International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Improved Production Mechanism for Dairy Plant Management Using Aggregated Rule Mining

[1]Ms. V. Manochitra, [2]Dr. A. Shaik Abdul Khadir

Submitted: 05/02/2024 **Revised**: 13/03/2024 **Accepted**: 19/03/2024

Abstract: The proposed method for supervising and proper management of milk processing processes to reduce utilization of water and energy. The primary process of this model is to measure and improve the pitfalls of the automated dairy plant that process the cow milk day-to-day activity values is monitored and processed using this proposed Improved Aggregated Rule Mining technique. The existing system updated by integrating proposed monitor system to the improved Automated Rule Mining system already deploy in the production process. So the system able to calculate the energy consumptions not directly computed from the available sensors. The efficiency of proposed system is high compared to existing system results. The records are used to improvising a monitoring system to reducing the water/energy consumption. This paper results shown the improved considerable water and energy in several milk processing processes by the introduction of new processing and management technologies by applying this proposed technique. The measurement of proposed approach was based on the comparison with the existing benchmark readings such as water consumption, milk wastage and human efforts.

Keywords*: Aggregated rule mining, Dairy farm, Data Mining, Farm Plant.*

Introduction

Currently, Data mining refers to the process of data analysis carried out from different contexts and then collating it into productive information[1]. Constant development in processing power of computers, disk storage, and statistical tools have been instrumental in the huge rise in the accuracy achieved with analysis while reducing the cost at the same time. Different data mining approaches exists which can be utilized in an application for the generation of knowledge based data[2,3,5]. In this research work, a data mining approach based on data aggregated rule mining has been suggested for the disease detection depending on the symptoms for dairy cow health data gathered with the help of WSN.

In this technical work, Section I studies the importance of the classification of dairy cow diseases. In Section II, the outline on the conventional techniques approaches used for dairy cow diseases classification has been discussed. In section III, the proposed research design model is explained. Section IV studies the simulation results. Section V provides the conclusion and work aimed for the future are discussed.

2. Related Works

Wagneret al [2020][6]studied about the application of Machine Learning (ML) in anomalous behaviour detection using constant monitoring. 14 cows (Bos taurus) are affected with Sub-Acute Ruminal Acidosis (SARA), which is a disease that can trigger behavioural changes. One more set of 14 control cows remained unaffected by SARA. A ruminal bolus is used for pH monitoring and if a cow is affected with SARA, it can be detected. A positioning system was used to understand the behavior of an animal in accordance with its location with respect to particular entities in the barn (feeder, resting area, and alleys). Variety of ML algorithms such as K Nearest Neighbours for Regression (KNNR); Decision Tree for Regression (DTR); MultiLayer Perceptron (MLP); Long Short-Term Memory (LSTM); and an algorithm in which the behavior is supposed to be the same right on consecutive days have been tested. At first, ML models have been developed to estimate the behavior on a particular day in the past 24 hours, taking all cows into account. Next, the error between the inferred and estimated values for a particular cow has been calculated. At the end, the error is compared with a threshold selected for optimizing the difference between normal and anomalous values. The performance of KNNR was found the best, with it able to detect 83% of SARA cases (true-positives), however, its false-positives was 66%, which restricts its usage practically. To conclude, ML can be efficient in anomalies detection with respect to behaviour. More improvements could likely be achieved with the usage of ML on massive datasets at animal level instead of group level.

[1]*Research Scholar, Department of Computer Science, Kadhir Mohideen College, Adirampatinam [Bharathidasan University, Tiruchirappalli] manokavishna@gmail.com*
[2]*Head & Associate Professor, Department of Computer Science, Kadhir Mohideen College, Adirampatinam [Bharathidasan University, Tiruchirappalli] hiqmath4u@gmail.com*

3. Proposed methodology

With regard to animal health monitoring employing WSN, there is a huge volume data that gets consistently gathered and stored. A data aggregated rule mining approach for this system is proposed. Finding intriguing correlations among large volume of data can be useful in rendering decisions, detect new diseases and make policies to monitore dairy cow health.

3.1.WSN based dairy how health monitoring using data aggregated rule mining method

In this work, a wellbeing checking and detailing framework it is proposed to utilize WSN rule. With this system, it is pointed toward observing the wellbeing and natural states of creatures found in the rustic locales in the State of Gujarat. Here, an appropriated information capacity model used for WSN based dairy cow wellbeing checking activity has been presented. This system is characterized into two levels alluded as the nearby level and a focal level. The essential goal of having neighborhood stockpiling is to get a quick answer for any inquiry that the clients raise. A higher level with unified information is useful in taking more time for long range, arranging and strategy making for the dairy cow wellbeing observing action. Whenever the information is within reach, it necessities to impeccably broke down.

