Cattle disease identification using Prediction Techniques

[1] **Noone Vijay Kishan,** [2] **Sai Trinath Y**, [3] **Sandeep Kavalur**,
[4] **Sangamesha V,** [5] **Mr. Sumanth Reddy**
[1][2][3][4]BE Students, Department of Computer Science and Engineering
[5] Asst. Professor, Department of Computer Science and Engineering
[1][2][3][4][5] Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

Abstract

As one of the earlier methods of occupation, non-industrial nations such as India, Bangladesh, Nepal and a lot more have dairy farming. Dairy farm automation plays a major role in the expansion of productivity in dairy production. Cattle animals are prone to many diseases, some of which can decrease productivity and lower the quality of dairy products and, if not identified at an early stage, can also contribute to the death of cattle, which is greatly impeded by the sustainable development of the national economy.

Significant numbers of cattle are found in many dairies. It is just too hard to take care of them and track the health of the dairy cow. This work is also heavily reliant on the owner of the dairy and municipal authorities. Continuously seeing the health of individual cattle, quickly diagnosing and handling sick cattle as early as possible is the main feature of a health management method. We use sensor technologies to chart the basic aspects of animal activity such as hotness, heartbeat, etc. This data is aggregated and submitted to a data mining model to validate whether any anticipated event of imminent disease is expected. Which raises the lowest level of veterinary inspection and potential costs of animal health care.

This paper presents a technique that explains how the use of IOT and data mining can diagnose cattle diseases that are rare in farm animal medical facilities that can have cost-effective medical solutions.

Keywords: IOT, Data Mining, dairies, health monitoring system, healthcare, economy .

I.INTRODUCTION

In raising a country's economic standing, agriculture plays a significant part. In India itself, Almost 18 percent of the gross domestic product is provided for by the agricultural sector.

Agriculture and the husbandry of livestock both go in parallel. Animal husbandry is an activity where food and non-food items are nurtured by livestock. Milk and its substitutes, eggs and beef are agricultural goods. Bone products, pharmaceuticals, wool, etc. contain non-food products. This includes thorough care of the animals on a regular basis. A part of its productivity comes from animal husbandry as a branch of agriculture. In our country, around 20.5 million people rely on livestock for their livelihoods, meaning that 2/3 of the village population gets their livelihoods from dairy farming. From livestock animals, we get variety of food and non-food products.

An significant factor in the production and preserving the quality of these goods is the wellbeing of livestock. Cattle diseases can, in terms of quality and quantities, have a detrimental effect on productivity. parturient paresis, Ketonemia, tension, limping, fever etc. include major cattle diseases. This diseases can spread rapidly where thousands of cattle are raised together in vast farms, which can cause a significant drop in income.

Many dairy farmers in developed countries are not well paying to have their cattle routinely diagnosed, which can predict disease at an early stage, because there is little medical infrastructure available in many areas. To get to a medical centre, many farmers from villages have to carry their cattle a very long distance away. An early detection with an electronic device that would log all the critical tracking information related to cattle and feed the data to a disease prediction data mining model that would help speed up the healing process and deter accidents due to local unavailability or shortage of veterinary specialists.

Speaking of India, which has a large cattle population, doctors in the rural sectors are robbed. In order to spread

knowledge in villages about the prevention of these cattle diseases, a scheme is required that would lead to a decrease in reliance on medical experts and would also minimize the expense of transporting livestock in major cities or towns during medical conditions.
The method discussed in this paper deals with an approach to providing a broad database of hardware that contains symptoms and records of various health care conditions that can report and document the health condition. With the aid of data mining strategies to classify the situation, this can be done and offers a pre-diagnosis of the issue with data obtained from sensors that are paired with the internal database.

