RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta dissertação será disponibilizado somente a partir de 10/12/2027.

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

PREDIÇÃO DA MASTITE SUBCLÍNICA EM SISTEMA DE ORDENHA AUTOMATIZADO UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

Mariana Massoni Fraga

Médica veterinária

2025

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

PREDIÇÃO DA MASTITE SUBCLÍNICA EM SISTEMA DE ORDENHA AUTOMATIZADO UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

Mariana Massoni Fraga

Orientador: Prof. Dr. Luiz Francisco Zafalon

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Ciências Veterinárias.

2025

F811p

Fraga, Mariana Massoni

Predição da mastite subclínica em sistema de ordenha automatizado utilizando técnicas de aprendizado de máquina / Mariana Massoni Fraga. -- Jaboticabal, 2026

61 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Luiz Francisco Zafalon

1. Bovinos de leite. 2. Mastite. 3. Ordenha. 4. Aprendizado do computador. 5. Microorganismos. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: PREDIÇÃO DA MASTITE SUBCLÍNICA EM SISTEMA DE ORDENHA AUTOMATIZADO UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

AUTORA: MARIANA MASSONI FRAGA
ORIENTADOR: LUIZ FRANCISCO ZAFALON

Aprovada como parte das exigências para obtenção do Título de Mestra em Ciências Veterinárias, área: Saúde Única pela Comissão Examinadora:

Pesquisador Dr. LUIZ FRANCISCO ZAFALON (Participaçao Virtual)
EMBRAPA Pecuária Sudeste (CPPSE) / São Carlos/SP

gov.br Documento assinado digitalmente
LUIZ FRANCISCO ZAFALON
Data: 11/12/2025 08:13:46-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. ANDERSON LUIZ ARA SOUZA (Participaçao Virtual)
Universidade Federal do Paraná (UFPR) / Curitiba/PR

gov.br Documento assinado digitalmente
ANDERSON LUIZ ARA SOUZA
Data: 10/12/2025 23:51:44-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. IVERALDO DOS SANTOS DUTRA (Participaçao Virtual)
Departamento de Producao e Saude Animal / FMVA UNESP Aracatuba

gov.br Documento assinado digitalmente
IVERALDO DOS SANTOS DUTRA
Data: 11/12/2025 07:34:25-0300
Verifique em https://validar.iti.gov.br

Jaboticabal, 10 de dezembro de 2025.

DADOS CURRICULARES DO AUTOR

Mariana Massoni Fraga, natural da cidade de Jaú, SP, nascida em 29 de agosto de 1999, é Médica Veterinária graduada pela Universidade Estadual do Norte do Paraná (2023), recebendo o título de aluna laureada em reconhecimento ao melhor desempenho acadêmico. Durante a graduação, realizou estágios na área de bovinocultura leiteira na Empresa Brasileira de Pesquisa Agropecuária (Embrapa Pecuária Sudeste), em São Carlos, SP, e atuou na clínica médica e cirúrgica de grandes animais no Hospital Veterinário da UENP, em Bandeirantes, PR. Desenvolveu e foi bolsista em dois projetos de iniciação científica: “Avaliação do comportamento, desempenho e estresse oxidativo de cordeiros em enriquecimento ambiental” e “Avaliação do hemograma em touros atletas pré e pós exercício”. Foi monitora da disciplina de Anatomia, Fisiologia e Melhoramento animal. No mesmo ano de formatura ingressou no Mestrado em Ciências Veterinárias na Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista (FCAV/UNESP), Campus de Jaboticabal, SP, concluindo em 2025.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela resiliência para chegar até aqui.

Ao meu orientador, Luiz Francisco Zafalon, pela acolhida desde o primeiro contato, ainda no e-mail em que manifestei interesse em participar do processo seletivo de mestrado. Agradeço pela confiança, incentivo constante e pelas orientações atentas, críticas e construtivas, que foram fundamentais para o meu crescimento acadêmico e profissional.

Aos meus pais, Regina e Edgard, por sempre me proporcionarem as melhores condições para estudar e pelo apoio incondicional em cada etapa da minha vida.