In dairy cow, every disease depicts its own symptoms. The bunch of different symptoms characterizes the disease/s. It is likely that one of the symptoms may be a part of multiple diseases. During the time of sickness, in order to find the disease, the set of symptoms must be identified. Let this set be called “Symptom Set”. For the classification of disease to be contagious or non contagious, it is important that their identification is done first.

The algorithm introduced in this research work includes two steps, which are generation of rules and usage of a classifier constructor. In the first step, the training data set is scanned to find the frequent independent elements, and next the elements created are recursively combined to generate elements having multiple attributes. Later, the rules are generated, ranked and stored. In the second step, the rules are utilized for forming a classifier by taking their efficiency with respect to the training data set into consideration. The algorithm is described clearly as follows[8].

Algorithm: 1.Improved Aggregated rule mining Technique

Input: Training data (X), Min Supp and Min Conf thresholds
Output: Ax set of rules
Scan X for the set Ax of frequent single items
Do
For each pair of disjoint items i_1,i_2 in Ax
If $< i_1 \cup i_2 >$ passes the MinSupp
threshold
Ax⇠ $s \cup < i_1 \cup i_2 >$
Until no items which satisfy MinSupp are found
For each item i in Ax
Generate all rules $i \longrightarrow C$ which satisfy the Min Conf threshold
Sort all rules generated
Delete all rules $i' \longrightarrow c'$ from Ax where
there is some rule $i \longrightarrow c$ of a higher rank and
$i \subseteq i'$.

For the productivity improvement as for the finding of successive things and age of rules, a methodology, which depends on a crossing point strategy, is utilized. This strategy plays out the checking the preparation informational index one chance to count the rates of individual components, from which the components not entirely set in stone to pass the MinSupp limit. The components are put away alongside their positions (rowIds) inside the exhibits. Afterward, through the convergence of the rowIds of the successive things found till now, the other regular components including numerous characteristics can be effectively acquired. Additionally, rowIds is utilized for incessant individual components to get the help and certainty values for decides that include various elements [10].

Take the items $<(A_1,X_1)>$ and $<(A_2,y_1)$, the following two sets indicate the rowIds of their incidence, {l, 2, 3, 4, 8} and {1, 3, 5, 6, 10}. The support of a new diseases can be determined, like $< (A_1,x_1),(A_2,y_1)>$ by carrying out an intersection operation on the rowId sets for items $< (A_1,x_1),(A_2,y_1)>$ .The resultant set {I, 3} will specify the tuples in which both the elements have been seen jointly in the training data[11].

In the event that the help related with the new thing < (A1,x1),(A2,y1)> for example 2/10, overcomes the

MinSupp limit, then, at that point, it turns into a possibility to be available as a standard in a standard. Components passing the MinSupp limit are inferred in a recursive way from the components with less number of highlights, starting with the incessant individual component acquired in one ignore the preparation informational index. After a component has been viewed as a regular thing, the MCAR calculation gets every one of the guidelines having that component to be the condition passing the MinConf. It should be seen that each time a regular component is got, MCAR calculation takes exclusively the standard having the highest certainty [12].

In the scenario of an element having two rules with similar confidence, the rules will be randomly selected. As the scanning of the training data set is done just one time to find and create the rules, this technique is extremely efficient during runtime and storage since it does not depend on the conventional technique, where data needs several scan counts. But, in scenarios of a huge number of candidate elements stored in the main memory, the probable number of intersections necessary for the generation of frequent elements may be very high. This makes for one setback in the proposed algorithm, whose resource consumption for storage etc may rise.

3.2 Regulation Order

Regulation Order has a predominant part to play in association mining. Thee rules are ranked primarily in terms of the confidence level of the rules. If the confidences or supports of different rules are similar, one of the rules may be randomly selected, due to which the accuracy may be reduced. To deal with this problem, this proposed work always searches for the best of the rules that can be used for the classification system finally. The top rules are not just the ones having bigger confidence values but also are represented highly in the training data set.

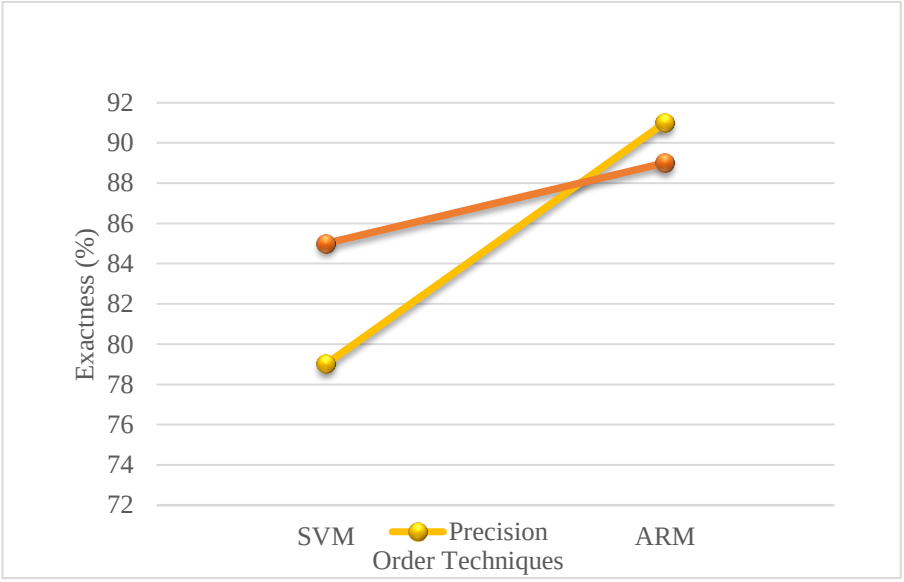
3.3 Rules Evaluation and dairy cows infections Classification

