

Data mining and decision support systems for efficient dairy production

Sunesh Balhara¹, Rishi Pal Singh² and A. P. Ruhil³

1. ICAR-Central Institute for Research on Buffaloes, Hisar, Haryana, India; 2. Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India; 3. ICAR-National Dairy Research Institute, Karnal, Haryana, India.

Corresponding author: Sunesh Balhara, e-mail: sunesh.balhara@icar.gov.in

Co-authors: RPS: pal_rishi@yahoo.com, APR: apruhil@gmail.com

Received: 22-01-2021, **Accepted:** 02-04-2021, **Published online:** 22-05-2021

doi: www.doi.org/10.14202/vetworld.2021.1258-1262 **How to cite this article:** Balhara S, Singh RP, Ruhil AP (2021) Data mining and decision support systems for efficient dairy production, *Veterinary World*, 14(5): 1258-1262.

Abstract

Gainful livestock farming requires selective breeding of animals with certain heritable desirable traits which gives profitability in terms of farm produce. Modern dairy animals are selected for traits which directly or indirectly contribute to high milk production. The concept of “feed conversion efficiency” in terms of milk production is now vigorously taken up by researchers and farm managers for recognizing and breeding efficient milk-producing animals. The whole concept of economic farming thus requires identification of “elite” animals, meeting above criteria as base population for the farm enterprise. Farmers and animal traders have been selecting best animals based on certain physical characters, which were also accepted by the breeding scientists as phenotypes. Data mining allows uncovering of hidden patterns in the data for better understanding of data relationship for developing suitable models for further improvements. Along with artificial intelligence techniques, data mining has opened new avenues for achieving high resource utilization efficiency and sustainable profitability in livestock production systems. The present review discusses and summarizes various data mining techniques and decision support systems for scientific dairy farming.

Keywords: dairy farming, data mining, decision support systems.

Introduction

The animal rearing has become highly competitive with narrow profit margins and therefore, there is increased focus on research for improving the efficiency of animal production. Proper selection of animals is the first and the most important step in dairy farming. Animal records, which can be utilized in data mining-based animal evaluation, form the basis of selection for breeding programs and/or a new dairy herd. For this, farmers need to be well informed about best practices in the selection of elite animals. It is not uncommon for the animal breeders to ask: Which is the best dairy animal? When an owner offers an animal for sale, he must be putting on animals that are being culled from the herd for undesirable traits. Therefore, the purchaser has to be well aware of the characters of a good animal.

For any dairy setup, most common criteria for the selection of dairy animals for breeding, production and sale/purchase include –

1. Milk production potential – High milk yield is preferred because of higher economic returns it gives
2. Milk composition – In India, fat content is a key factor in determining price of milk. However,

in western countries, fat, protein, and solid not fat are considered. Buffaloes with high milk fat and protein content are preferred for Mozzarella cheese production

3. Fertility – Animals with a pedigree history of regular breeding are desired
4. Sex of the offspring – Female offspring is preferred for dairying while male is preferred for raising bulls for natural service or semen production
5. Performance of the breed in the local environment – Better climate resilient breeds are preferred
6. Resistance against prevalent diseases – Disease resistant animals incur lesser expenses on health-care and hence preferred over vulnerable animals
7. Age of the animal – Age of the animal determines the longevity of the animal in farm and hence its production life. Therefore, younger animals are preferred
8. Feed conversion efficiency – It is an indicator of how efficiently the animal utilizes feed offered for conversion into animal produce (for lactating animals, it is milk). Thus, animals with higher efficiency will produce more with same amount of feed offered and hence selected
9. True to breed characters – These are indicative of good production potential and high resale value
10. Disposal of spent/unproductive animals – Cow slaughtering is not allowed in India but buffalo slaughtering is permissible, so spent/unproductive buffaloes realize higher price and therefore preferred by many farmers for dairy.

A good farm animal is not universal and much depends on what you want and where you are.

Copyright: Balhara, *et al.* Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Researchers have used data mining tools on recorded data for the development of intelligent decision support systems. The present review provides an insight into the progress made with regard to finding smart computer-based solutions to various problems related to commercial animal farming.