À minha irmã, Giovana, minha melhor amiga e parceira inseparável, cuja presença e incentivo constante foram indispensáveis nesta trajetória.

À minha avó Angelina, inspiração e motivação diária, e ao meu avô Guilherme Onofre (*in memoriam*) que me acompanha em pensamento.

Ao meu namorado, Marcos Antonio, por ser meu braço direito, oferecendo apoio nos momentos de maior desafio.

À UNESP, pelo aprendizado, crescimento pessoal e profissional proporcionados ao longo do mestrado.

Ao professor Anderson Ara, da Universidade Federal do Paraná, e seus orientandos, Débora Pompilio Leite, Alan Vitor Lima Silva e Arthur Hideio Noguti, pela valiosa contribuição e colaboração prestadas, fundamentais para a concretização deste trabalho.

À CAPES pela concessão da bolsa. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ao CNPq pelo financiamento da pesquisa, processo nº 403108/2023-3..

A todos os funcionários da Embrapa Pecuária Sudeste, em especial às pesquisadoras Teresa Cristina Alves e Simone Cristina Méo Niciura, e meus colegas, Larissa Cristina Brassolatti, Maria Laura da Silva Soarde e Mateus Galli da Silva, que foram essenciais para a realização deste trabalho.

SUMÁRIO

CERTIFICADO DA COMISSÃO DE ÉTICA NO USO DE ANIMAIS..... ii
RESUMO.....iii
ABSTRACT..... iv
CAPÍTULO 1 – Considerações gerais.....1
1.1 Introdução.....1
1.2 Revisão de literatura.....3
1.2.1 Pecuária leiteira.....3
1.2.2 Ordenha robótica.....4
1.2.3 Mastite bovina.....6
1.2.4 Aspectos etiológicos da mastite bovina.....7
1.2.5 Métodos diagnósticos da mastite subclínica.....8
1.2.6 Diagnóstico microbiológico da mastite bovina.....10
1.2.7 Principais técnicas de aprendizado de máquina.....12
1.2.8 Aprendizado de máquina aplicado à mastite bovina.....16
REFERÊNCIAS.....17
CAPÍTULO 2 – Prediction of subclinical mastitis in automated milking system using machine learning techniques.....31

CERTIFICADO DA COMISSÃO DE ÉTICA NO USO DE ANIMAIS (CEUA)

Embrapa
Pecuária Sudeste

CERTIFICADO
CEUA PRT Nº 01/2024

Certificamos que o projeto de Pesquisa intitulado: "Ordenha robótica em vacas criadas a pasto: estratégias para o diagnóstico, predição e controle da mastite subclínica por técnicas de aprendizado de máquina", sob responsabilidade do pesquisador científico Dr. Luiz Francisco Zafalon, que envolve a produção, manutenção ou utilização de animais pertencentes ao filo Chordata, subfilo Vertebrata (exceto humanos), para fins de pesquisa científica encontra-se de acordo com os preceitos da lei nº 11.794, de 8 de outubro de 2008, do Decreto nº 6.899, de 15 de julho de 2009 e com as normas editadas pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA) e foi aprovado pela Comissão de Ética no Uso de Animais da Embrapa Pecuária Sudeste.

(We hereby declare that the research project titled "Robotic milking in cows at pasture: strategies for the diagnosis, prediction, and control of subclinical mastitis using machine learning techniques" has been registered under the responsibility of Dr. Luiz Francisco Zafalon involving production, management or utilization of animals from phylum Chordata, subphylum Vertebrata (except humans). The described experimental protocol is in accordance to the Brazilian Federal Law on Animal Experimentation (#11.794, enacted on 8th October 2008), to the Decree 6.899 (enacted on 15th July 2009) and the corresponding rules of National Council for Animal Experimentation Control (CONCEA), and it was approved by the Committee of Animal Experimentation of Embrapa Southeast Livestock.)

São Carlos, 14 de março de 2024.