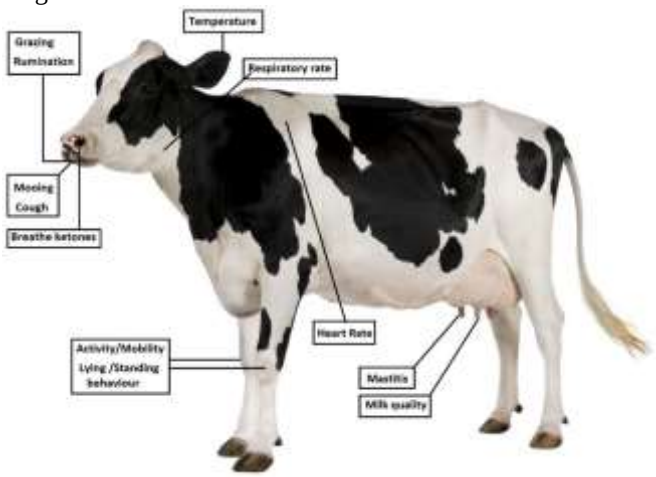

Figure 1. Different monitoring parameters of cattle

The technology allows early diagnosis of the disease and can avoid delays in the recognition of heinous diseases. The system further conducts an intelligent examination of a hardware device's sensor data and detects whether or not the cattle is suffering from a disease. The system uses multiple sensors such as temperature, vibration accelerometers, etc. to measure.
In order to generate results from the obtained data, the method utilizes the intelligent analysis mechanism functionality of data mining. This program is a first aid mechanism that analyzes the signs to send you results with algorithms for the detection of cattle diseases on the basis of data computation.

II.PROBLEM STATEMENT

A disease is a state of illness, it affects the functionality of the body parts. If disease is present in the cattles, this intern causes problems like reduction in the yield of milk, and the economy of the country.
If the present disease is a communicable disease, then the impact is very high. That might destroy all the cattles. The early detection of the disease present is very important to overcome the above problems.

In this research work, we combine data mining techniques and IOT technologies to solve the problems.
The model developed takes the symptoms of the cattles and output from the IOT sensors as input, and does the analysis using data mining algorithms to predict the accurate disease. This model helps in early detection of some of the dangerous diseases.

III.RELATED WORKS

Many review papers have been published on disease prediction using data mining techniques and cattle health monitoring systems. They used diverse data mining methods for prediction and obtained dissimilar outcomes for different methods:
Dr. S. Anitha1 , Dr. N. Sridevi2 in their report on "HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES ", Three different classification methods for data mining are used, namely k-nearest neighbors , Naive Bayes and SVM .to analyse the dataset with 76.67%,86.6%, 77.7% accuracy respectively for predicting Heart Disease.

S.Vijiyarani, S.Sudha in their work on "Disease Prediction in Data Mining Techniques ", In this survey paper numerous algorithms in data mining that are used in the field of medical prediction that have issue of summarizing are debated

K. P. Suresh1, Dhemadri2, Rashmi Kurli3, R. Dheeraj4 and Parimal Roy5 in their work on "Application of Artificial Intelligence for livestock disease prediction",
here they have trained various AI models like GLM, RF, BRT, ANN, CTA, etc for disease modelling.

Chinmayi Chitnis, Roger Lee in their work on "Improving Health-Care Systems by Disease Prediction", they have used Neural Networks to train on a dataset and achieved 80.2% accuracy.

Seema Kumaria , Dr. Sumit Kumar Yadavb in their work on "Development of IoT Based Smart Animal Health Monitoring System using Raspberry Pi", Using the Raspberry Pi3, heartbeat sensor,MCP30008 , body temperature sensor and rumination sensor, it has developed a system capable of monitoring body temperature, heartbeat and rumination in real time.

Vaishnavi Shinde, Prof. Sweta Jha, Amruta Taral in their work on "IOT Based Cattle Health Monitoring System", where they developed a system where In order to ensure their fitness as it directly impacts the health of consumers, they developed constant surveillance of cattle health.

Emi Mwaka Katemboh, Raed Abdulla, Vikneswary Jayapal, Sathish Kumar Selvaperumal, Dhakshyani Ratnadurai in their work on "Integrated Animal Health Care Using Iot", where they developed a system which has massager system which turns on depending on BPM and a fan system which cools the environment when needed.

Data Mining

Within computer science, data mining is defined as an interdisciplinary subfield. It involves discovering patterns through computation in large data sets combining methods used in several disciplines such as machine learning, artificial intelligence, database systems and statistics. The general objective of the data mining process is to reveal hidden predictive information and unknown data, relationships, patterns and knowledge by searching through large data sets where it is hard to search and determine using classical methods of statistics.

Data mining comprises three steps of knowledge discovery in databases (KDD); preprocessing, modeling and post-processing of data (24). The goal of pre-processing is to prepare raw data for mining. The relationships between various data are discovered in the modeling step, in order to extract a pattern. In the post-processing step, the extracted pattern is evaluated in order to be verified and stated as knowledge. For the interactive and iterative KDD process model shown in Figure 2.

