

## СЪЗДАВАНЕ НА ЕКСПЕРТНА СИСТЕМА ЗА РАННА ДИАГНОСТИКА И ПРЕВЕНЦИЯ НА МАСТИТ

### DEVELOP OF EXPERT SYSTEM FOR EARLY DIAGNOSIS AND PREVENTION OF MASTITIS

**Yuri Stoynov**

*St. Cyril and St. Methodius University of Veliko Tarnovo,  
Bulgaria,  
juri@celtis.net*

#### **Abstract**

*Expert systems (ES) are one of the prominent research domains of artificial intelligence (AI). They are applications developed to solve complex problems in a particular domain, at the level of extra-ordinary human intelligence and expertise. We present design and development of expert system for data collection, analysis and decision making for early mastitis detection. It focuses on both milk quality and animals health.*

**Keywords:** early mastitis diagnosis, mastitis prevention, expert system, information technologies.

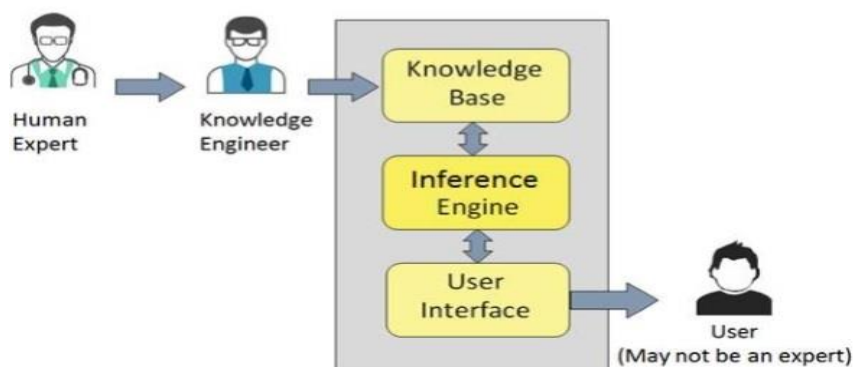
#### **INTRODUCTION**

Expert systems (ES) are characterised with high performance, high responsiveness and reliability combined with ease of use and ease of understanding, which makes developing of such system challenging and involves different type of specialists [1].

The ES are capable of advising, instructing and assisting humans in decision making. Demonstrating, explaining and interpreting inputs, deriving solutions, diagnosing problems, predicting results and justifying conclusions. Suggesting different options for solving the problems in their domain. Despite their capabilities, they also have all kinds of limitations, most significant of them being that they can not refine their own knowledge base, and that they can not produce accurate output for inadequate knowledge base. This makes them incapable of substituting human decision making. Expert systems are well integrated in the animal health management [2], [3], [6], [7], [8].

#### **COMPONENTS OF EXPERT SYSTEM**

ES are composed of three main components that are responsible for different things and are interacting in a different ways [9]. They are the knowledge base, the interface engine and the user interface (fig.1 ).



*Figure. 1. Components of ES*

## KNOWLEDGE BASE

Knowledge is termed as data, information and past experience combined together. The data is a collection of facts, and the information is organised as data and facts about the task domain.

The knowledge base should contain domain-specific and high-quality knowledge. The success of an ES majorly depends upon the collection of highly accurate and precise knowledge. The knowledge base in ES is stored of both factual and heuristic knowledge. Heuristic knowledge is about practice, accurate judgment, and ability of evaluation and informed guessing.

Knowledge representation is the method used to organise and formalise the knowledge in the knowledge base, usually in the form of if-then-else rules [4].

The success of any ES majorly depends on the quality, completeness and accuracy of the information stored in the knowledge base. The knowledge base is formed by collecting information from different experts in the specific domain, scientists and the knowledge engineers. The knowledge engineer is a person with case analysing skills and a quick learner. Knowledge engineer is responsible of categorisation and organisation of the information in a meaningful way, often in the form of if-then-else constructs to be used by the inference machine. He also monitors the development of the ES and is responsible for the quality control.

## INFERENCE ENGINE

The use of efficient procedures and rules by the inference engine is essential in deducing a correct and flawless solution. In case of knowledge based ES the inference engine acquires and manipulates the data in the knowledge base to arrive in particular solution. In case of rule based ES it applies rules repeatedly to the facts, which are obtained from earlier rule execution. It might also add new knowledge to the knowledge database if required and resolve

rule conflicts when multiple rules are applicable to a particular case. In order to recommend a solution, the engine might use forward chaining or backward chaining strategy.