A standard is significant given that its inclusion is in some measure for one preparation test. When the standard age and positioning is done, an assessment step approves each standard individually against the preparation informational index so those which isn't effective in arranging essentially a solitary example are taken out. At each progression, every one of the lines that are appropriately gathered by the current rule will get eliminated from the preparation informational collection. Every single time, a standard isn't effective in characterizing any of the lines of the information, it will get erased from the guidelines set since a standard that is profoundly positioned has arranged its examples right.

In this procedure, it is assured that this classifier contains just the high confidence rules. During classification, assume R refer to the bunch of generated rules and D indicates the test data. The fundamental concept of the proposed technique is to select a set that represents high confidence and common rules present in R to provide coverage for D. During the classification of a test object, the first rule present in the group of ranked rules matching its criterion performs the classification

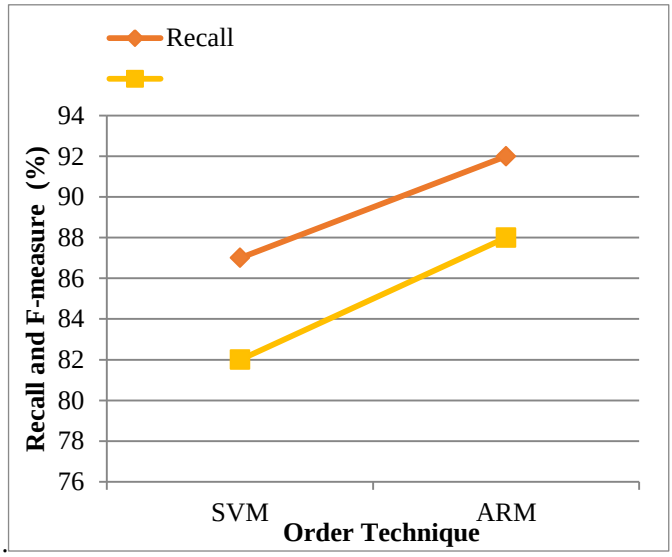
4. Results and Discussion

In this section, the results of the simulation tests carried out with the proposed model are studied. The proposed model is simulated in the java based application software. Here, the comparison between the proposed improved aggregate rule mining model and the available SVM is done in terms of precision, accuracy, recall and f measure metrics for dairy cow dataset.



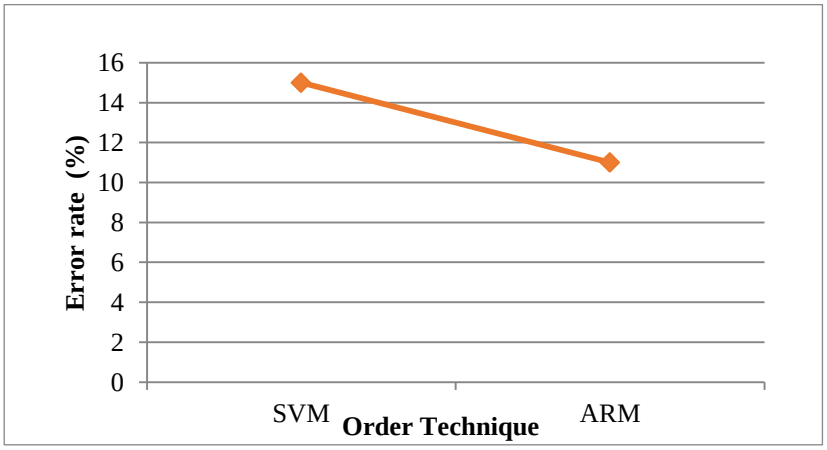
Graph:1. Exactness and accuracy results versus order techniques

Above graph shows the chart for the exhibition examination between the current classifier SVM and the proposed Improved aggregated Rule Mining approach with regards to exactness and accuracy measurements. In the chart, different strategies are brought the X-pivot and the Accuracy and accuracy values are brought the Y-hub. It is deduced from the outcome, it is guaranteed that the proposed model yields an expansion in the Accuracy which is 89% though the current SVM approach gives simply 85% correspondingly. Essentially, as for accuracy, this proposed Improved aggregated Rule Mining model yields further developed accuracy consequence of 91%.