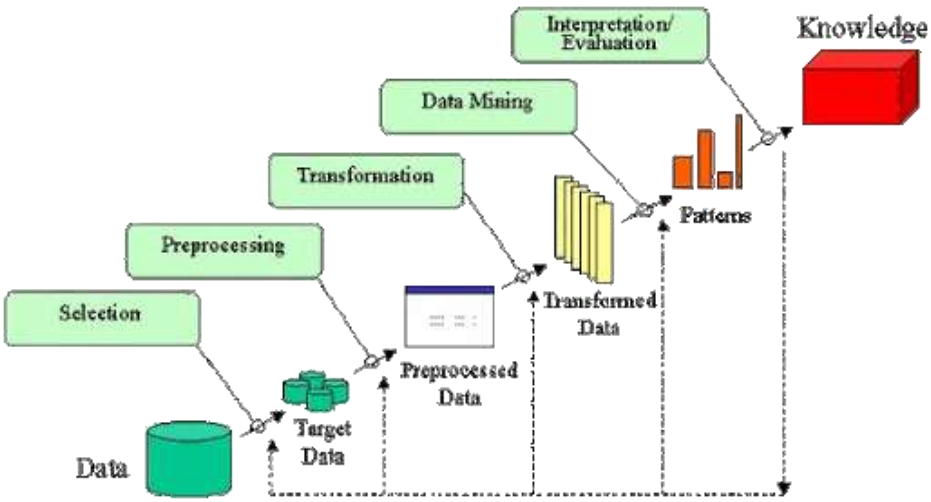

Figure 2. KDD

IOT

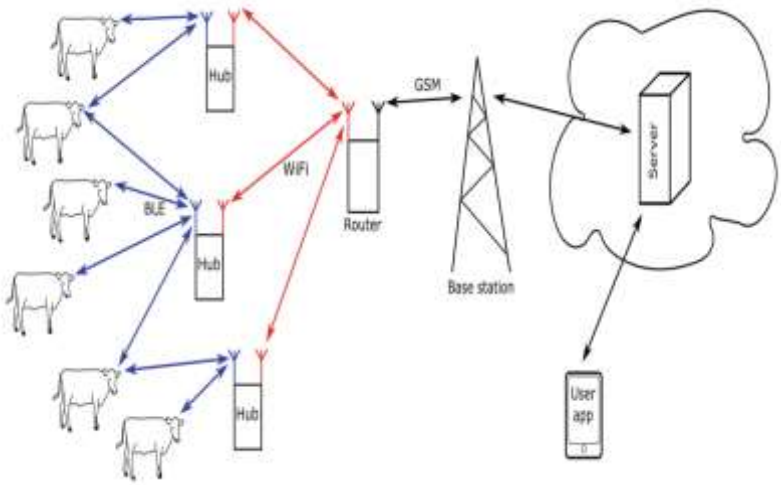

Figure 3. IOT

The Internet of Things (IoT) is more than a buzzword; it is a mechanism that influences how technology communicates with humans and how humans behave. The meaning of IoT is to link networks and computers that have not been connected until now. In other words, IoT brings artificial wisdom to otherwise stupid things. A homeowner can lock or unlock their front door from a phone with an IoT smart home protection device, change the temperature of their home or run a security camera to see who knocks at their door.

The Internet of Things (IoT), which is now a mature and powerful technology, seems to be one of the alternatives to low performance and competitiveness in agriculture and livestock. Farming field surveillance, greenhouse monitoring, agricultural drones, cattle monitoring, cultivation warehouse monitoring, and soil monitoring are the following fields of IoT use in agriculture. The goal of this paper is to implement a new IoT-based animal tracking device dedicated to the automatic evaluation of the health of dairy cows in a traditional loose-housing cow shed.

IV.MATERIALS AND METHODOLOGY

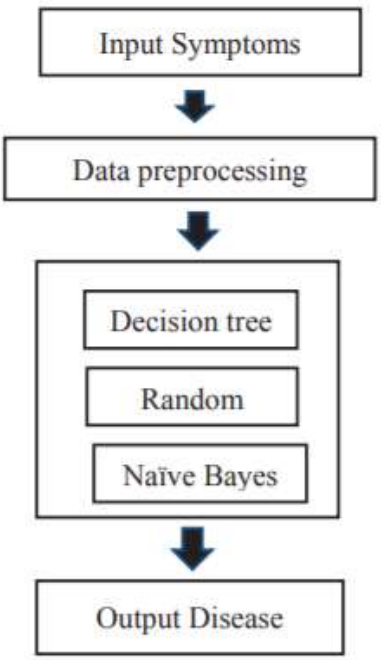

Figure 4. Flow Diagram

i) Input -Symptoms

When creating the model, we believed that the customer had a good awareness of the symptoms he was feeling. The developed prediction recognizes 95 symptoms in which the consumer will treat the symptoms as an

input.