Forward chaining is a strategy based on the question “What can happen next?”. The Inference engine run the chain of conditions and derivation in order to deduce the outcome. It considers all the facts and rules, and sorts them before concluding a solution. This strategy is preferred when the ES is trying to make a conclusion, calculate result or prediction of a result as an effect of changes of some system parameter. For example trying to predict the effect of increases of greenhouse gasses to the climate.

With backward chaining strategy the ES is trying to find the answer of the question “Why this happened?”. On the basis of what has already happened, the Inference tries to find out which conditions could have happened in order to end up with this result. It is used to find the cause for a reason. For example reasons for the increasing rates for type 2 diabetes.

As expert systems evolved, many new techniques were incorporated into various types of inference engines. Some of the most important are:

**Truth maintenance.** These systems record the dependencies in a knowledge-base so that when facts are altered, dependent knowledge can be altered accordingly. For example, if the system learns that Socrates is no longer known to be a man it will revoke the assertion that Socrates is mortal.

**Hypothetical reasoning.** In this, the knowledge base can be divided up into many possible views, a.k.a. worlds. This allows the inference engine to explore multiple possibilities in parallel. For example, the system may want to explore the consequences of both assertions, what will be true if Socrates is a Man and what will be true if he is not?

**Fuzzy logic.** One of the first extensions of simply using rules to represent knowledge was also to associate a probability with each rule. So, not to assert that Socrates is mortal, but to assert Socrates may be mortal with some probability value. Simple probabilities were extended in some systems with sophisticated mechanisms for uncertain reasoning and combination of probabilities.

**Ontology classification.** With the addition of object classes to the knowledge base, a new type of reasoning is possible. Along with reasoning simply about object values, the system could also reason about object structures.

## **USER INTERFACE**

The user interface (UI) provides an interaction between user of the ES and the ES itself. The user of the ES does not need to be an expert to use the system. The UI needs to have some clear and user friendly way to display how the system reached to a particular conclusion.

The UI should also make it easy for the user to trace the credibility of the deductions that lead to the final decision. It should be designed towards the users, in order to allow them to accomplish their goals in the shortest possible way, to work with their existing or desired work practices, and makes efficient use of the users input. The technology should be adapted to the user’s requirement, and no the other way around.

## **APPLICATIONS, LIMITATIONS, BENEFITS**

Expert systems like any other technology have its limitations. Large systems are costly, because they require significant development time and compute resources. Basic limitation can

be classified in four basic categories. Limitation of the technology, Difficult knowledge acquisition, they are difficult to maintain, and have high development costs.

Expert systems are widely used in design, medicine, monitoring systems, process control systems, knowledge domain, finance and commerce, for things like camera lens design, automobile design, diagnosis systems to deduce cause of disease from observed data, leakage monitoring in long petroleum pipeline, finding out faults in vehicles, computers, detection of possible fraud, suspicious transactions, stock market trading, airline scheduling, cargo scheduling.

Main benefits from such systems include their availability, less production cost, often they provide great speed, and are able to reduce the amount of work, their error rate is less compared to human errors, they reduce risk, since they can work in environment that is not human friendly and they provide a steady response with-out getting motional tensed or fatigued.

## **DEVELOPMENT OF ANIMAL MONITORING EXPERT SYSTEMS**

The process of ES development is iterative. Steps in developing the ES include:

- Identify a problem domain. This includes checking if the domain is suitable for an expert system to solve it, do we have access to experts in the area of the domain, determine the cost-effectiveness of the system in the given domain.
- Design the system. This includes identifying the ES technology, establishing a degree of interaction with other systems and databases, investigation how the concepts can represent the domain knowledge best.
- Develop a prototype. This includes acquiring the domain knowledge from the experts and representing it the needed format for the inference engine.
- Test and refine the prototype. This includes using sample cases to test the prototype for any deficiencies in performance
- Develop and complete the ES. This includes tests to ensure that the interaction of the ES with all elements of its environment, including end users, databases, and other information systems, documenting the ES project, training the users to use the ES.
- Maintain the system. This includes keeping the knowledge base up to date by regular review and updates, search for integrations with other information systems as those systems evolve.

