# Intelligent System for Solving Problems of Veterinary Medicine on the Example of Dairy Farms

Shopagulov Olzhas, Tretyakov Igor and Ismailova Aisulu

Dept. of Information Systems, Kazakh Agro Technical University named after S. Seifullin, Nur-Sultan, Kazakhstan

## Abstract

*This article describes an automated expert system developed to diagnose cow diseases and assist veterinarians in treatment. We set before a diagnostic method based on the analysis of observed symptoms and experience of veterinarians. The system represents a web interface for maintaining a database of diseases, their symptoms and treatment methods, as well as a smartphone application for the diagnostics in offline mode. The developed intelligent system will allow agricultural producers to make specific decisions based on automated data analysis. Also presented in the article the information on the developed expert system, and the results of tests and testing during its use. The economic efficiency and importance of the work is determined by the possibility of automated recording of data on the livestock of animals, zoo technical and veterinary operations.*

## Keywords

*Intelligent system, diagnosis of diseases, application evaluation, milk yield, herd management*

## 1. Introduction

Modern animal husbandry is a transition to new production approaches. The basis of this transformation is agricultural machinery and equipment, supplemented by telemetry systems. Telemetry and monitoring systems for agricultural production are considered one of the most innovative technologies for precision animal husbandry and are automatic systems for collecting and transmitting information, as well as data analysis and remote decision making.

Nowadays, the use of digital technologies in intensive dairy farming is playing a key role in the proper management of the herd to improve animal welfare and increase the milk production of cows. Italian scientists E. Tullo, I. Fontana, D. Gottardo, K. H. Sloth, M. Guarino [1] conducted research on the application of the GEA CowView system (Gea Farm Technologies, Benen, Germany) on dairy cows.

Similar studies were carried out by Turkish scientists on the use of the SCADA application in dairy farming, which allows you to evaluate all records of animal production in a digital environment (behavior, production, health, feeding, and other events). With the introduction of automation systems into industrial production, it creates new work areas in the livestock sector for various applications of emerging technologies [2].

A huge role in the productivity of cattle is played by various diseases. These factors are a significant obstacle to the healthy and sustainable development of livestock breeding. Cattle, in

particular cows, have a number of specific features in the diagnostics and establishing diagnosis. Taking into account the world experience in developing intelligent systems, the conclusion on their creation has been made. The uniqueness of this system is determined by the development methods used and the knowledge base of veterinary experts. The study objective was to develop an automated tool for diagnostics of cattle diseases and, and study its impact on the diagnostics of cattle diseases in Kazakhstan [3].

Contrary to humans, animals are unable to describe their feelings, and diagnosis of diseases is based on externally observed symptoms and laboratory tests. For example, if a lacrimation is detected in a cow, it may indicate conjunctivitis. But practice shows that in most cases there are other factors that affect the development of disease. Animals do not exhibit obvious clinical symptoms when they suffer from a disease, the former are usually misdiagnosed as a result of the disease. In spite of the improved diagnostics of cow diseases in recent years, most veterinarians do not have much experience, which leads to serious losses due to delayed disease control, as well as serious problems in accounting diseases on paper. Health of cows is a key factor in dairy herd productivity. Mastitis, ketoses, fattening problems and other diseases significantly reduce dairy production and treatment will be expensive if diagnosed and detected late. Thus, the financial component and disease prevention are interlinked phenomena on any farm. Improving diagnostic accuracy and reducing losses caused by disease are the most serious problems at present (3).

Input data needed to diagnose cattle diseases should be classified, i.e. we need to know what information a farmer should have in order to make a decision [3]. Several groups of input data can be identified:

1.  External characteristics:
- Sex, breed, age (date of birth), live weight of the animal;
- Genealogical tree of the animal (breeding record);
- Animal data: individual number, body type, live weight, age, color, sex, photo of the animal,
- Place of birth, date of birth and location of the animal;
- Method of birth of the animal (natural / artificial breeding);
- Date of slaughter (and disposal actions following the slaughter).
2.  Animal productivity:
- Volume of dairy products produced and milking schedule;
- Live weight at present and graphs of live weight changes;
- Date and method of last animal cover (natural or artificial breeding), date of pregnancy test and its result, date of start in the dry;
- Date of last calving and number of calves (live, stillborn);
3.  Animal's medical information:
- Full medical history of the animal (dates of diagnosis, treatment);
- Immunization and vaccination of the animal with the dates and type of medication injected;
- Graph of body temperature changes, mobility, recent PH in the stomach.
4.  Observed symptoms and organ and body system lesions: skin cover, musculoskeletal system, nervous system, cardiovascular system and others.

By processing the received input and output data, we should obtain a set of expert conclusions (output data) accepted by the system:

1. A diagnosis of the animal's disease obtained by analyzing the input data and comparing it with the symptom data library;
2. Recommendations for further actions on treatment, prevention, use of medications, etc.;

## 2.  GENERAL SYSTEM DESIGN CONCEPT

The expert system architecture was developed according to the method of structured systems development. It consists of a knowledge base, a knowledge subsystem, a decision-making subsystem, an administration subsystem and a user interface (Figure 1) [3].

The system uses N-tier web architecture (Figure 1), the structure has been developed according to the method of development of structured systems, consisting of:

1. Knowledge databases and knowledge generation subsystems (interface for working with veterinary experts);
2. System administration block and the administrator work interface correspondingly;
3. Database and data management block are the core of the system;
4. Work with input and output data as well as decision making and work with the user interface is located in the disease diagnostic subsystem.
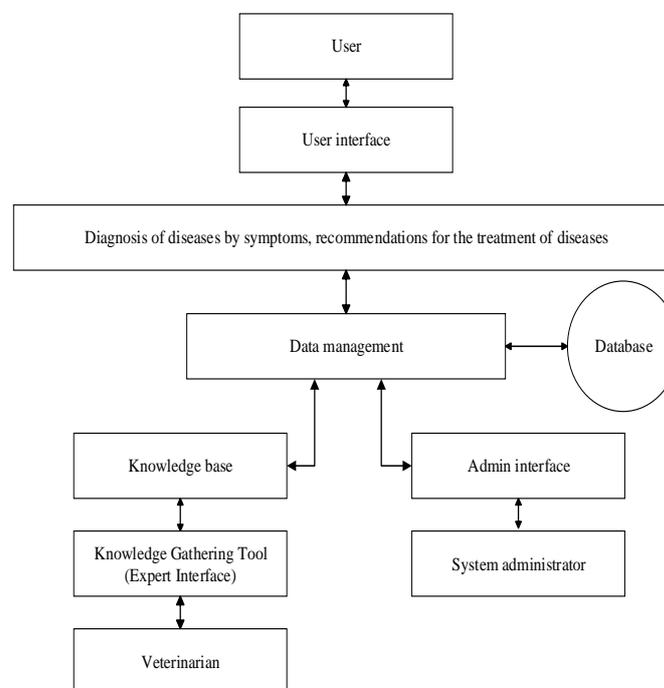
Figure 1: Structure of a cow disease diagnosis system