ii) Preprocessing of data

Technique of data mining that transforms the raw data from dataset or the data is encoded into a form that the algorithm can effectively translate is called preprocessing of data. The methods of data preprocessing which are used in the discussed work are:

- Cleaning of Data: Data is cleaned by procedures like filling in missed values, data is cleaned, therefore fixing the discrepancies in the data.
- Reduction of Data: When working with a large archive, the research becomes difficult. We thus delete certain independent variables (symptoms) that may have little to no effect on the goal variable (disease). 95 out of one hundred and thirty two signs which are strictly correlated to diseases are picked in this work.

iii) Selected Models

The model is taught to use three algorithms to forecast diseases:

- Random forest Classifier
- Decision Tree Classifier
- Naïve Bayes Classifier

iv) Output(diseases)

A rule set is established when the model is trained on the dataset using the described algorithms, and once the user provides the symptoms as an input to the model, those symptoms are interpreted according to the developed rule set, thus classifying and predicting the most probable illness.

METHODOLOGY

The scheme of predicting disease is applied using three algorithms for data mining, i.e. Classifier of decision tree, Random forest classifier and classifier of Naïve Bayes. Below is the definition and features of the algorithms.

A. Decision Tree Classifier

The classification models constructed by the decision tree are identical to the tree structure. The dataset is broken down into small and small subsets by studying the sequence of explicit rules(if-then) on function values (In our situation, symptoms), where the target value(in our case disease) is predicted as a result. The decision nodes and leaf nodes are made up of a decision chain.

- Decision node: It has two divisions or more. All of the signs are known as judgment nodes in our presented study.
- Leaf node: It reflects the grouping, that is, the decision of any branch. The leaf nodes apply to diseases here.

$$Gain(S, A) = Entropy(S) - \sum_{vcValues(A)} \frac{|S_v|}{|S|}.Entropy(S_v)$$

A is an attribute, where S is a set of instances, Sv is subset of S with Such that A = v, and Values (A) will be the set of all A's potential values.

B. Random forest Classifier:

Random forest is an easy-to-use, scalable ML algorithm which without the need of any hyper-tuning produces exceptional results most of the time. The major weakness of decision tree algorithms is overfitting, as stated in the Decision Tree. It seems as if the information was memorized by the tree.

This issue is avoided by Random Forest: It is a variant of ensemble learning. It applies several times to using several algorithms or similar algorithms. Random Forest is a Decision Tree squad. And the larger the number of Random Forest decision trees, the greater the generalization.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (fi - yi)^2$$

Where N is the number of data points,
fi is the value returned by the model and
yi is the actual value for data point i.

C. Naïve Bayes Classifier:

The basic principle of Naïve Bayes is that each function renders an:

- Independent
- Equal

Contribution to the performance. The value is that, even on a huge dataset, it operates easily because it takes less processing resources.

Bayes theorem

The algorithm for Naïve Byes is based on the Bayes theorem given by:

$$P(s/h) = \frac{P(h/s)P(s)}{P(h)}$$

Where

$P(s/h)$= Posterior probability
$P(h/s)$=Likelihood
$P(s)$ = Class prior probability
$P(h)$= Predictor Prior probability

In the mentioned equation, class denotes 's' and features denotes 'h' . The denominator in "P(h)" consists of the single term which is a function of "data(features)"-it is not a function of the class with which we are actually concerned. Thus, with all the classes, it would be identical. Traditionally, we disregard this denominator in the naive Bayes classification since it does not influence the classifier's outcome in order to make the prediction.:

$$P(s/h) \propto P(h/s)P(s)$$

V. CONCLUSION

There are some cattle diseases which cause the reduction in yield of milk. This research work is made to predict those diseases early using some of the supervised learning models.

The missing data and noisy data present in the dataset can alter the performance of the model. Therefore, we should use statistical approaches to clean the dataset. The performance of these models is calculated using accuracy and confusion matrix.

This research shows that we can use data mining methods efficiently to forecast/predict the cattle diseases. The result of this model can be used by the veterinary doctors to make accurate predictions of the disease.

According to our research, the naive bayes algorithm is more accurate than the decision tree and random forest.

