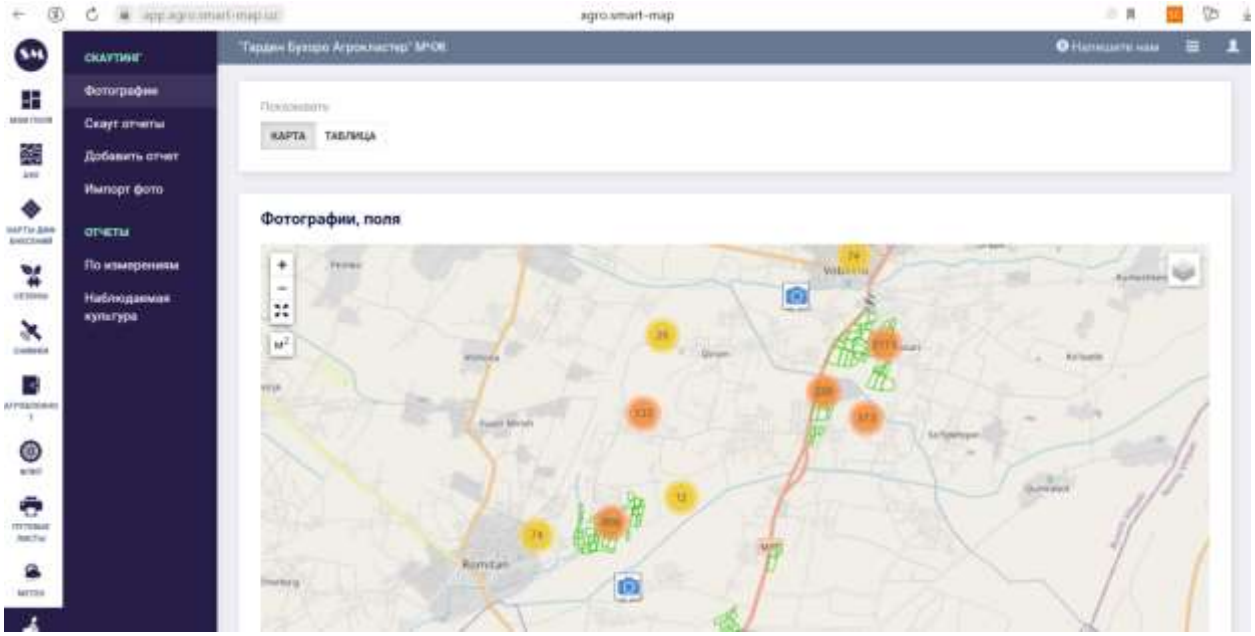

Fig 7. The system for taking photos of agricultural products in the contour section with mobile applications.

- A system for storing land registers, elevations, gradients and other data in the section of contours;