Decision making diagram of the system is used for interpretation of user interface functions. Diagram of variants of use of expert system of diagnostics of diseases of the cows consists of six compound blocks shown in Fig. 2.
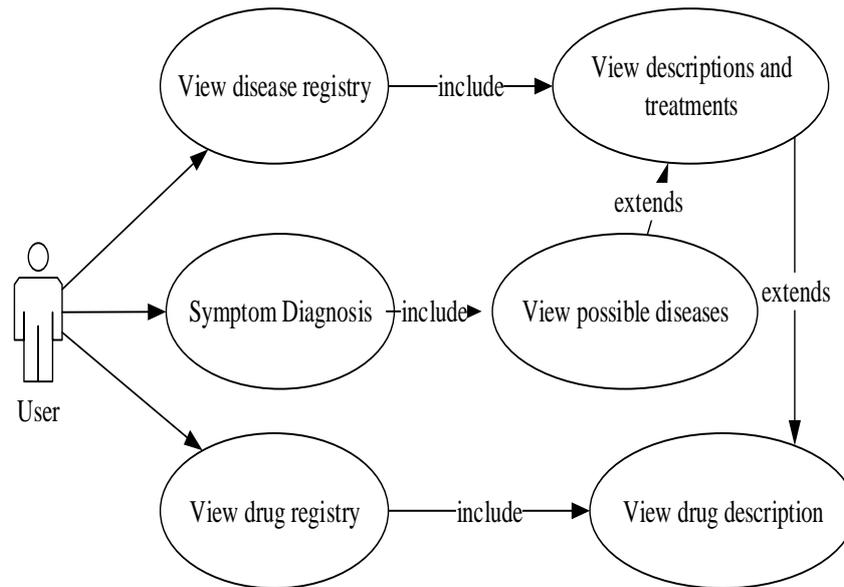
Figure 2: The main scenarios for using the expert system

Figure two shows that a user entity in the system can perform a number of actions, such as viewing the disease registry, performing diagnostics of diseases by symptoms and viewing the medication registry. On this basis, having defined a diagnosis by initial symptoms, the system will offer a description, methods of treatment and prevention of this disease. After selecting a treatment method, it is possible to view the descriptions of the medications used [3].

## 3. ACQUISITION AND PRESENTATION OF KNOWLEDGE

As expert system, it should contain knowledge obtained from experts in the subject area. Acquisition and presentation of knowledge is the most important stage in building expert systems. The main task is to create a knowledge database to meet the requirements of the expert system for the solution of set tasks. Many methods were developed for obtaining knowledge from experts in the subject area.

In this study, we have analyzed and summarized the aggregate of knowledge by conducting literature reviews and interviewing experts using a questionnaire for disease analysis. It consisted of questions about symptoms, diagnoses and treatments. Based on their experience and requirements to develop an expert system, experts modified the questionnaire and provided information in tabular form [8].

In this study the diagnosis of 16 most dangerous infectious diseases of cattle is considered: Anthrax (D1), FMD (Foot-and-Mouth Disease) (D2), Tuberculosis (D3), Brucellosis (D4), Rabies (D5), Pasteurellosis (D6), Trichophytosis (D7), Leukosis (D8), Infectious Rhinotracheitis (D9), Viral Diarrhea (D10), Lumpy Skin Disease of Cattle (D11), Emphysematous Carbuncle in Cattle (D12), Salmonellosis (D13), Colibacillosis (D14), Rotavirus (D15), Coronavirus (D16).

| Symptom codes | Anthrax | FMD | Tuberculosis | Brucellosis | Rabies | Pasteurellosis | Trichophytosis | Leukosis | Infectious Rhinotracheitis | Viral Diarrhea | Lumpy Skin Disease of Cattle | Emphysematous Carbuncle in Cattle | Salmonellosis | Colibacillosis | Rotavirus | Coronavirus |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Amount: | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | |
| F01 | 1% | 1% | - | - | 1% | 2% | - | - | - | 10% | 1% | 5% | 1% | 1% | 1% | 1% |
| F02 | 1% | 1% | - | - | 1% | 2% | - | - | - | - | - | 5% | - | - | - | - |
| F03 | - | 8% | - | - | 8% | 2% | - | - | - | 10% | - | - | - | - | 8% | 10% |
| F04 | - | 8% | - | - | 6% | - | - | - | - | 10% | - | - | - | - | - | 12% |
| F05 | 8% | - | - | - | 2% | - | - | - | - | - | - | - | - | 5% | - | - |
| F06 | - | 8% | - | - | 6% | - | - | - | - | 10% | 5% | - | - | - | - | - |
| F07 | - | 8% | - | - | - | - | - | - | - | 8% | - | - | - | - | - | 8% |
| F08 | - | 8% | - | - | - | - | - | - | - | - | 5% | - | - | - | - | - |
| F09 | 8% | 1% | - | - | - | - | - | 1% | - | - | - | - | - | - | - | - |
| F10 | 1% | 1% | - | - | 1% | 1% | - | 1% | - | 10% | - | - | - | - | - | - |
| F11 | 4% | - | - | - | 1% | - | - | - | - | 10% | - | - | - | 5% | - | - |
| F12 | 4% | - | - | - | - | - | - | 1% | - | 10% | - | - | 8% | 6% | 12% | 12% |
| F13 | 1% | - | - | 1% | 1% | - | 1% | - | - | - | - | - | - | - | - | - |
| F14 | 8% | - | - | - | 2% | - | - | - | - | 10% | - | - | 8% | 6% | 12% | 12% |

(Row group label: Digestive system)

Figure 3: Table of symptoms and their weighting values

Symptoms of disease are grouped according to the affected organs and systems, and each symptom is assigned a code for ease of work:

Skin: Skin Lesions (S01), Rumpleness of Hair (S02), Dermatitis (S03), Lumpy Skin, Extuberances (S04), Papules (S05), Vesicles (S06), Pustules (S07), Sloughs (S08), Scaliness (S09), Scratch (S10), Dark Red Skin Stains (S11).

Musculoskeletal system: Lesion of Limbs (M01), Arthritis (M02), Bursitis (M03), Limp (M04), Bone Deformity (M05), Suppurative Discharge (M06), Edema of Limbs (M07), Edema in Groats, Lungs, Neck, Chest, Lower Jaw (M08), Edema of Joints (M09), Joint Deformity (M10).

Digestive system: Loss of Appetite (F01), Loss of Chewing Cud (F02), Salivation (F03), Stomatitis (F04), Oral Mucosa Hemorrhages (F05), Oral Lesions (F06), Aphta, Oral Ulcers (F07), Vesicles, Tubercules in Mouth (F08), Gastric Timpany (F09), Gastric Atony (F10), Abdomen Wall Disease (F11), Diarrhea (F12), Constipation (F13), Excrements admixed with blood, mucus, gas bubbles (F14).