Quantifying Traits through Selection Indices

As elaborately above, for success of any dairying program, identification of the best animals is a prerequisite. Conventionally, farmers have developed certain assumptions with respect to good dairy characters and this indigenous technical knowledge has been passed on from generation to generation. Animal breeders and dairy managers have added a lot of information on scientific assessment of the animal for their selection. It has now been widely accepted that there is high level of correlation between physical characters (“phenotypes”) and the production potential of the animal. Farmers, animal lovers, researchers, and breeders have been following “Score Card” methods for various type of animals’ selection during judging in animal championships. Scientifically, animals are evaluated in terms of heritability of characters and genetic gain in subsequent generations [1]. Globally, animal breeding programs have identified and quantified heritability of different traits for improvements in production in subsequent generations. These traits are controlled by certain genes, which allow calculations of genetic correlation. These heritability scores have been used for classification of traits in Holstein and Jersey cows in USA [2]. Data collected on such traits along with genetic parameters have paved way for the development of indices, which can be applied at dairy farms for breed improvement programs, for example, “Net Merit Dollars” index used by the United States Department of Agriculture ranks dairy cows in economic terms. Similarly, the selection indices Jersey Performance Index for Jersey and Total Performance Index for Holstein cattle were developed by the American Jersey Cattle Association and Holstein Association, USA, respectively. These breed-specific selection tools are updated periodically to include latest scientific knowledge and economic aspects of dairying.

Dairy Production Expert Systems Based on Artificial Neural Network (ANN) and Fuzzy Logic

Statistical analysis of the historical animal production data and subsequently its data mining applications gave impetus to for developing smart dairying solutions including automation of operations. ANN and fuzzy logic are most important arithmetic tools that have been widely used in developing artificial intelligence (AI) for decision making machines. Introduced by Lotfi Asker Zadeh in 1965, fuzzy logic (or fuzzy rule) models are based on the principal of logic, which is similar to the human

reasoning ability [3]. The earliest efforts for developing AI focused around comprehensive breeding plan for productivity improvement in livestock population. Animal production systems are greatly affected by diseases, and this necessitated development of expert systems for disease diagnosis and surveillance. Animal nutrition management is another area where expert systems are in great demand as this particular activity accounts for the largest portion of recurring costs in livestock farming.

Many studies have indicated higher accuracy of ANN and fuzzy system models for calculating breeding estimates: (1) First lactation 305-day milk yield [4-6], (2) first lactation 305 days or less milk yield (FL305DMY) from early body weights [7], (3) multiple traits – AFC, FL305DMY, first lactation length, first service period, first dry period, etc. [8], and (4) estimated breeding values [9]. Conventionally, the best linear unbiased prediction method has been used for estimating milk and fat yields in dairy cows. A multilayer program has been used to build ANN and neuro-fuzzy systems to predict breeding values in cows [9]. The model, although not user friendly, had higher degree of accuracy over conventional breeding evaluation methods.

Automation in milking at large-sized dairy farms facilitated accumulation of large volumes of accurate data and development of expert systems for its earliest detection of mastitis. Mastitis, affecting dairy production globally, is one of the most extensively studied dairy animal diseases. Early detection of mastitis helps in checking its further spread and minimizing chances of permanent udder losses. A fuzzy logic model for classification and control of mastitis was developed using a data set of 403,537 milking’s from 478 cows [10]. Mastitis was determined according to somatic cell counts (SCC) standards. Mastitis alerts were generated by a fuzzy logic model using MATLAB software with electrical conductivity, milk production rate and milk flow rate as input data. After evaluation on specific parameters, the model was proposed as a useful tool for the development of an intelligent mastitis detection method. In another study, ANN and adaptive neuro-fuzzy inference system models were developed for raising mastitis alerts (subclinical stage) on the basis of SCC in Holstein cows [11]. A cost-effective neural network (NN) model using combination of variables – pH, electrical conductivity, temperature (udder, milk and skin), milk somatic cells, milk yield, and dielectric constant, was developed to classify healthy and mastitis Murrah buffaloes [12]. A fuzzy logic based automated monitoring system for sorting mastitis and estrus alerts in cows [13]. The output was in the form of “true” or “false,” for each alert. The number of false-positive alerts did not change the level of detected cases of mastitis and estrus. Lameness is another gray area in dairy production, which requires immediate animal replacements in the herd. A very basic fuzzy model

to predict lameness was developed for cows using correlation between claw length and dietary intake in cows [14]. Since the model was very basic, it was not of much use. More sophisticated, inclusive and accurate lameness detection systems have now come up which can alert up to three days in advance of manual observation [15].

Data science is of enormous importance in poultry production. A neuro-fuzzy algorithm – combination of ANN and fuzzy logic, for poultry feed formulation is a useful tool wisely used in African poultry industry [16]. The algorithm in this model was trained with African feed ingredients composition on MATLAB 2009 platform. Outputs from the system were compared with some available standards and the data analyzed on NCSS 2000. The system produced high accuracy of results and fulfills the objectives of reducing feed cost and supplementing all essential minerals and amino acids in poultry diet – Arginine, Histidine, Isoleucine, Leucine, Lysine, Phenylalanine, Tyrosine, Cystine, Methionine, Threonine, Tryptophan, Valine; Crude protein; Ether Extract; and Crude Fiber.