Graph2: .Recall and F-measure results vs. Order Technique

The graph above illustrates the performance comparison between the available classifier SVM and the proposed Improved aggregated Rule Mining approach in terms of recall and F-measure metrics . In the graph above, various techniques are plotted along the X-axis and the recall and f measure values are taken along the Y-axis. From the results, it can be inferred that the proposed Improved aggregated Rule Mining framework achieves an increased recall value of 92% whereas the available SVM approach provides just 87% correspondingly. Similarly, it is shown from the f measure results that this proposed Improved aggregated Rule Mining model achieves excellent f measure values of 88% and the available SVM yields just 82% .



Graph 3: Error rate results vs. Order Technique

In Figure:4 above, the results for the Performance comparison between the available SVM and the proposed Improved aggregated Rule Mining approach in terms of error rate is illustrated. As per the above graph, the various techniques are taken along the X-axis and the error rate values are taken along the Y-axis. From the results, it can be noticed that the proposed Improved aggregated Rule Mining framework yields reduced error rate that is 11% whereas the existing SVM approach provides 15% correspondingly.

5. Conclusion

The most recent few years has seen the fast expansion in ranch computerization starting with the traditional wellbeing observing framework in which human work who are in close contact with the dairy cow are accused of the judgment of their wellbeing to the present

industrialized ranch where cutting-edge medical services approaches are being utilized for remote exchange of the information to a medical services place found from a distance. Likewise the checking of the dairy cow proprietor and the dairy cow can be performed effectively and consequently time, cost and work are diminished. In this examination work, an information amassed rule digging approach is proposed for dairy cows wellbeing observing framework established on WSN. In this, an Improved Association rule based mining procedure is proposed which assists with distinguishing the infection based on side effect. The proficiency of the proposed approach is displayed from the consequences of investigations as far as accuracy, review, exactness and f-measure measurements. However, this proposed model has not been checked with huge volume dataset to exhibit the effectiveness of this proposed model and hence it very well may be the turn out planned for what's to come.

References

[1] FoodDrinkEurope, "Data & Trends of the European Food and Drink Industry",2011.

[2] Eurobarometersurvey,"SMEsareimportantforasmoothtransitiontoa greener economy". MEMO/12/218, March2012.

[3] European Commission, "Integrated Pollution Prevention and Control", Reference Document on Best Available Techniques in the Food, Drink and Milk Industries,2006.

[4] DG Environment – European Commission, "Water Scarcity and Droughts,In-DepthAssessment,SecondInterimReport",June2007.

[5] Sevenster, M. and de Jong, F., "A sustainable dairy sector: Global, regional and life cycle facts and figures on greenhouse-gasemissions", CE Delft, 2008

[6] USDairyWaterUse,"Understandingthegeographicalhotspotsfordairy operations with regard to water use impacts",2011.

[7] CIAA,"ManagingEnvironmentalSustainabilityintheEuropeanFood& Drink Industries",2001.

[8] Bilgen, B., Dogan, K., "Multistage production planning in the dairy industry: A mixed-integer programming approach. Industrial & EngineeringChemistryResearch54(46),pp.11709-11719,2015.

[9] Doganis, P., Sarimveis, H., "Optimal scheduling in a yogurt production line based on mixed integer linear programming". Journal of Food Engineering 80 (2), pp. 445-453,2007

[10] Grossmann,I.E.,Hooker,J.,Mendez,C.,Sand,G.,Wassick,J.,"Scope forindustrialapplicationsofproductionscheduling modelsandsolution methods".Computers&ChemicalEngineering62,pp161-193,2014

[11] Hazaras, M. J., Swartz, C. L., Marlin, T. E., "Industrial application of a continuous-time scheduling framework for process analysis and improvement".Industrial&EngineeringChemistry Research53(1),259- 273,2013.

[12] Kopanos,G.M.,Puigjaner,L.,Georgiadis,M.C.,"Resource-constrained production planning in semicontinuous food industries", Computers & chemical engineering 35 (12), 2929-2944,2011

[13] Mendez, C. A., Cerda, J., Grossmann, I. E., Harjunkoski, I., Fahl, M., "State-of-the-art review of optimization methods for short-term schedulingofbatchprocesses".Computers&ChemicalEngineering30 (6), pp 913-946,2006

[14] Okubo,H.,Miyamoto,T.,Yoshida,S.,Mori,K.,Kitamura,S.,Izui,Y., "Project scheduling under partially renewable resources and resource consumption during setup operations". Computers & Industrial Engineering, 83, pp. 91-99,2015.