Respiratory system: Respiratory System Injury (B01), Rapid Pulse (B02), Short Wind (B03), Nasal Cavity Excretion (B04),  Injury of Nasal Cavity and Tapetum Lucidum Cellulosum (hemorrhages, wounds, etc.), (B05), Coughing (B06), Lung Rale (B07), Lung Inflammation (B08), Edema in Larynx, Chin, Abdomen (B09).

Central nervous system (CNS): Injury of the CNS (N01), Excitation (N02), Inhibition (N03), Muscular Tremor (N04), Eclampsia (N05), Paresis (N06), Paralysis (N07), Ataxia (N08), Scratch, Pruritus (N09).
Cardiovascular system (CVS): CVS Lesion (H01), Asphyxiation of Mucous Membranes (H02), Mucous Membrane Anemia (H03), Mucous Membrane Hemorrhage (H04), Mucous Membrane Hyperaemia (H05), Tachycardia (H06), Arrhythmia (H07), Myocarditis (H08).

Urogenital system (G-U System): Lesions of G-U System (U01), Nebulous Urine (U02), Urine Erythrocytes (U03), Frequent and Painful Urination (U04), Abortions (U05), Orchids and Epididymitis (U06), Swelling of External Genitals (U07), Hyperemia of External Genitals (U08),

Viral Shedding of External Genitals (U09), Retention of Placenta (U10), Endometritis (U11), Lesion of Ovaries and Fallopian Tubes (U12), Vulvovaginitis (U13), Balanopostitis (U14).

Visual organs: Lesions of Visual Organs (E01), Mucous membrane bleeding (E02), Eyeball Retraction (E03), Photophobia (E04), Serous or Suppurative Discharge (E05), Conjunctival Hyperemia (E06), Swelling and Edema of Conjunctival (E07), Corneal Ulceration and Clouding (E08).

Lacteous Gland: Lacteous Gland Lesion (J01), Decreased or Stopped Secretion (J02), Afta, Dug Skin Erosions (J03), Dug Inflammation (J04), Udder Edema (J05), Udder Pain (J06), Mastitis (J07), Enlarged Supramammary Glands (J08), Watery Milk admixed with Blood or Curds (J09).

Lymphoid system: Lesion of Lymphoid System (L01), Inflammation of Submandibular and retropharyngeal Lymphnodes (L02), Inflammation of Prescapular Lymph Glands (L03), Inflammation of Precrural Nodes (L04), Inflammation of Parotid Lymphnodes (L05), Inflammation of Supramural Nodes (L06), Inflammation of Internal Lymphnodes (L07), Tumour Proliferation (L08), Enlarged Spleen (L09).

Based on their own experience, the experts attached a weighting factor (w) to each symptom. The concept of "symptom complex" has also been introduced to reflect a group of symptoms combined into a single value argument, which is also given a weighting value (w). Figure 3 shows some symptom values and weighting values [3].

In order to determine the most probable diseases, the system calculates the sums of the weighting values for the observed symptom complexes and individual symptoms, after which the calculated values and the corresponding diseases are ranked (sorted) in descending order.

## 4. METHOD OF CALCULATION

According to the method of weighting calculation, we can analyze a given case on a concrete example. Viral diarrhea disease, with a certain number of symptoms and symptom complexes. According to the knowledge database and the questionnaire received from veterinary experts, the following symptoms appear in case of viral diarrhea:

C01 (Fever) W(d, s)=1%;
M01 (Lesion of Limbs) W(d, s)=1%;
M04 (Limp) W(d, s)=1%;
F01 (Loss of Appetite) W(d, s)=10%;
F03 (Salivation) W(d, s)=10%;
F04 (Stomatitis) W(d, s)=10%;
F06 (Oral Lesions) W(d, s)=10%;
F07 (Aphta, Oral Ulcers) W(d, s)=8%;
F10 (Gastric Atony) W(d, s)=10%;
F11 (Abdomen Wall Disease) W(d, s)=10%;
F14 (Excrements admixed with blood, mucus, gas bubbles) W(d, s)=10%;
N01 (Injury of the CNS) W(d, s)=6%;
N03 (Inhibition) W(d, s)=1%;
N08 (Ataxia) W(d, s)=1%,
where C01 is the symptom code, W (d, s) is the weight value of symptom s for disease d.

Thus, it is possible to calculate the sum of weighting coefficients of symptoms by disease. A simple sum of weights of symptoms for the disease is calculated using a formula:

$$W(d, S^o) = \sum_{s \in S^o} W(d, s) \quad (1)$$

where d is the disease, So is the observed set of symptoms, W (d, s) is the weighting value of symptom s for disease d.

For the above example with the disease "viral diarrhea», the values of the arguments will be equal to:

With So = 14:
 W(d, s0)= W (C01) + W (M01) + W(M04) +  + W (F01) + W (F03) + W (F04)  +W (F06) + W (F07) + W (F010) + W+ +(F011) + W (F014) + W (N01) + W (N03) +  +W (N08)=100%

Due to the fact that several symptom complexes k with different weights w can be  defined for disease d, the symptom complex with the highest weighting is taken into account, each symptom of which includes the many symptoms observed:

$$k_{max} \in K(d), где\ S(k) \in S^o, и\ W(d, k), \quad (2)$$
$$i.e.\ W(d, k) \geq W(d, k_j)\ for\ \forall k_j \in K(d) \quad (3)$$

where: K (d) - symptom complexes of the disease d, S (k) is the set of symptoms of symptom complex k, W (d, k) - weight coefficient of symptom complex k, for disease d.

Thus, for viral diarrhea, the symptom complex with the highest weighting factor is $k_{max}$, which includes symptoms F01, F03, F04, F06, F07, i.e. $S(k_{max})$=5.

In total for the disease "viral diarrhoea" the symptom complex $k_{max}$ will give W (d, $k_{max}$) =65%, according to the knowledge base provided by veterinary experts.

Taking into account the above mentioned number of symptoms not included in the symptom complex will be calculated as Sx=S0-S($k_{max}$) difference, i.e. for viral diarrhea Sx=9. Thus, it is easy to calculate W(d, Sx)=52%.

The total sum of weights R for observed symptoms So and symptom complexes S($k_{max}$) for disease d is calculated by formula [12] :

$$W_r(d, S) = W(d, k_{max}) + W(d, S^y) \quad (4)$$

For viral diarrhoea, the total sum of R weights for the observed symptom group Sy (F11, F14) and symptom complexes S($k_{max}$) at the same time will be Wr (d,S)=85%.

After calculating the total sums, the obtained data are sorted in descending order.

Thus, we can conclude that the introduction of such a parameter as a symptom complex leads to a more accurate definition of this or that disease. The results obtained as a percentage may show that the observed symptoms belong to a certain diagnosis [3].