The animal monitoring system is planed as expert system for the dairy industry. It is mainly targeted for cows but with adaptation of the algorithm should be able to work with goats and buffalos. It focuses on both milk quality and animals health. It tries to detect early symptoms of mastitis, based on specific physical qualities of the milk samples and start early treatment of the animal and early separation of infected milk, preventing it from mixing with other milk in the tanks which can potentially lead to infecting the whole tank with pathogens and cause significant loses for the farmers from the order of thousands of euro depending on the tank size. In worst case it can reach the market and have an effect on the public health since milk and other diary products have the potential to transmit pathogenic organisms to humans [5].

Cost effective equipment is used to feed the data into the system. The selected equipment for field measurements is EkoMilk Horizon hybrid mid-infrared and ultrasound analyser. Clustered version of it with sample feeder is able achieve 350 milk samples per hour

for price 0.04 euro per test, which makes it perfect choice for a mini lab. Milking robot version that is also available can work fully unattended as in farm setup controlled by the milking robot automation software [10]. With the partners from BulTeh 2000 we were able to add a wifi module to all modifications of the EkoMilk Horizon analyser and connect it to the internet. That allowed us to feed the system with real-time raw data from each measurement made with the analyser.

## COMMUNICATION AND DATA STORAGE

Communication between the analyser and the expert system is made possible over internet with the integrated ESP8266 wifi module from Olimex [11]. Wifi connection has to be set up from the analyser menu. Once the connection is made, and instrument have the connection to the internet, HTTP protocol is used for communication with the system, since it is widely used and all network equipment including firewalls are preconfigured to allow this kind of traffic. Cryptographic hashes are used for signing every message. This ensures the message integrity and validates and authenticates the sender instrument in the system. Each instrument has to be registered in the system using its AMP identification code that is printed on the box and on the instrument label. Four digit pin is also associated with it. This gives us another layer of security and guarantees that fake data could not enter the system and compromise its operation and results.

In order to standardize the communication and make it compatible and extendable for this and future generation equipment, the AMP 1.0 protocol was defined. One of the main goals of the developed protocol is to confirm the identity of the sender as well as the integrity of the data, since the HTTP protocol is vulnerable to man in the middle attacks. To achieve this we use two hashing functions - the faster MD5 for signing the message because of it's size, and the more reliable SHA1 for creating the identity token, which the server uses to verify the identity of the terminal device and determine the access rules for that device. The terminal device calculates the token for every message sent to the server and attaches its value as HTTP header. This way the token is different for every single HTTP request.

Once the data is received and all signatures checked, it is stored securely in the system database. Blockchain proof-of-work technology and smart contracts based on the Ethereum API [12] are used to additionally secure the measurements data, in a way that any further modification of any of the parameters of the measurement can be detected and the affected measurements and decisions that are made based on them invalidated.

Using the same communication channel the analyser send all errors and log messages that happened during operations. They are also stored in the system for future analysis, and to find some correlation between them and problems in the measured results. The information system use them for notifying the users or warn them about missed procedures that might have negative effect on the measurements and break the warranty of the analyser. Notification are especially helpful when the instrument is working in unattended mode paired with milking robot. This way the operator will know when it is time for maintenance or when his attention is needed because of errors.

Another important part of collecting the data is the identification of the sample and the animal or tank from where the sample is taken. In most cases the animal identification code comes with the measurement. In rare cases, when the animal or tank identification data is missing by the time of the measurement, user will need to interact with the system and do the identification manually via appropriate template that he can download, filled with the animal

or the tank identification number next to the measurement sequence number, and upload back to the system. Using the unattended or the manual identification process, The system needs this identification data, because it later plays important role connecting the newly made measurement with the existing knowledge base.

## KNOWLEDGE BASE

The system keeps a big knowledge base consisted of different types of data. Some of the important entities are the instruments and their full history of raw measurements, their logs and errors and also their location. As it turns out the location is very important parameter, since there are different practices in the dairy industry in the different countries. The cultural, technological and legal differences in the different countries gives significant changes in the results calibration and treatment of the animals. The difference is mostly noticed between Western farms and the farms in the developing countries in Asia. There are sometime differences between farms in the same country, this is why we have also information about the farm in the knowledge base. At the end the hope is that the decisions of the expert will be consistent with the local practices and compliant with local market and best practices of the vets that are serving the given farm.

