Can algorithms help us manage dairy cows?

Marianne Cockburn [1]

Abstract: Digitalisation has reached agricultural production and specifically dairy farming, where a wide range of sensing technologies are now available. From farm management systems over body condition scoring systems to those that detect behavioural changes. All these systems have one aim: to offer decision support to the farmer and aid his management decisions. Currently, however, little is known about the return of investment that these systems offer, or even the effectiveness of their functionality. Only little information is available about the underlying algorithms, despite them presenting the essence of performance. Thus, we can only consider the published literature to get an impression of such systems' outcome. In the current study, we therefore evaluated machine-learning related studies published in the scientific literature between 2015 and 2020. We found that machine-learning algorithms were implemented across all fields of dairy science, but only a minority of them could reliably aid management decisions in practice. In this publication, we aim to give an overview of the achievements of current machine-learning algorithms published in dairy science literature and give an outlook on how they could develop further in the future.

Keywords: dairy, smart farming, machine learning, digitalisation, decision support systems

1 Introduction

Over the past decades, the structure of dairy farming has altered significantly. A trend from small family-run farms to large herd operations can be found across the globe. Dairy farms can only operate in an economically efficient way if the cows are healthy. Monitoring dairy cow health and wellbeing has therefore become critical, but also increasingly challenging. Digitalisation in dairy farming offers new opportunities to monitor animal health. Sensors can nowadays be installed both directly attached to the animal or indirectly in the milking machine or the barn. These sensors continuously collect data, which are stored in a variety of applications, including herd management systems and smartphone applications. Machine-learning algorithms could make additional use of such data in dairy science.

The availability of statistical packages that make it easier to apply machine-learning methods has supported their application in most scientific fields. These algorithms allow to predict data and perform data analysis without investigating a perceived outcome.

Recent studies have used machine-learning algorithms in the dairy sector, yet overall it is unclear which additional value these methods have in comparison to traditional

[1] Agroscope, Competitiveness and System Evaluation, Tänikon 1, 8356 Ettenhausen,
Marianne.cockburn@agroscope.admin.ch, https://orcid.org/0000-0003-3193-0641

approaches. We therefore aim to give the reader an idea of the current ability, potential and risks of such algorithms.

2 Material and methods

The data addressed in the current publication was published and discussed previously [Co20]. The previous publication summarises the current findings and potential of machine-learning methods in dairy science and explores the potential of future research. The current publications shall focus on specific aspects and thus allow an outlook on future application potentials and risks.

2.1 Data collection

Scopus was used to discover machine-learning related studies in the field of dairy science. We used the search string "machine learning dairy", as well as method-related search strings ("Random forest dairy", "Cluster* dairy", "Neural networks dairy", "Deep learning dairy", "K-Nearest neighbor dairy", "Bayesian models dairy", "Support vector dairy", "Decision tree dairy", "Ensemble learning dairy") to find relevant literature. Due to the large number of studies available we decided to screen those in which the search strings appeared in the abstract, title or keywords of the document. Further, only articles published between 2015-2020 were considered. These articles were screened for their relevance to the field.

3 Results and discussion

The results show that machine-learning algorithms have become a common tool in dairy science. The Scopus search resulted in 1821 articles that included the search strings in their abstract title or keywords (Tab.1). Hereby, the search strings for specific method searches resulted in higher numbers of search results than the general term "machine learning", where "cluster*dairy" resulted in the highest number of articles found. The original publication discussed 101 articles [Co20]. In the current publication, we will present a confounded overview of these results.

Search string	Number of search results
Machine learning dairy	109
Random forest dairy	46
Cluster* dairy	1174
Neural networks dairy	112
Deep learning dairy	25
K-Nearest neighbor dairy	10
Bayesian models dairy	213
Support vector dairy	51
Decision tree dairy	68
Ensemble learning dairy	13

Tab. 1: Number of publications found in abstract, title or keywords. Searches last performed in Scopus on 25th of June 2020

3.1 Overview of publications

Studies using machine-learning algorithms covered a vast majority of topics concerning the management of dairy cows. These included feeding-, breeding-, and pasture management, physiology and health, as well as behavioral analysis. The main advantage of machine learning algorithms is their ability to extract information without a preconceived outcome and to predict data. This made it possible to predict milkyields [JVH18], onset of calving [Fa17, Bo17, Fe17], improve grazing management [Hy18] or to predict disease outbreaks through geospatial information [Ro17]. Anomaly detection, particularly for monitoring cow health, is a field with potential. Despite the work performed, the sensitivity and specificity of the algorithms were not satisfying. Glatz-Hoppe et al., [GML19] have been able to improve traditional feeding strategies by evaluating a large dataset (7.3 million recordings) with linear regression models and thus have been able to improve feeding strategies.

3.2 Evaluation of implementation

It became apparent that although models, sensors and IoT devices have become available, their implementation in practice is not satisfactory [OO17]]. A possible reason for this is the inefficiency of sensors, which may be due to sensors addressing only one subfield of management, where none related data to that of different subdomains [OO17]. Hardly any of the reviewed studies integrated their data with that of additional data from sensors or herd management software. Such integrated data sources would offer great potential for the application of machine-learning methods.