## 5. DATABASE, THEIR IMPLEMENTATION

For the central database as well as for its local version, relational databases (MS SQL Server 2019 and SQLite) are used on the user device. Figure 4 presents the database structure in terms of knowledge storage about diseases, symptoms and symptom complexes [3].
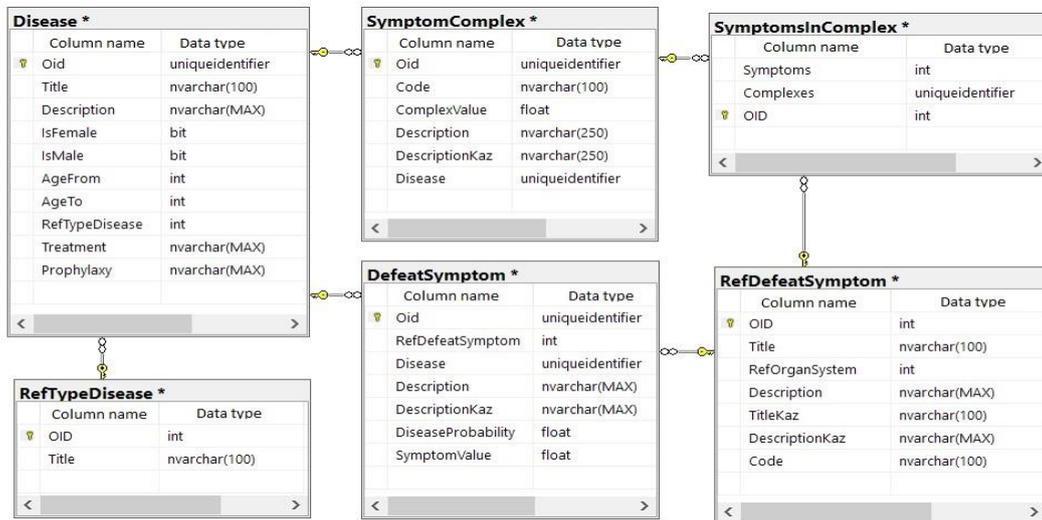


Figure 4: Database structure

Knowledge database contains information about 16 major infectious diseases and 103 symptoms of diseases. The database stores the subject area knowledge needed to solve problems, including age, cow breed, symptoms, photographs, and other relevant information. The database is developed on:



Figure 5: Veterinarian tablet web interface

- Operating system: Windows Server 2019 Standard;
- Web server: Internet Information Services;
- DBMS: SQL Server 2017 Standard;
- Platform: NET 4.5.2, language C #;

- Framework: DevExpress XAF 18.2 - a set of libraries to help the developed program with modern high-quality functionality [3].

## 6. EXPERT SYSTEM EVALUATION – TEST AND RESULTS

The evaluation process was carried out due to the user-friendliness of the user interface and system utilization efficiency testing. The reliability of the system diagnosis was evaluated with the participation of two groups of senior veterinary students of KATU named after S. Seifullin.

| № | Anthrax | Foot and mouth disease | Tuberculosis | Brucellosis | Rabies | Pasteurellosis | Trichophytosis | Leukemia | Infectious Rhinotracheitis | Viral diarrhea | Nodular dermatitis of cattle | Emphysematous carbuncle of cattle | Salmonellosis | Colibacillosis | Rotaviruses | Coronaviruses | Number of correct answers | % correct answers |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | - | + | - | + | + | - | + | - | + | - | + | + | - | - | - | - | 7 | 43,75 |
| 2 | - | + | + | + | + | - | - | - | - | - | - | - | - | - | - | - | 4 | 25 |
| 3 | + | + | - | - | + | - | - | - | - | - | + | - | - | + | - | - | 5 | 31,25 |
| 4 | - | - | - | + | + | - | + | - | - | - | + | + | - | + | - | - | 6 | 37,5 |
| 5 | - | + | - | + | + | - | + | - | + | - | - | + | - | + | - | - | 7 | 43,75 |
| 6 | - | + | - | + | + | + | + | - | - | - | + | - | - | - | - | - | 5 | 31,25 |
| 7 | - | - | - | + | + | + | + | - | - | + | + | - | - | - | - | - | 6 | 37,5 |
| 8 | - | + | - | + | + | - | + | - | - | - | + | + | - | + | - | - | 7 | 43,75 |
| 9 | - | + | - | + | + | - | + | - | + | + | + | + | + | - | - | - | 9 | 56,25 |
| 10 | - | + | - | + | + | + | + | - | - | + | + | - | - | - | - | - | 7 | 43,75 |
| 11 | - | + | - | + | + | + | - | - | - | - | - | - | - | - | - | - | 5 | 31,25 |
| 12 | - | + | - | + | + | - | + | - | - | - | - | - | - | - | - | - | 4 | 25 |
| 13 | + | + | + | + | + | - | - | + | - | - | + | + | - | + | - | - | 9 | 56,25 |
| 14 | + | + | + | + | + | - | - | + | - | - | - | - | - | - | - | - | 6 | 37,5 |

Figure 6:  Questionnaire of Veterinary Students

Table 1: Changes Characteristic in the Disease Diagnosis Using Tablet versus Diagnosis without Tablet

| Job Number | Name of the disease | Characterization of changes in the diagnosis of the disease when using a tablet, compared with diagnosis without a tablet |
|---|---|---|
| 001 | Anthrax | A significant increase in the probability of diagnosing anthrax – from 20% to 69% (3.43 times) |
| 002 | Foot and mouth disease | The increase in the probability of diagnosis of foot and mouth disease - from 70% to 93% (1.33 times). |
| 003 | Tuberculosis | A significant increase in the probability of diagnosis of tuberculosis - from 12% to 89% (7.41 times) |
| 004 | Brucellosis | A slight decrease in the probability of diagnosis of brucellosis - from 82% to 81% |
| 005 | Rabies | A slight decrease in the probability of a correct diagnosis of rabies is from 100% to 92%. Probably by providing subjects with additional options |
| 006 | Pasteurellosis | An increase in the probability of correct diagnosis of pasteurellosis - from 44% to 54% (1.23 times) |
| 007 | Trichophytosis | A slight increase in the probability of a correct diagnosis from 68% to 74% |
| 008 | Leukemia | A significant increase in the probability of a correct diagnosis of leukemia - from 10% to 86% (8.64 times) |
| 009 | Infectious Rhinotracheitis | Decrease in probability of correct diagnosis of infectious rhinotracheitis - from 54% to 44% |
| 010 | Viral diarrhea | A slight increase in the probability of a correct diagnosis of viral diarrhea - from 20% to 22% |
| 011 | Nodular dermatitis of cattle | Increasing the likelihood of a correct diagnosis of nodular cattle dermatitis - from 80% to 90% (1.13 times). |
| 012 | Emphysematous cattle carbuncle | An increase in the probability of a correct diagnosis of cattle emphysematous carbuncle - from 58% to 67% (1.15 times). |
| 013 | Salmonellosis | Reducing the likelihood of a correct diagnosis of salmonellosis - from 40% |

| | | to 35%. |
|---|---|---|
| 014 | Colibacillosis | An increase in the probability of correct diagnosis of colibacillosis - from 38% to 58% (1.52 times) |
| 015 | Rotaviruses | A significant increase in the probability of correct diagnosis of rotavirus infections - from 8% to 25% (3.13 times). |
| 016 | Coronaviruses | The correct diagnosis when using the tablet was 25%, without using the tablet, no one was tested correctly |