Each animal has a lot of meta data initially entered and later collected in the system. We are tracking also the animal species, but for now the research is based on cows. The inference engine takes into account parameters such as breed, heard, the barn that the animal lives in when, full history of previous treatments and previous measurements when it is deducting treatments or calibrating the results. We also often know the parents of the animal and their full history, when the cow is born in the farm under the control of the expert system.

The system also distinguish between tank and animal milk and keeps different parameters for the tanks. For example it knows the location - the country and region, the volume, the milk origin by animal and also it has full history of previous measurements over time and all the tank maintenance data.

One of the most important information in the knowledge base are the old measurements. There is a lot of meta data connected to each measurement, starting with the raw data about physical properties of the sample from the analyser, the official lab milk quality results for the same sample, the exact analyser used, is it cow or tank sample and which, and in case of cow we store also information for the udder quarter of the animal, the time that he sample is taken and some others. The first decision that the expert system needs to do on every measurement is how to properly calibrate it.

The other very important information is the history of the treatments for each individual animal and the affect of it to the sub-sequential milk samples. The systems keeps a record when and with what medicine the animal was treated and know how effective the treatment was and the illness progressed. This information is used from the experts and from the system itself to fine tune the prescribed treatments in the future.

The system stores also full history of the official lab results to a set of control samples which are used in the first step of the data processing called measurement calibration These results are uploaded by the farm owners, vets or the labs themselves in machine readable format.

## MEASUREMENT CALIBRATION

Once we receive the data from the measurement and we have the sample identified, the expert system can start a process that we call measurement calibration (fig. 2). This is process of deduction of the basic milk quality parameters that official lab measures based on the physical milk properties that the analyser reads. Milk quality parameters includes Fat, Protein and the main milk quality indicator - somatic cells count [13]. The analyser has ultrasonic, infra red, conductivity and other sensors and also measures the flow time at a given temperature.

|               |   |   |   |      |      |     |      |      |      |      |
|---------------|---|---|---|------|------|-----|------|------|------|------|
| Farm1<br>Tank | – | – | A | 22.9 | 22.9 | 140 | 4.64 | 4.17 | 5.50 | 9.04 |
| Farm1<br>Tank | – | – | A | 22.7 | 21.0 | 146 | 4.63 | 3.54 | 5.48 | 9.02 |

*Figure. 2. Calibrated and uncalibrated measurement*

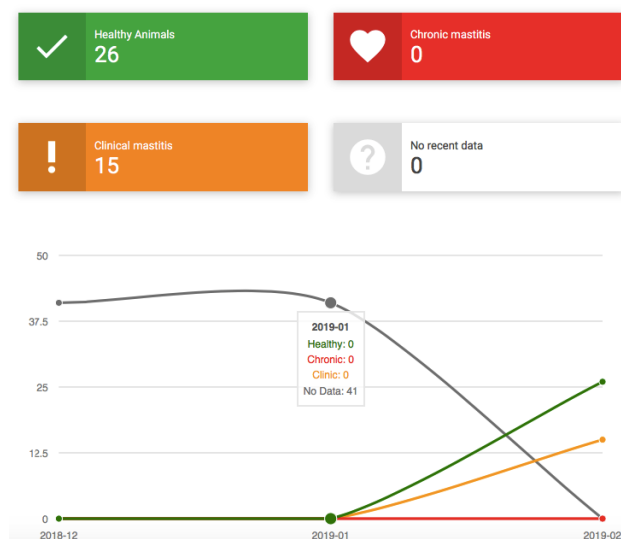
The inference engine (fig. 3) gets various types of data from the knowledge base in order to calibrate the measurement correctly. It uses the instrument measurements count to account for sensors wearing, the country, the farm and the animal but mostly the physical properties mapped to the official lab milk quality results, which are constantly updated from control samples sent to the lab on a recurring schedule.