Furthermore, farmers are sceptical as to whether they can expect a return of investment if they purchase sensors [OO17]. However, the implementation of Big Data is still developing in the area of smart farming [WVB17, KKP17]. New ideas, such as the "dairy brain", which continuously evaluates data from feeding schedules, herd management systems, and the milking parlour or automatic milking system to improve management decisions, are already being implemented [FCW20].

3.3 Risks of machine learning in dairy science

As promising as these methods are, they do come at a risk. I would like to point out two issues: 1. Machine learning algorithms require large datasets to make accurate predictions. Furthermore, predictions will predict from the data provided, and thus, it can become a problem if these predictions are primarily created from the data of specific countries. To date we can see that many of the studies in the dairy sector are performed on American and Irish data [Co20]. These two countries present the perfect example to explain why this may become a problem. While most farms in America are large herd operations without grazing, the number of cows per farm in Ireland is lower, and most cows are kept in a full grazing system. Thus, algorithms developed on American data may not show satisfying results for Irish farms and vice versa. 2. Many academics were not trained to apply or evaluate the correct application of machine-learning methods as the application of these methods has only become popular in recent years. However, machine-learning methods are increasingly being applied, because better data has become available through the ongoing development of sensors, smartphone applications and software. Additionally, statistical packages that make it easier to apply machine-learning algorithms have been published on open source programming sites. This implies that it could become difficult for supervisors or reviewers to evaluate the quality of the analysis performed and therefore present a risk to the quality of scientific publications.

3.4 Outlook

Large datasets would bring a great step forward to exploit machine-learning algorithms in dairy science. Data need to be well-described, multifactorial, of high-quality, and freely accessible. Better data would allow developing better algorithms. Recently initiated

projects such as “Dairy Brain” and “Smart Dairy Tracer” are working towards demonstrating the potential of such datasets [FCW20, Al20].

4 Conclusion

Machine-learning has become a part of dairy science and is useful for most fields of dairy management. Hereby, machine-learning algorithms can perform predictions or detect anomalies and therefore improve the management of dairy cows. However, to date only few algorithms are ready for application in practice. Large multiparameter datasets, ideally from many different countries would be required to develop better and more useful algorithms.

References

[Co20] Cockburn, M.: Review: Application and Prospective Discussion of Machine Learning for the Management of Dairy Farms. Animals, 10, 1690, 1-22, 2020.

[JVH18] Jensen, D.B.; van der Voort, M.; Hogeveen, H.: Dynamic forecasting of individual cow milk yield in automatic milking systems. Journal of Dairy Science, 101, 10428–10439, 2018.

[Fa17] Fadul, M.; Christopher, B.; Alsaaod, M.; Hasler, J.; Alexander, S.; Adrian, S.; Hirsbrunner, G.: Prediction of calving time in dairy cattle. Animal Reproduction Science, 187, 37–46, 2017.

[Bo17] Borchers, M.R.; Chang, Y.M.; Proudfoot, K.L.; Wadsworth, B.A.; Stone, A.E.; Bewley, J.M.: Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. Journal of Dairy Science, 100, 5664–5674, 2017.

[Fe17] Fenlon, C.; O’Grady, L.; Mee, J.F.; Butler, S.T.; Doherty, M.L.; Dunnion, J.: A comparison of 4 predictive models of calving assistance and difficulty in dairy heifers and cows. Journal of Dairy Science, 100, 9746–9758, 2017.

[Hy18] Hyland, J.J.; Heanue, K.; McKillop, J.; Micha, E.: Factors influencing dairy farmers’ adoption of best management grazing practices. Land Use Policy, 78, 562–571, 2018.

[Ro17] Rossi, G.; de Leo, G.A.; Pongolini, S.; Natalini, S.; Zarenghi, L.; Ricchi, M.; Bolzoni, L.: The potential role of direct and indirect contacts on infection spread in dairy farm networks. PLOS Computational Biology., 13, e1005301, 2017.

[GML19] Glatz-Hoppe, J.; Mohr, E.; Losand, B.: Use of milk recording data for characterization of dairy cow supply situation Second part: Evaluation scheme for the assessment of milk ingredients at farm level. Zuechtungskunde, 91, 449–473, 2019.

[OO17] O’Grady, M.J.; O’Hare, G.M.P.: Modelling the smart farm. Information Processing in Agriculture, 4, 179–187, 2017.

[WVB17] Wolfert, S.; Ge, L.; Verdouw, C.; Bogaardt, M.-J.: Big Data in Smart Farming—A review. Agricultural. Systems. 153, 69–80, 2017.

[KKP17] Kamilaris, A.; Kartakoullis, A.; Prenafeta-Boldú, F.X.: A review on the practice of big data analysis in agriculture. Computer and Electronics in Agriculture, 143, 23–37, 2017.

[FCW20] Ferris, M.C.; Christensen, A.; Wangen, S.R.: Symposium review: Dairy Brain—Informing decisions on dairy farms using data analytics. Journal of Dairy Science, 103, 3874–3881, 2020.

[Al20] Alonso, R.S.; Sittón-Candanedo, I.; García, Ó.; Prieto, J.; Rodríguez-González, S.: An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. Ad Hoc Networks, 98, 102047, 2020.