Decision Support Systems (DSS) in Livestock Production

The decision support systems help a setup (organization, individual, etc.) in decision making processes by the use of computer-based models to identify and solve problems [17]. DSS essentially work on the principal of knowledge acquisition from current or historical data followed by creating appropriate algorithm for logical decision-making for identifying and solving problems in the system. The decision support systems are being recognized as tools for improving resource utilization across businesses and other enterprises. For livestock, mostly “data-driven” or “knowledge-driven” DSS are in use or developed, as these are based on analysis of large amounts of structured data/information [18].

During the last decade of twentieth century, a very basic DSS was developed for pasture-based livestock production [19]. This model was a good effort in simulating biological processes (grass growth, pasture-animal interface, and cow model), population dynamics (stochastic herd model), management practices, and economics for suggesting options on efficient use of resources for optimum farm production.

With accumulation of scientific data over the years, better understanding of animal biology and recent developments in data mining applications, DSS for specific purposes are now in use in livestock production. Most DSS have been developed from simulation models using multiple computational methods.

Nevertheless, these systems offer great benefit to the dairy enterprises by assisting decision-making process in the management programs. Table-1 [20-27] lists some of the commercially available DSS which have been well adopted by farmers and frequently updated incorporating scientific breakthroughs.

Table-1: Some common commercially available DSS in dairy farming.

Name of DSS	Application	DM/AI techniques	Reference
GrazPlan suite – GrazFeed, GrazGro and AusFarm	For farmers to make decisions about farm management principally in grazing enterprises	Simulation model	[20]
EUDE	Assist rural business stakeholders in building their own target-relevant decision-making tools in Australia	Generic knowledge model	[21]
VetMet	Surveillance and control of airborne animal diseases	Linearized flow model	[22]
Farmax Dairy Pro	A pastoral grazing model of a dairy farm	Whole-farm system models	[23]
DairyMGT suite - UW-Dairy	A collection of tools for decision support systems in Dairy Farm Management	Matrix-based	[24]
Repro\$	Livestock manure management	ANN	[25]
ValorE	For assisting control of animal diseases – provide information about infected animals, infected places and diseases outbreaks	Apriori algorithm	[26]
ADSDSS	To optimize the management of feed resources and feed supply systems within organic and low input dairy systems in Europe to minimize the risk of feed shortages	Simulation model	[27]

ANN=Artificial neural network, EUDE=End user enabled design environment, ADSOSS=Animal Diseases Spatial Decision Support System

Efforts to develop an expert system for replacing human expert in dairy cattle judging (based on 12 traits – chest width, fore udder attachment, loin, suspensory ligament, angularity, udder depth, rear leg side view, front teat placement, rear leg rear view, rear teat placement, foot angle, and body depth) by linear approach using data mining applications has been tried in Iran [28]. Judging rules were represented as knowledge and fuzzy logic arithmetic was developed to overcome uncertainty during automatic judging. A similar system was developed as a fuzzy logic-based decision support system using reproduction parameters and milk yield records of Holstein Friesians [29] - 305-day milk yield (305 DMY), service period (SP), calving interval (CI), artificial insemination (AI), and dry period (DP) as input parameters. Output parameter of the system is defined as classification decision. Kappa statistics were used to investigate the similarities between expert and decision system and fitting value was obtained as 92.6% ($p < 0.05$).

Computational intelligence techniques have also been used for development of a decision support system in dairy cattle health management [30]. Here, ANNs were used to classify health of cattle into three classes (regular, surveillance, and risk) based on the quality of milk produced. DSS has made way into the concept of precision farming. The Long Short-Term Memory (LSTM) NN based DSS [31] predict forage mass using historical pasture growth data collected directly and association with meteorological data. The results show that LSTM NNs are capable of making a reasonable estimate of the dry mass variation over time. A unique buffalo behavior decision support system, based on threshold value of various acoustic features of vocalization for estrus detection was developed using machine learning techniques J48 classifier along with C4.5 Algorithm to address the problem of silent estrus in buffaloes [32].

Conclusion

Sustainable dairy farming requires comprehensive decisions in husbandry practices for efficient utilization of resources, better health, and productivity. The software programming should enable us to manage a database consisting of every animal on the farm and utilize the database using data mining techniques and decision support systems. Advances in arithmetical systems have given rise to modeling farm situations close to real and therefore provide highly efficient DSS framework to the stakeholders. The tech-savvy dairy farmers are investing in computer-based tools and software products to improve farm economics through monitoring farm products quality and quantity, breeding management, feed – fodder and pasture availability, disease control, and prevention and environmental parameters in their surroundings. AI acquired through data mining techniques, decision support systems and algorithms are bound to make smart decisions for efficient and sustainable dairy production systems.