The performance of algorithms may vary in studies conducted on different data sets. The most important reason is the fact that it depends on the data source, the preprocessing on data and the selection of algorithm parameters. In studies, the researchers may use successfully tested algorithms and they may also use any algorithms which are proper for the dataset. These algorithms may produce different performances depending on the dataset.

There are many areas and trends in the veterinary field where data mining experts can work. Technologies for automation have the ability to allow scientific information and to advance it. But in the veterinary field, there are limited studies. For example; Early detection / automate the detection of compromised health and welfare in animals using sensor studies, automatically diagnosing the health status of animals used by images studies, animal detection by video camera studies and automatically animal identification studies.

REFERENCES

[1] An IoT Solution for Cattle Health Monitoring :-*Akhila Suresh, Sarath T V.
[2] IOT Based Cattle Health Monitoring System : Vaishnavi Shinde,Prof. Sweta Jha, Amruta Taral
[3] IoT-Based Cow Health Monitoring System : Olgierd Unold1(B) , Maciej Nikodem1 , Marek Piasecki1 , Kamil Szyc1 , Henryk Maciejewski1 , Marek Bawiec1 , Pawel Dobrowolski2 , and Michal Zdunek3.

[4] Integrated Animal Health Care Using Iot : Emi Mwaka Katemboh, Raed Abdulla, Vikneswary Jayapal, Sathish Kumar Selvaperumal, Dhakshyani Ratnadurai.
[5] Development of IoT Based Smart Animal Health Monitoring System using Raspberry Pi :Seema Kumaria , Dr. Sumit Kumar Yadavb
[6] Disease Prediction using Machine Learning Algorithms : Sneha Grampurohit, Chetan Sagarnal
[7] Analysis of Data Mining Tools for Disease Prediction: Kauser Ahmed P
[8] Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction : Jyoti Soni , Ujma Ansari
[9] Disease Prediction in Data Mining Techniques: 1K. Aparna, 2Dr. N. Chandra Sekhar Reddy, 3I. Surya Prabha, 4Dr. K. Venkata Srinivas
[10] Improving Health-Care Systems by Disease Prediction: Chinmayi Chitnis, Roger Lee
[11] Disease Prediction in Data Mining Techniques : S.Vijiyarani, S.Sudha.
[12] Machine Learning Methods for Disease Prediction with Claims Data : Tanner Christensen∗, Abraham Frandsen†, Seth Glazier∗, Jeffrey Humpherys∗‡, David Kartchner∗
[13] Disease Prediction by Machine Learning over Big Data from Healthcare Communities : Min Chen, Yixue Hao, Kai Hwang, Fellow, IEEE, Lu Wang, and Lin Wang* .
[14] IoT-Based Cow Health Monitoring System :Olgierd Unold1(B) , Maciej Nikodem1 , Marek Piasecki1 , Kamil Szyc1 , Henryk Maciejewski1 , Marek Bawiec1 , Pawel Dobrowolski2 , and Michal Zdunek3.
[15] An IoT Solution for Cattle Health Monitoring : Akhila Suresh, Sarath T V.
[16] Prediction of postpartum diseases of dairy cattle using machine learning : A.M. Hidalgo1 , F. Zouari1 , H. Knijn1 & S. van der Beek1.
[17] Application of Artificial Intelligence for livestock disease prediction : K. P. Suresh1, Dhemadri2, Rashmi Kurli3, R. Dheeraj4 and Parimal Roy5.
[18] Babesia and Babesiosis in Livestock of Karnataka State, India- An Overview : K. Muraleedharan
[19] Common Cattle Diseases: Symptoms, Treatment And Prevention : Dr Ashok Singh (Professor & Head)
[20] Influence of season, age and breed on prevalence of haemoprotozoan diseases in cattle of Tamil Nadu, India : R. Velusamy, N. Rani, G. Ponnudurai, T. J. Harikrishnan, T. Anna, K. Arunachalam, K. Senthilvel and P. Anbarasi.
[21] Opportunistic Wireless Networking for Smart Dairy Farming : Chamil Kulatunga, Laurence Shalloo, William Donnelly, Eric Robson, and Stepan Ivanov.
[22] Classification of the cattle's behaviors by using accelerometer data with simple behavioral techniques : Sanya Kuankid, Tanadon Rattanawong and Apinan Aurasopon.
[23] THEILERIOSIS IN CATTLE: PREVALENCE AND SEASONAL INCIDENCE IN JALANDHAR DISTRICT OF PUNJAB (INDIA) : Vipan Kumar1, Parvinder Kaur2, V.M.Wadhawan3, Heigo Pal4 Hanish Sharma