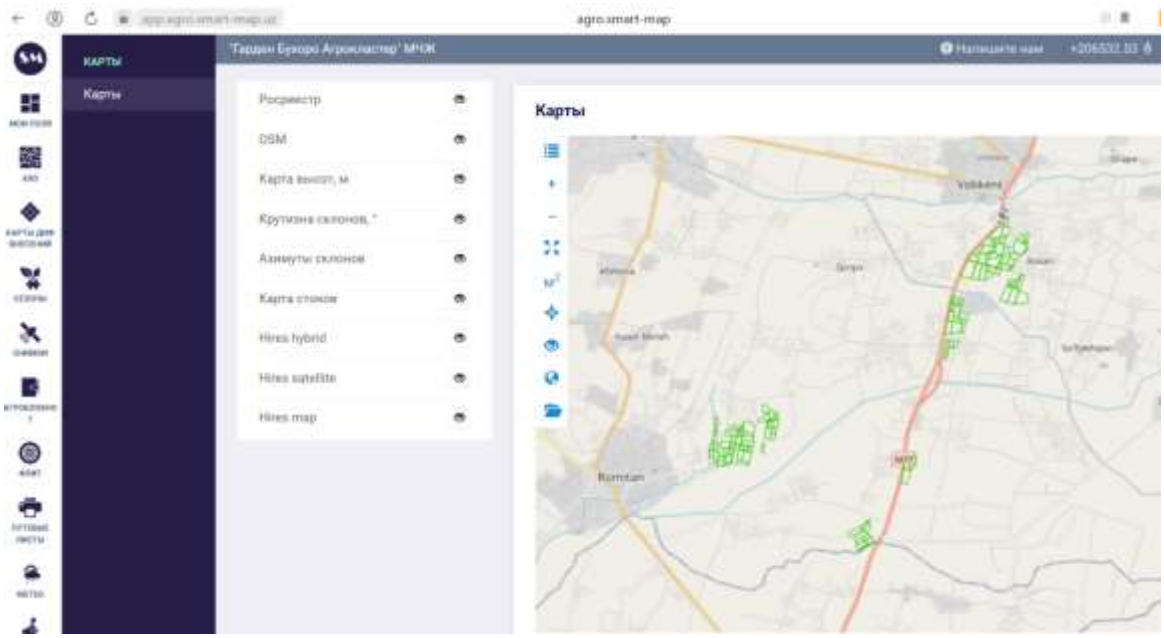

Fig 8. The system for storing land registers, elevations, gradients and other data in the section of contours.

- A system of reports on the statistical analysis of the above-mentioned systems and their data, etc.

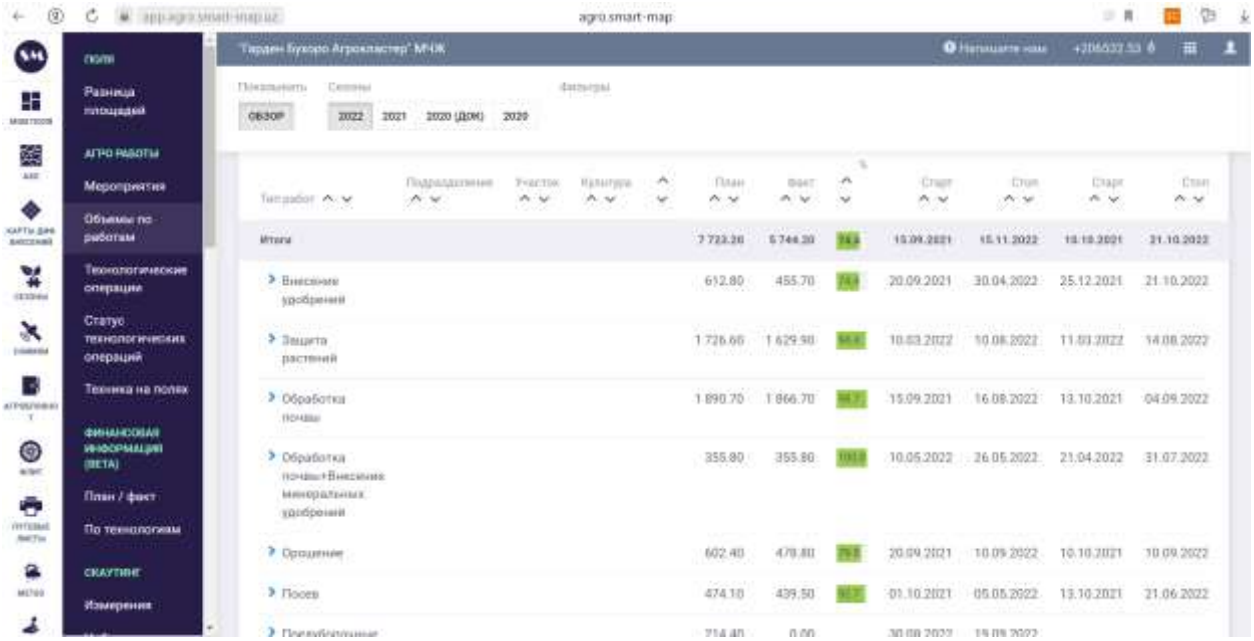

Fig 9. The system of reports on the statistical analysis of the above-mentioned systems and their data.