In Figure 6: green indicates that this student has answered the question correctly, and at the end of the table is information on the questions correctly answered and their proportion compared to the total number of tasks.
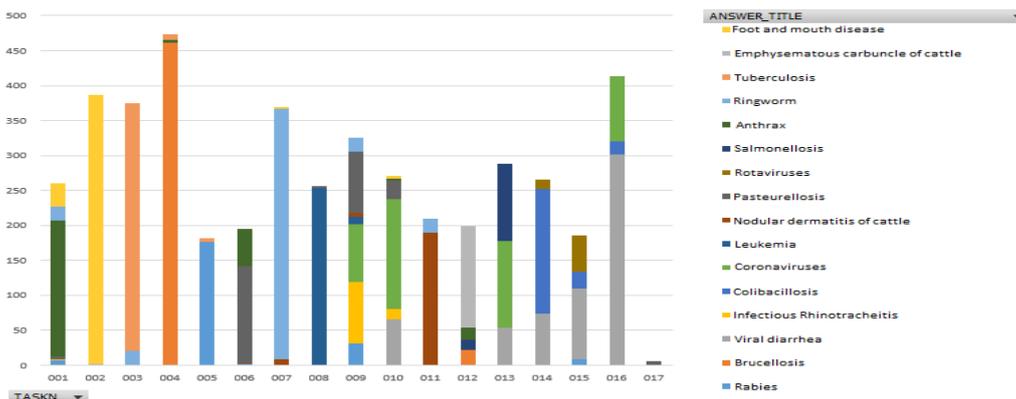


Figure 7:  Diagram of Correct Answers Quality Representation

The second group of students worked with the "Veterinary Tablet". This group worked on the same test tasks as the first one. Figure 7 provides information on the quantity and quality of correct answers. After calculating and analysis of the data obtained, we concluded that the percentage of correct answers using the "Veterinary Tablet" was 69%.

As the diagram in Figure 7 shows, most students answered the questions correctly. Colors show the ratio of their answers (diagnosis) depending on the task number. The right column of the diagram shows the colors that correspond to a certain diagnosis of the disease, for example, brucellosis is indicated in orange, and etc.

We have analyzed the most common symptoms that students chose when answering questions using the tablet, analyzed the number of selected symptoms and their types [3].

According to Figure 8, the most frequent symptoms as a result of the test were: fever, stomatitis, salivation, lameness. The rest of the symptoms were less frequent.

After comparing all the data obtained, we came to the conclusion about the probability of making the right diagnosis in cases with and  without the application. This information is shown in Table 1.

Table 1 shows the changes characteristic in the disease diagnosis using tablet versus diagnosis without Tablet. The analysis of changes in correct answers taking into account each symptom and set of symptoms is performed.

The average time to answer each question has also been calculated. This information is reflected in Table 2.

According to Table 2, we can conclude that the average time to conduct diagnostics for known symptoms is 2-5 minutes. There are no dependencies between the quality of the diagnosis and the time spent on the test task.

For diseases such as leukaemia, tuberculosis, anthrax, rotavirus infections, a clear improvement in the correct diagnosis with help of the veterinary tablet was found (a total of 12 out of 16 test questions showed an improvement in the quality of diseases diagnosis). Separately considering rotavirus infections, the developed software allowed the correct diagnosis of diseases in a quarter of cases, given that without the veterinary tablet no test subjects in this case answered the questions correctly.
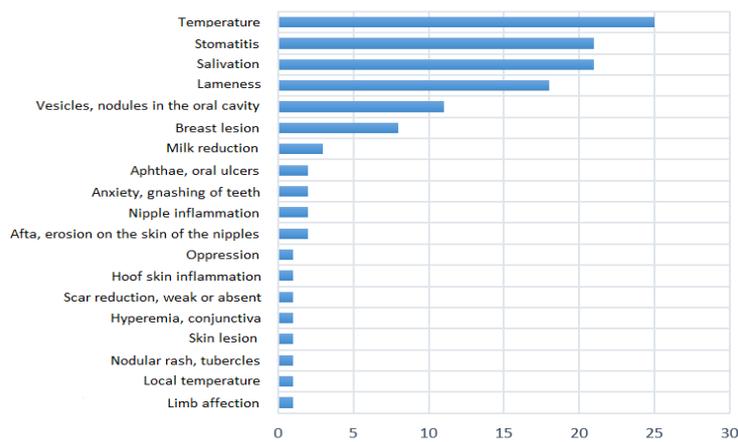


Figure 8: Symptom Distribution Using "Veterinary Tablet"

Table 2: Characteristics of disease diagnosis time with the use of veterinary tablet

| Row names | Time to make an erroneous diagnosis (seconds) | | | Time to make the correct diagnosis (seconds) | | |
|---|---|---|---|---|---|---|
| | Min. | Avg. | Max. | Min. | Avg. | Max. |
| 001 | 72 | 255 | 938 | 82 | 213 | 375 |
| 002 | - | - | - | 62 | 125 | 284 |
| 003 | 181 | 181 | 181 | 65 | 113 | 201 |
| 004 | 123 | 301 | 478 | 61 | 165 | 540 |
| 005 | - | - | - | 64 | 163 | 771 |
| 006 | 95 | 185 | 700 | 61 | 142 | 294 |
| 007 | 111 | 201 | 338 | 62 | 148 | 539 |
| 008 | 89 | 137 | 185 | 64 | 162 | 622 |
| 009 | 86 | 246 | 937 | 94 | 255 | 584 |
| 010 | 71 | 207 | 456 | 78 | 160 | 346 |
| 011 | 83 | 110 | 136 | 60 | 132 | 276 |
| 012 | 111 | 157 | 205 | 62 | 244 | 1255 |
| 013 | 70 | 227 | 900 | 75 | 112 | 198 |
| 014 | 60 | 93 | 146 | 74 | 142 | 313 |
| 015 | 63 | 159 | 632 | 64 | 129 | 243 |
| 016 | 77 | 123 | 319 | 99 | 131 | 185 |
| **Total:** | **92** | **184** | **468** | **70** | **158** | **439** |

Summing up, we can conclude that the software implemented has improved, on average, the results of correct diagnosis from 42% to 69% [3].

In the process, some shortcomings of the selected software testing method were revealed. Firstly, it is the impossibility of veterinarians to work directly with the studied animal directly, as well as the limited initial data on disease symptoms and the lack of visibility of the whole picture in general.