The screenshot shows a 'Create Rule' form with the following fields:

- Farm:** All farms
- Type:** User Defined
- Result Type:** Viscosity (LT[s])
- Compare Type:** Greater (>)
- Value:** 5
- Message:** blockage is possible

*Figure. 3. Adding a rule to inference engine*

Every value that the inference engine pulls from the knowledge base is checked against the normal operating ranges of the system. The operating ranges are also stored in the knowledge base in the form of rules. These rules are added by the experts in the area and the analyser manufacturers, and if such abnormal conditions are detected the inference engine is



**Figure. 4.** UI for managing calibration tables

marking the sample as unreliable and attach a warning message to the measurement with a suggested maintenance procedure that needs to be performed on the analyser. This flow as many others triggers an action in the notification engine in the expert system.

| Name                      | Created At          | Updated At          | Farms | Instruments |    |    |   |
|---------------------------|---------------------|---------------------|-------|-------------|----|----|---|
| Initial                   | 2018-04-16 07:32:43 | 2018-07-02 09:13:36 | 0     | 13          | 🗑️ | ✎️ | ➔ |
| Africa March 18           | 2018-05-02 15:46:02 | 2018-05-24 08:38:36 | 0     | 0           | 🗑️ | ✎️ | ➔ |
| Africa 24May18            | 2018-05-24 08:37:18 | 2018-05-29 11:09:57 | 0     | 1           | 🗑️ | ✎️ | ➔ |
| 06Jun18                   | 2018-06-05 14:22:52 | 2018-07-09 13:52:10 | 0     | 0           | 🗑️ | ✎️ | ➔ |
| SCC corr for UHT and past | 2018-09-26 09:23:09 | 2018-09-26 09:23:36 | 0     | 1           | 🗑️ | ✎️ | ➔ |
| EHE(R) Initial            | 2019-02-08 08:56:43 | 2019-02-08 08:58:25 | 0     | 2           | 🗑️ | ✎️ | ➔ |

**Figure. 5.** Health categories

The notification subsystem is used for notifying the users and other external systems for circumstances that the expert system detects. It uses various channels including API calls to external systems, sms, e-mail messages and in system UI notifications.

If all parameters are in normal working ranges, the engine starts checking its if-then-else constructs in order to select the right calibration table for the given conditions. The calibration tables are stored in the knowledge base, and are added by experts in the area of milk analysis and chemistry. There is also the option to generate a calibration table (fig. 4, 5, 6) based on the data



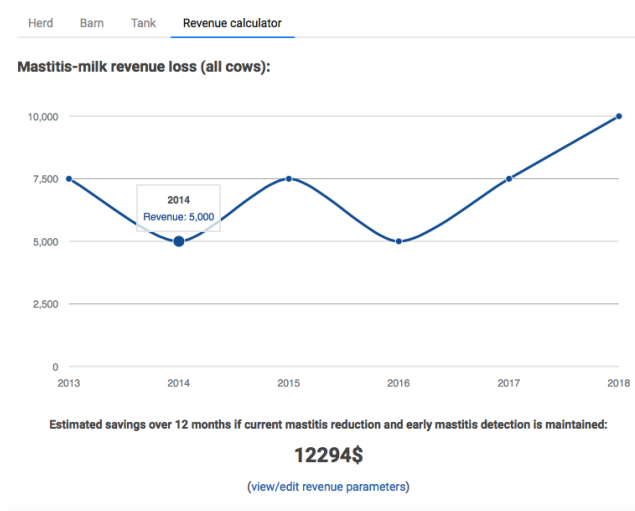
in the knowledge base, and the results can be compared with the tables created from the expert. Selected set of data is fed to an external machine learning service that calculates the calibration table based on the past experience. The system has appropriate user interface for adding, exporting, importing and modifying calibration tables. Once the right calibration table is selected, it is used to calculate the final values for the milk quality parameters, that are added to the measurement and back to the knowledge base.

## HEALTH AND TREATMENT ANALYSIS

The expert system analysis about heard health status is based mainly on the somatic cells count in the milk samples for the given animal over time. The inference engine is working with another set of predefined if-then-else statements, but from different kind of experts - veterinarians. The current set of rules are categorising the animals in four categories - healthy, clinical mastitis, chronicle mastitis and no data available.

If the health status of an animal changes, it triggers an action in the notification engine, and notification to the farmer or veterinarian is sent. The notification contains the appropriate treatment plan for the animal, based on the country or region best practices and law regulations. The local veterinarian experts are consulted for the creation of the rules for each farm, in order to match the best practices so far for the given farmer and heard. In the modern farms the notification engine sends an API call to the milking robot, notifying it that the milk from this animal should not be mixed with the other milk, in order to limit the contamination and limit the losses.