All this data is very effective and helps to use it as big data for processing, but in recent days, the analysis and processing of geofields based on GIS data systems and information systems filling in the database associated with geofields around the world have begun to apply new methods for analyzing big data[4].

There are two main tasks in supervised machine learning: classification and regression.

The purpose of classification is to predict a class label, which is a choice from a predefined list of possible choices. Classification is sometimes divided into binary classification, which is a special case of division into two classes, and multiclass classification, when more than two classes are involved in the classification. Binary classification can be thought of as an attempt to answer a given question in a "yes/no" format[5].

The purpose of regression is to predict a continuous or floating-point number in programming terms, or a real number in mathematical terms.

In supervised machine learning, we need to build a model on the training data and then make accurate predictions for new data we haven't seen yet that has the same characteristics as the training set we used. If the model can make accurate predictions on previously unseen data, we say that the model has the ability to generalize the result to the test data. We need to build a model that will have the maximum generalizing ability.

The more complex the model, the better it will perform on the training data. However, if our model gets too complex, we start paying too much attention to every single data point in our training set, and that model won't generalize well to new data.

There is an optimal point that allows getting the best generalization ability. Actually, this is the model that we need to find[6,7].

We propose to use these big data analysis methods in information systems as an additional functioning module in the development or modification of digitalization systems for the agro-industry.

For example, in the Bukhara region (Uzbekistan), the management of Bukhara Agrocluster LLC developed and implemented the smart-map system. We plan to use this methodology to intelligently determine crop maturity. To apply these algorithms and methods in the smart-map system, which has been used in the region for several years, there is a module called Monitoring the state of crops, determining the plant mass index (NDVI). Images are taken from unmanned aerial vehicles (UAV) and satellites. This module in the system collects big data for future analysis[2].

Fig 10. Interface of the smart-map system.

In addition to using machine learning for clustering, a hierarchy of conditions will also be included in practical development and checked against data obtained from internal systems through APIs, exported data, databases and third-party systems to get more accurate decisions and results.

References:

1.R.A.Gulyaev. A.A.Sultonov, R.F.Yunusov, D.R.Rafikov. on the development of a system for digital remote monitoring of agricultural land. The American Journal of Agriculture and Biomedical Engineering. ISSN – 2689-1018. VOLUME 04 ISSUE 03 Pages: 42-49.

2. R.A.Gulyaev, A.A.Sultonov, H.B.Ismoyilov. Comprehensive hardware and software approach to digitalization of agricultural industry. GALAXY INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (GIIRJ) ISSN (E): 2347-6915. Vol. 9, Issue 12, Dec. (2021).

3. Sh.T. Shukurov, U.M. Ibragimov. Algorithms of the Systems of Big Data Storage and Analysis. European Journal of Innovation in Nonformal Education(EJINE) Volume 2| Issue 5| ISSN:2795-8612. 149-153 pp.

4. Sh.T. Shukurov, U.M. Ibragimov. Architecture for building the systems of storage and analysis of BIG DATA. "Economics and society" №5(96) 2022(www.iupr.ru).

5. Ж.Ю. Саидов, У.М.Ибрагимов, К.Х. Гафуров, Математический анализ и алгоритм автоматизации процесса сортировки фруктов. Современные научные исследования и инновации. eISSN: 2223-4888.

6. Perri Lea. Internet of Things for Architects. ДМК Издательства. Москва. 2019. -455 с.

7. Vinit Yadav. Processing Big Data with Azure HDInsight. Building Real-World Big Data Systems on Azure HDInsight Using the Hadoop Ecosystem. APress. Ahmedabad, Gujarat, India. 2017. -221 p.