## 7. DEVELOPMENT OF AN ALGORITHM FOR PROCESSING DIGITAL DATA FOR MONITORING THE PRODUCTIVITY OF FARM ANIMALS WITH ELEMENTS OF MACHINE LEARNING

One of the main tasks of the work is the development of software for predicting the productivity of farm animals using data from the smaXtec bolus system and data on milk production of animals. The smaXtec system allows real-time reception of such indicators as: pH, body temperature of the animal, measurement of movement activity, etc. The bolus is injected into the rumen of the animal, and then enters the ruminant stomach, the reticulum, and transfers the data from there. An orally administered bolus measures temperature and activity (via an accelerometer), continuously at 10-minute intervals, while the activity measurement is independent of rumen mobility. Typical increases in activity during sexual activity are detected immediately and trigger notifications. The individual activity levels of the cow are taken into account when processing the data. Fever events are presented to the farmer as a graph or list in the cow's profile and in the fertility section of the dashboard. In this way, the history of previously successful inseminations of dairy cows can be documented in the software for calculating the expected lactation. We used the data of the cattle herd LP "Mambetov and K", consisting of 800 heads of the "simmental" population in the period from 2019 to 2020. (for about 30 thousand data points for all animals).
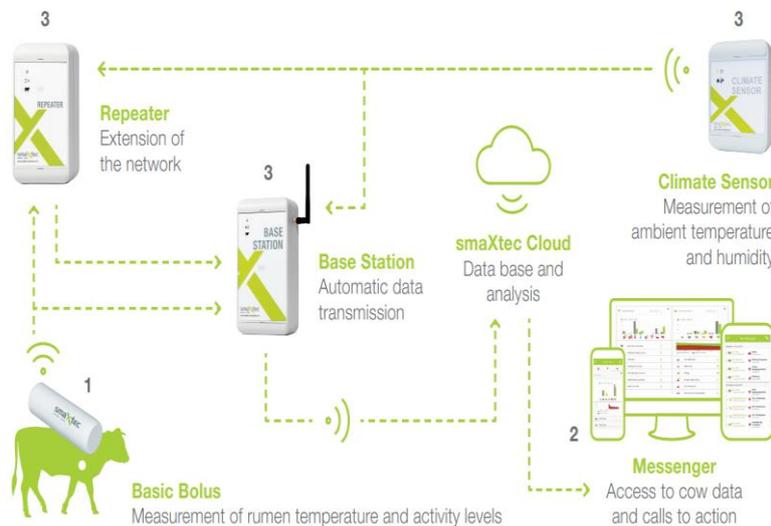


Figure 9: Architecture of the smaXtec system. The smaXtec solution consists of a reticuluminal bolus (1), an online data platform (2), infrastructure including readers and climate sensors (3).

Table 3: Data for each animal

| Animal ID | Date | Body temperature (° C) | Normal body temperature (° C) | Index of the activity | Body temperature without water consumption (° C) | Milk yield (kg)) |
|-----------|------|------------------------|-------------------------------|-----------------------|--------------------------------------------------|-------------------|
| DE0667033081 | 18.10.2019 | 39.19 | 40.000000 | 13.223118 | 39.826875 | 6.35 |
| DE0667033081 | 18.10.2019 | 39.19 | 40.000000 | 13.223118 | 39.826875 | 10.34 |
| DE0667033081 | 19.10.2019 | 39.34 | 40.000000 | 14.251181 | 39.871319 | 6.45 |
| DE0667033081 | 19.10.2019 | 39.34 | 40.000000 | 14.251181 | 39.871319 | 0.00 |
| DE0667033081 | 19.10.2019 | 39.34 | 40.000000 | 14.251181 | 39.871319 | 10.49 |
| ... | ... | ... | ... | ... | ... | ... |
| DE0953378924 | 21.12.2020 | 38.69 | 40.000000 | 10.522472 | 39.016806 | 0.00 |
| DE0953378924 | 21.12.2020 | 38.690951 | 40.000000 | 10.522472 | 39.016806 | 4.99 |
| DE0953378924 | 23.12.2020 | 38.826500 | 40.000000 | 14.486500 | 39.186528 | 5.14 |
| DE0953378924 | 23.12.2020 | 38.826500 | 40.000000 | 14.486500 | 39.186528 | 0.00 |
| DE0953378924 | 23.12.2020 | 38.826500 | 40.000000 | 14.486500 | 39.186528 | 4.77 |

In accordance with the tasks set, such indicators were selected:
1. Body temperature of the animal;
2. Normal body temperature of the animal;
3. Activity index;
4. Body temperature without drinking cycle;
5. Milk yield;

Thus, a library in the Python programming language was developed to extract data from the smaXtec system. The library retrieves data by using REST API technology.

Based on the data obtained and the purpose of the work (forecasting milk production), the most suitable methods are: "Learning with a teacher", exactly regression methods and forecasting methods.

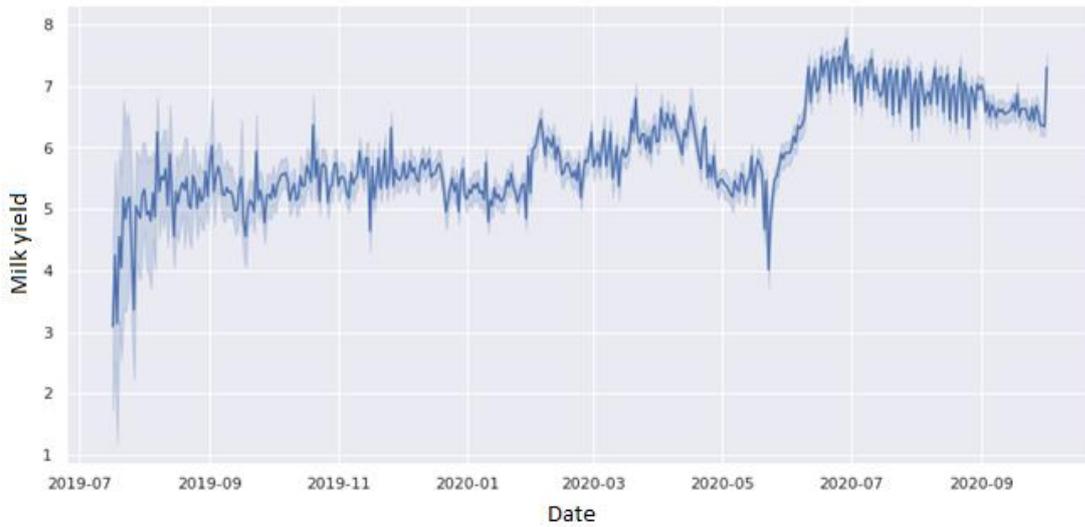According to the graph of the distribution of milk production by day (Figure 10), you can see useful data.

Figure 10: Distribution of milk yield sorted by days.

The stage of applying machine learning methods follows after the initial analysis and data preprocessing. Based on the data, the most suitable machine learning method is polynomial regression.

Figure 3 is a block diagram of the overall software architecture.