Treatment plans are stored back to the knowledge base, and the animal somatic cell count monitored and if the condition does not improve, the farmer should consult with a veterinarian. The inference engine rules are not still well tested in different control groups in order for the expert system to change the current treatment based on the old treatment plan and the somatic cell count readings after the start of the treatment.



**Figure. 6. Revenue calculator**

Tank milk condition is also monitored, but the data is used just for creating a tank dashboard where specialist can monitor the tank condition over time, work with the current collected data and try to create rules for the tank maintenance.

The system also has a revenue calculator that indicates approximately how much money the farmer saved by the early mastitis detection by calculating the amount of milk that will be lost due contamination and late animal treatment.

## CONCLUSION AND FUTURE WORK

Most of the work already done so far was focused on the data collection and measurement calibrations techniques. Blockchain technology combined with the secure data transfer methodology from the analyser to the system, it can guarantee that the origin of the data is genuine and no external modifications are possible. This will additionally ensure that the input data is correct and prevent the system working with incorrect data which is one of the main problems of the expert systems, because working with corrupted data not only will produce wrong output, but it can contaminate the knowledge base with wrong data and compromise future predictions. The extra defence mechanisms that check that all parameters are in normal operation ranges additionally ensures the reliability of the system, and also helps detecting malfunctions in the analysing equipment. These bad measurement will be excluded from the blockchain and will not contribute to the final system decisions.

The knowledge base about treatment plans and early mastitis diagnosis based on the somatic cell count should be updated with more accurate rules, so the system can prescribe more accurate treatments, and also give it the ability to change treatment plans and suggest termination of animals that have chronicle mastitis that is not affected by the treatments prescribed.

Another not directly connected area with the early mastitis detection in which the system could be expanded is related to the support and maintenance of the milk analysers. With all the data collected the system can monitor important deadlines for maintenance and also detect the common failures in the analysers by catching drifts in the data.

## REFERENCES

- [1] Benson M., 2015. *Handbook of Expert Systems*, Clanrye international.
- [2] Dagnino A., Allen J. I., Moore M. N., 2007. Development of an expert system for the integration of biomarker responses in mussels into an animal health index, *Biomarkers*.12(2):155-72. doi: 10.1080/13547500601037171.
- [3] Enting J., Huirne R.B.M., 1999. A knowledge documentation methodology for knowledge-based system development: an example in animal health management, *Computers and Electronics in Agriculture*, 22(2-3), 117-129.
- [4] Ertel W., 2017. *Introduction to Artificial Intelligence*, Springer International Publishing.
- [5] Hameed, K. Sender G., Kossakowska A., 2006. Public health hazard due to mastitis in dairy cows, *Animal Science Papers and Reports* 25(2), 73-85.
- [6] Jampour M., 2011. A Fuzzy Expert System to Diagnose Diseases with Neurological Signs in Domestic Animal, *Eighth International Conference on Information Technology: New Generations*, DOI: 10.1109/ITNG.2011.217

- [7] Michele R., Scioscia F., Sciascio E. Di, 2009. A knowledge-based framework enabling decision support in RFID solutions for healthcare. *2010 IEEE International Symposium on Industrial Electronics (ISIE)*, 68–71. DOI: 10.1109/ISIE.2010.5637492
- [8] Nuthall P.L., Bishop-Hurley G.J., 1996. Expert systems for animal feeding management. *Computers and electronics in agriculture*, 14(1), 9-22.
- [9] [https://www.tutorialspoint.com/artificial\\_intelligence/artificial\\_intelligence\\_expert\\_systems.htm](https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_expert_systems.htm)
- [10] <https://www.bulteh.com/ekomilk-horizon-unlimited-hybrid-analyzer.html>
- [11] <https://www.olimex.com/Products/IoT/ESP8266/MOD-WIFI-ESP8266/open-source-hardware>
- [12] <https://etherscan.io/apis#contracts>
- [13] <https://dairy.ahdb.org.uk/technical-information/animal-health-welfare/mastitis/symptoms-of-mastitis/somatic-cell-count-milk-quality-indicator/#.XHJ-zS2B1QI>