Authors' Contributions

SB: Did all the literature search and manuscript writing. RPS and APR: Read the manuscript and made corrections. All authors read and approved the final manuscript.

Acknowledgments

The authors are thankful to the Director, ICAR-CIRB, Hisar, for providing necessary funds and facilities for undertaking this study.

Competing Interests

The authors declare that they have no competing interests.

Publisher's Note

Veterinary World remains neutral with regard to jurisdictional claims in published institutional affiliation.

References

- Hill, G.W. (2014), Applications of population genetics to animal breeding, from wright, fisher and lush to genomic prediction. *Genetics*, 196: 1-16.
- Bennet, C. (2009), Using Heritability for Genetic Improvement. College of Agriculture and Life Sciences, Virginia Polytechnic Institute and State University, Publication No. 404-084. Available from: <https://www.goo.gl/ezmwct>. Retrieved on 02-11-2020.
- Zadeh, L.A. (1965) Fuzzy sets. *Inf. Control*, 8(3): 338-353.
- Sharma, A.K., Sharma, R.K. and Kasana, H.S. (2007) Prediction of first lactation 305-day milk yield in Karan Fries dairy cattle using ANN modelling. *Appl. Soft Comput.*, 7(3): 1112-1120.
- Dongre, V.B., Gandhi, R.S., Singh, A. and Ruhil, A.P. (2012) Comparative efficiency of artificial neural networks and multiple linear regression analysis for prediction of first lactation 305-day milk yield in Sahiwal cattle. *Livest. Sci.*, 147(1-3): 192-197.
- Gandhi, R.S., Monalisa, D., Dongre, V.B., Ruhil, A.P., Singh, A. and Sachdeva, G.K. (2012) Prediction of first lactation 305-day milk yield based on monthly test day records using artificial neural networks in Sahiwal cattle. *Indian J. Anim. Sci.*, 65(3): 229-233.
- Manoj, M., Gandhi, R.S., Raja, T.V., Ruhil, A.P., Singh, A. and Gupta, A.K. (2014) Comparison of artificial neural network and multiple linear regression for prediction of first lactation milk yield using early body weights in Sahiwal cattle. *Indian J. Anim. Sci.*, 84(4): 427-430.
- Gandhi, R.S., Raja, T.V., Ruhil, A.P. and Kumar, A. (2009) Prediction of lifetime milk production using artificial neural network in Sahiwal cattle. *Indian J. Anim. Sci.*, 79(10): 1038-1040.
- Shahinfar, S., Mehrabani-Yeganeh, H., Lucas, C., Kalhor, A., Kazemian, M. and Weigel, K.A. (2012) Prediction of breeding values for dairy cattle using artificial neural networks and neuro-fuzzy systems. *Comput. Math. Methods Med.*, 2012: 127130.
- Cavero, D., Tolle, K.H., Buxade, C. and Krieter, J. (2006) Mastitis detection in dairy cows by application of fuzzy logic. *Livest. Sci.*, 105(1-3): 207-213.
- Mammadova, N.M. and Keskin, I. (2015) Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. *Indian J. Anim. Sci.*, 49(5): 671-679.
- Panchal, I., Sawhney, I.K., Sharma, A.K. and Dang, A.K. (2016) Classification of healthy and mastitis Murrah buffaloes by application of neural network models using yield and milk quality parameters. *Comput. Electron. Agric.*, 127: 242-248
- DeMol, R.M. and Wolde, W.E. (2001) Application of fuzzy logic in automated cow status monitoring. *J. Dairy Sci.*, 84(2): 400-410.
- Carvalho, V., Naas, I., Mollo, M. and Massafra, V. (2005) Prediction of the occurrence of lameness in dairy cows using a fuzzy-logic based expert system-Part I. *Agric. Eng. Int.*, 7: 1-12.
- Warner, D., Vasseur, E., Lefebvre, D.M. and Lacroix, R. (2020) A machine learning-based decision aid for lameness in dairy herds using farm-based records. *Comput. Electron. Agric.*, 169(7): 105193.
- Aderounmu, G.A., Omidiora, E.O., Adegoke, B.O. and Taiwo, T.A. (2013) Neuro-fuzzy system for livestock feed formulation (African Poultry). *Int. J. Eng. Sci.*, 2(5): 25-32.
- Burstein, F. and Holsapple, W.C. (2008) Handbook on Decision Support Systems 1: Basic Themes (International Handbooks on Information Systems). 2008th ed. Springer Publishing, Berlin.
- Power, D.J. (2002) Decision Support Systems: Concepts and Resources for Managers. Quorum Books, Connecticut, United States. p12-20.