The software architecture consists of 3 main components:
1)      System for remote monitoring of the condition of animals smaXtec;
2)      Computing part (machine learning);
3)      Interface for the farmer;
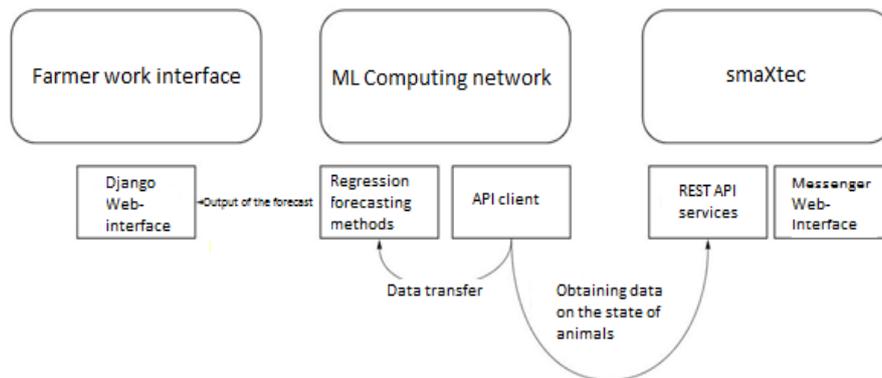


Figure 11: General architecture of the software.

A simple and straightforward architecture makes it faster to make any changes, as well as easier to integrate with other systems.
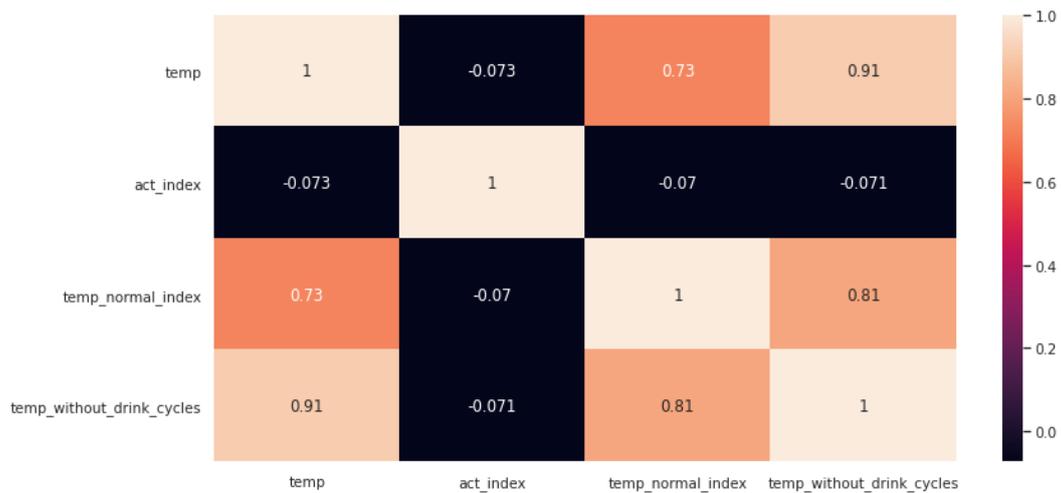
Figure 12: Correlation matrix of the obtained data.

As a result of visualization of the correlation matrix in the form of a heat map, presented in Figure 12, we concluded that there is no statistical relationship between such parameters as: body temperature, normal body temperature, body temperature without a drinking cycle, movement index. In this case, changes in the values of one or more of the above values do not accompany a systematic change in other values.
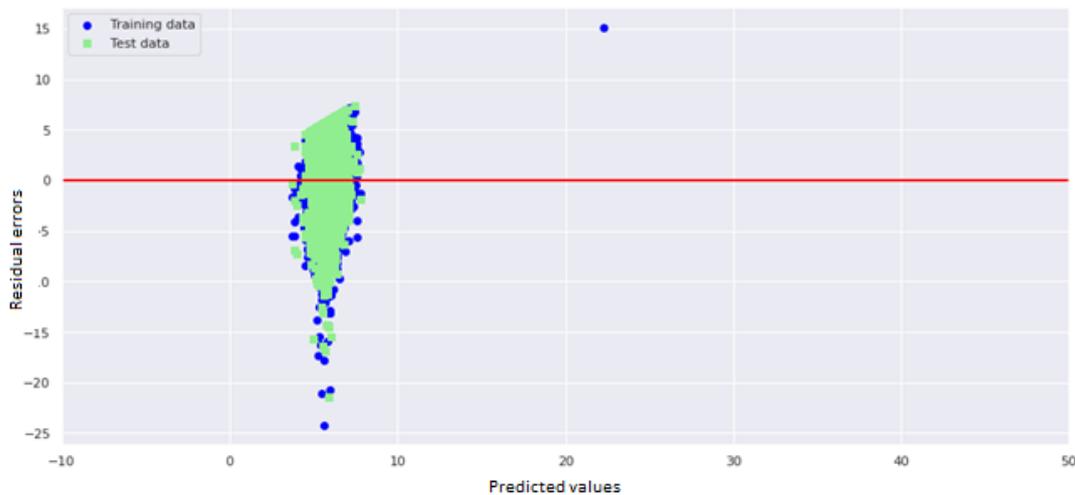


Figure 13: Regression analysis of the obtained data.

As a result of the conclusions made, we came to the need to use a simple linear regression for possible dependencies of the parameters of the initial data and the need to predict the volume of milk yield. In Figure 13, we see that the resulting errors do not have a normal distribution and the data variations around the regression line are not constant. Thus, prediction of the milk production of animals based on the above parameters (body temperature, normal body temperature, body temperature without a drinking cycle, movement index, milk yield) is also impossible.

Thus, we can conclude that the initial data (body temperature, normal body temperature, body temperature without a drinking cycle, movement index, milk yield) obtained from the smaXtec

bolus system and data on the milk production of animals are not dependent. As a result, there is no need to use the initial data parameters for the effect on the volume of milk yield.

## 8. CONCLUSIONS

Upon analyzing the problem of cow diseases diagnostics, we came to the conclusion that it is necessary to develop an expert system of cattle diseases diagnostics. When setting the main tasks to build an expert system, the main one was to determine the input and output data of this system. By using the induction method, we have identified separate groups of input and output data, which will be used to build this system. The next stage was the creation of a generalized web architecture, with the indication of individual functional blocks and equally developed the basic scenario of the use of the intellectual system. When diagnosing a disease, the way the knowledge base is presented plays an important role, which in turn depends on the experience of team of veterinarians. Information on main symptoms and diseases has been collected through questionnaires and this information is structured and presented for better understanding. Thus, a model of knowledge representation has been developed, which leads to an accurate diagnosis. Together with a team of veterinarians, each symptom and symptom complex was given the weight coefficients required for a more accurate diagnosis of the disease.

Thus, we can conclude that the developed expert system for addressing veterinary medicine challenges is effective. By comparing the percentage ratios of the results of the questionnaire of two groups, it becomes obvious that its use is expedient. A detailed analysis of the test subjects' answers has been made and all regularities in both cases of testing have been taken into account. Conclusions were made that the process of diagnosing diseases is simplified in terms of speed of decision-making and their reliability, a direct correlation between the number of detected initial symptoms of the disease and the correct formulation of the diagnosis was revealed. Also, with the participation of veterinary students, an evaluation of the user interface was conducted, which included checking system design and correct compilation of the knowledge base to meet user requirements.