Dr Alexandre Rossetto Garcia
Presidente da Comissão de Ética no Uso de Animais
Embrapa Pecuária Sudeste

Finalidade	Pesquisa Científica
Vigência da Autorização	05/12/2023 a 31/12/2026
Espécie / Linhagem / Raça	Bovino / Bos taurus / Holandesa e Jersolanda
Número de Animais	70
Peso / Idade	Em fase de lactação
Sexo	Fêmeas
Origem	Embrapa Pecuária Sudeste

Empresa Brasileira de Pesquisa Agropecuária
Ministério da Agricultura, Pecuária e Abastecimento
Rod. Washington Luiz, km 234 Caixa Postal 339 13560-970 São Carlos, SP
Telefone (16) 3311-5600
E-mail: cppse.ceua@embrapa.br

PREDIÇÃO DA MASTITE SUBCLÍNICA EM SISTEMA DE ORDENHA AUTOMATIZADO UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

RESUMO - A detecção precoce da mastite subclínica é essencial para manter a qualidade do leite, evitar a disseminação entre vacas em lactação e prevenir infecções crônicas. Os sistemas de ordenha automatizados possibilitam o monitoramento contínuo do rebanho, registrando variáveis como produção e fluxo de leite, condutividade elétrica, frequência de ordenhas e presença de sangue. Esses dados permitem a geração de alertas automáticos sobre a saúde do úbere, favorecendo intervenções rápidas. Com o avanço do aprendizado de máquina, tornou-se possível integrar essas informações em modelos capazes de identificar padrões complexos e prever a ocorrência de doenças com maior precisão. O objetivo deste estudo foi aplicar técnicas de aprendizado de máquina supervisionado para detectar mastite subclínica, combinando dados produtivos, zootécnicos, ambientais e microbiológicos obtidos em um sistema de ordenha robótica. Foram acompanhadas 103 vacas em lactação, em uma propriedade de São Carlos, SP, durante um ano. Nesse período, foram registradas variáveis referentes à data da coleta, teor de gordura, condutividade elétrica (CE), Índice de Detecção de Mastite (MDi), produção e fluxo de leite, presença de sangue, sujidade do teto, tempo de ordenha, paridade, estágio de lactação, temperatura e umidade relativa do ar. O conjunto de dados final incluiu 955 registros de quartos mamários classificados por posição anatômica e descritos por variáveis explicativas. A mastite subclínica foi classificada de três maneiras distintas: (1) adoção de um valor de triagem para a contagem de células somáticas (CCS) ≥ 200 × 10³ células/mL; (2) combinação de CCS ≥ 200 × 10³ células/mL com resultados microbiológicos positivos; e (3) diferenciação dos quartos mamários em sadios (cultura microbiológica negativa e CCS < 200 × 10³ células/mL) ou acometidos por mastite subclínica (quartos com isolamentos bacterianos e/ou CCS ≥ 200 × 10³ células/mL). Seis algoritmos supervisionados foram avaliados, *Support Vector Machine*, Regressão Logística, *Decision Tree*, *Random Forest*, XGBoost e Rede Neural *Perceptron* Multicamadas, para modelar a relação entre as variáveis e a ocorrência da doença. Entre os modelos testados, o *Random Forest* apresentou o melhor desempenho com a classificação 3, demonstrando maior robustez e precisão. A análise dos valores de *SHapley Additive exPlanations* (SHAP) indicou as variáveis MDi e CE como as mais relevantes na predição, evidenciando o potencial do aprendizado de máquina como ferramenta de apoio à gestão da saúde mamária e à adoção de práticas preventivas na produção leiteira.

Palavras-chave: Diagnóstico Precoce, Qualidade do Leite, Modelos Preditivos, Saúde da Glândula Mamária.