19. Herrero, M., Fawcett, R.H. and Dent, J.B. (1996) Integrating simulation models to optimise nutrition and management for dairy farms: A methodology. In: Dent, J.B., McGregor M.J. and Sibbald A.R. editors. *Livestock farming Systems: Research, Development, Socio-Economics and the Land Manager*. European Association for Animal Production Publication 79. Wageningen Pers, Wageningen. p322-326.
20. Donnelly, J.R., Moore, A.D. and Freer, M. (1997) GRAZPLAN: Decision support systems for Australian grazing enterprises. I. Overview of the GRAZPLAN project and a description of the MetAccess and LambAlive DSS. *Agric. Syst.*, 54(1): 57-76.
21. Miah, S.J., Kerr, D. and Gammack, J. (2006) A design environment ontology for stakeholder-developed decision support tools in the Australian dairy industry. In: ACIS 2006 Proceedings No. 16. Available from: <https://www.aisel.aisnet.org/acis2006/16>. Retrieved on 10-04-2020.
22. Havskov, J., Alexandersen, S., Astrup, P., Christensen, K.E., Mikkelsen, T., Mortensen, S., Pedersen, T.S. and Thykier-Nielsen, S. (2008) Air pollution modelling and its application. In: XIX-The VetMet Veterinary Decision Support System for Airborne Animal Diseases. p199-207.
23. Bryant, J.R., Ogle, G., Marshall, P.R., Glassey, C.B., Garcia, S.C. and Holmes, C.W. (2010) Description and evaluation of the Farmax dairy pro decision support model. *N. Z. J. Agric. Res.*, 53(1): 13-28.
24. Cabrera, V.E. (2012) DairyMGT: A suite of decision support systems in dairy farm management. In: Jao, C., editor. *Decision Support Systems*. IntechOpen, UK. Available from: https://www.intechopen.com/books/decision-support-systems_2012/dairymgt-a-suite-of-decision-support-systems-in-dairy-farm-management. Retrieved on 26-02-2020.
25. Acutis, M., Alfieri, L., Giussani, A., Provolo, G., Guardo, D.A., Colombini, S., Bertocini, G., Castelnuovo, M., Sali, G., Moschini, M., Sanna, M., Perego, A., Carozzi, M., Chiodinia, M.E. and Fumagalli, M. (2014) ValorE: An integrated and GIS-based decision support system for livestock manure management in the Lombardy region (Northern Italy). *Land Use Policy*, 41: 149-162.
26. Bakr, M., Nasr, M. and Hazman, M. (2015) Cooperative spatial decision support system for controlling animal diseases outbreaks in Egypt. *Int. J. Adv. Netw. Appl.*, 6(6): 2533-2541.
27. Vaillant, J. and Baldinger, L. (2016) Application note: An open-source JavaScript library to simulate dairy cows and young stock, their growth, requirements and diets. *Comput. Electron. Agric.*, 120: 7-9.
28. Alizadeh, H., Hasani-Bafarani, A., Parvin, H., Minaei, B. and Kangavari, M.R. (2008) Dairy cattle judging: An innovative application for fuzzy expert system. In: Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008, October 22-24, 2008, San Francisco, USA. p1-4.
29. Akilli, A., Atil, H., Takma, C. and Ayyilmaz, T. (2015) Fuzzy logic-based decision support system for dairy cattle. *Kafkas Üniv. Vet. Fakültesi Derg.*, 22(1): 13-19.
30. Pimpa, A., Eiamkanitchat, N., Phatsara, C. and Moonmanee, T. (2019) Decision support system for dairy cattle management using computational intelligence technique. In: Proceedings of the 2019 7th International Conference on Computer and Communications Management, ACM, New York, USA. p181-185.
31. Schulte, L., Perez, N., Pinho, L. and Trentin, G. (2019) Decision support system for precision livestock: Machine learning-based prediction module for stocking rate adjustment. In: SBSI'19: Proceedings of the XV Brazilian Symposium on Information Systems. p1-8.
32. Devi, I., Dabas, P., Dudi, K., Lathwal, S., Ruhil, A.P., Singh, Y., Malhotra, R., Baithalu, R. and Sinha, R. (2019) Vocal cues-based decision support system for estrus detection in water buffaloes (*Bubalus bubalis*). *Comput. Electron. Agric.*, 162(1): 183-188.