In summary, the developed software has shown its need for use. In future, the database of expert system on diseases and symptoms will be expanded, all deficiencies related to the convenience of the user interface and the operation of the program in general will be taken into account and eliminated.

An expert system under development provides information on 16 major infectious diseases and 103 symptoms, which is currently being developed and populated in the database. The development works are carried out in the S. Seifullin Kazakh Agro Technical University, at the faculties of computer systems and veterinary medicine.

In the future, work will also be done to integrate the system under development with existing animal control systems, to automate the processes of their interaction and data exchange. It is necessary to work with symptoms and weight values, that is, it is necessary to work on choosing an index of the significance of the manifestation of a symptom, instead of its usual manifestation or absence.

part of the project "Transfer and adaptation of innovative technologies to optimize production processes at dairy farms in Northern Kazakhstan" (BR06349515).

## REFERENCES

[1] E. Tullo, I. Fontana, D. Gottardo, K.H. Sloth, M. Guarino. Technical note: Validation of a commercial system for the continuous and automated monitoring of dairy cow activity// J.Dairy Sci. 99:7489–7494. http://dx.doi.org/10.3168/jds.2016-11014. American Dairy Science Association, 2016.

[2] Emre Aydemir, İnci Bilge. Automation Applications in Integrated Animal Production System. Turkish Journal of Agriculture - Food Science and Technology, 8(3): 643-644, 2020. DOI: https://doi.org/10.24925/turjaf.v8i3.643-644.3133.

[3] O. Shopagulov, I. Tretiakov, A. Ismailova, "An expert system for diagnosis cow diseases," Journal of Theoretical and Applied Information Technology, No 15, Vol.98., 2020, pp. 3106-3115.

[4] Ermekov Aidar. Meat March, 2013, [Online]. Available at: http://mk-kz.kz/article/2013/02/11/810619-myasnoy-marsh.html/ (in Russian).

[5] Committee on Statistics of the Ministry of National Economy of the Republic of Kazakhstan: [Online]. Available at: http://www.kazagro.kz/analiticeskij-obzor-po-zivotnovodstvu/ (in Russian).

[6] The Committee on Statistics of the Ministry of National Economy of the Republic of Kazakhstan: [Online]. Available at: https://www.zakon.kz/4951625-obem-veterinarnyh-uslug-v-kazahstane.html/ (in Russian).

[7] L. I. Zubkova, "The effect of diseases of the udder on the milk productivity of cows," Dairy and beef cattle breeding, vol. 4, 2005, pp. 35-37. (in Russian).

[8] H. Qin, J. Xiao, X. Gao, H. Wang, "Horse-Expert: An aided expert system for diagnosing horse diseases," Veterinary Sciences, vol. 4, 2016, pp. 907-9015.

[9] M. Dorosh, Diseases of cattle, Veche, Moscow, 2007, 7 p. (in Russian).

[10] Yu. N. Kozlov, N. M. Kostomakhin, Genetics and selection of farm animals, Kolos, Moscow, 2013, 100 p. (in Russian).

[11] O. V. Zavyazkin, Breeding and keeping cattle, BAO, Kiev, 2012, 100p. (in Russian).

[12] Fu Zetian, Xu Feng, Zhou Yun, Zhang Xiao Shuan, "Pig-vet: a web-based expert system for pig disease diagnosis," Expert Systems with Applications, vol. 29, pp. 93-103, 2005.

[13] Daoliang Li, Zetian Fu, Yanqing Duan, "Fish-Expert: a web-based expert system for fish disease diagnosis," Expert Systems with Applications, vol. 23, pp. 311-320, 2002.

[14] Paolo Liberati, Paolo Zappavigza, "Improving the automated monitoring of dairy cows by integrating various data acquisition systems," Computers and electronics in agriculture, vol. 68, pp. 62-67, 2009.

[15] D. Rice, Common dog diseases and health problems 4-H Companion Animal Health, 2014, [Online]. Available at: https://www.extension.purdue.edu/extmedia/4H/4-H-852-W.pdf/

[16] E. B. Hunt, Artificial Intelligence, New York, San Francisco, London, Academic Press, 558 p, 1975.

[17] M. I. Makarov, V. M. Lokhin, Intelligent Automatic Control System, Moscow: Fizmatlit, 2001, 576 p. (in Russian).

[18] T. A. Gavrilova, V. F. Khoroshevsky, Knowledge Base of Intelligent Systems, St. Petersburg: Piter, 2000, 384 p. (in Russian).

[19] D. Zeldis and S. Prescott, "Fish disease diagnosis program – Problems and some solutions" Aquacultural Engineering, vol. 23, no. 1–3, 2000, pp. 3-11.

# AUTHORS

**Shopagulov Olzhas Almatovich**, Doctoral student, Department of Information Systems, Kazakh Agro Technical University named after S.Seifullin, Nur-Sultan, Kazakhstan. In 2014 he graduated the Eurasian National University named after L.N. Gumilyov, received a bachelor's degree in radio engineering, electronics, telecommunications. Also in 2017, he graduated from S. Seifulin Agro Technical University with a radio engineering, electronics, telecommunications master's degree. He is project assistant on the development of an expert system for solving the problems of veterinary medicine (on the example of a dairy farm in Northern Kazakhstan). His research interests include the development of information systems, researches related to big data, as well as works related to data transfer.

**Tretyakov Igor Igorevich** recieved bachelor in technics and technologies from the Eurasian National University, Astana, Kazakhstan, in 2009. From 2009 to 2012, he was a software developer of various state information systems for electronic government. From 2012 to 2016 he was a software developer, data analyst and project manager for project in animal husbandry, animal recording in Kazakhstan. Participated in projects for animal genomic (gBLUP) and index evaluation in collaboration with INRA - French National Institute for Agricultural Research (dairy cattle) and Agricultural Business Research Institute, Australia (beef cattle). Since 2018, he is a lead of IT group in S.Seifullin Kazakh Agro Technical University in the "Transfer and adaptation of innovative technologies to optimize production processes at dairy farms in Northern Kazakhstan" project. His interests include software development for web and mobile, automation systems and data analytics in agriculture.

**Ismailova Aisulu Abzhapparovna**, Doctor of Philosophy (PhD), Senior Lecturer. Received the bachelor degree in Physics and Mathematics from the Korkyt Ata Kyzylorda Pedagogical Institute, in 1993. Received the master's degree in Information Technology from the Eurasian National University named after L. N. Gumilyov, in 2006. From 2011 to 2014, she graduated a doctoral program at the Faculty of Information Technology at the Eurasian National University named after L.N. Gumilyov. His research interests are 3D modeling, working with spatial data, development of methods and software for creating information systems.