PREDICTION OF SUBCLINICAL MASTITIS IN AUTOMATED MILKING SYSTEM USING MACHINE LEARNING TECHNIQUES

ABSTRACT - Early detection of subclinical mastitis is essential to maintain milk quality, prevent its spread among lactating cows, and avoid chronic infections. Automated milking systems enable continuous herd monitoring by recording variables such as milk yield and flow, electrical conductivity, milking frequency, and the presence of blood. These data allow the generation of automatic alerts regarding udder health, facilitating timely interventions. Advances in machine learning have made it possible to integrate this information into models capable of identifying complex patterns and predicting disease occurrence with higher accuracy. The objective of this study was to apply supervised machine learning techniques to detect subclinical mastitis by combining productive, zootechnical, environmental, and microbiological data obtained from a robotic milking system. Approximately 103 lactating cows were monitored over one year at a farm in São Carlos, SP, Brazil. During this time, data were collected on sampling date, milk fat content, electrical conductivity (EC), Mastitis Detection Index (MDi), milk yield and flow, blood presence, teat cleanliness, milking time, parity, stage of lactation, temperature and relative humidity. The final database comprised 955 quarter-level records classified by anatomical position and described by explanatory variables. Subclinical mastitis was determined in three different ways: (1) adoption of a screening threshold for somatic cell count (SCC) ≥ 200 × 10^3 cells/mL; (2) combination of SCC ≥ 200 × 10^3 cells/mL with positive microbiological results; and (3) grouping mammary quarters as healthy (negative microbiological culture and SCC < 200 × 10^3 cells/mL) or affected by subclinical mastitis (including quarters with positive culture and/or SCC ≥ 200 × 10^3 cells/mL). Six supervised algorithms, Support Vector Machine, Logistic Regression, Decision Tree, Random Forest, XGBoost, and Multilayer Perceptron Neural Network, were evaluated to model the relationship between the variables and disease occurrence. Among the models tested, Random Forest achieved the best performance with classification 3, demonstrating greater robustness and accuracy. SHapley Additive exPlanations (SHAP) values analysis identified MDi and EC as the most relevant variables for prediction, highlighting the potential of machine learning as a tool to support udder health management and the implementation of preventive practices in dairy production.

Keywords: Early Diagnosis, Milk Quality, Predictive Models, Udder Health.

CAPÍTULO 1 – Considerações gerais

1.1 Introdução

A mastite é uma das doenças mais prevalentes em bovinos leiteiros, comprometendo a qualidade do leite, reduzindo a produção e representando uma das principais causas de perdas econômicas e descarte de vacas (Hogeveen et al., 2011; Hertl et al., 2018). Além disso, compromete a saúde geral do rebanho e está fortemente associada ao uso de antibióticos, os quais, embora eficazes no controle das infecções, podem favorecer o desenvolvimento de resistência antimicrobiana quando utilizados de forma excessiva ou inadequada (Vanderhaeghen et al., 2010).

Diferentemente da mastite clínica, que apresenta alterações evidentes no úbere ou no leite, a forma subclínica é geralmente diagnosticada por métodos indiretos, como a contagem de células somáticas (CCS) ou testes como o *California Mastitis Test* (CMT) (Viguier et al., 2009). A mastite subclínica apresenta um desafio particular, pois é muito mais comum que a forma clínica, afetando de 20 a 30% das vacas anualmente (Heringstad et al., 2000; Ebrahimie et al., 2018), mas não apresenta sintomas visíveis, o que dificulta a detecção precoce.

Estratégias para reduzir a dependência de medicamentos devem priorizar a saúde da glândula mamária, uma vez que grande parte do uso de antimicrobianos na bovinocultura leiteira está relacionada ao tratamento de afecções do úbere (Petrovski et al., 2006; Krogh et al., 2020). A detecção precoce da mastite subclínica surge, portanto, como uma alternativa promissora para impedir a progressão da doença para o estado clínico, contribuindo para a melhoria da saúde do rebanho e consequentemente a redução do uso de antimicrobianos (Pakrashi et al., 2023).

Nos últimos anos, a adoção de sistemas automáticos de ordenha (*Automatic Milking System* – AMS) e de procedimentos rotineiros de controle leiteiro tem recebido crescente atenção na pecuária leiteira. Esses sistemas permitem a coleta contínua e detalhada de informações, como volume de produção, condutividade elétrica (CE), frequência e intervalo entre ordenhas. Embora as fazendas com AMS coletem um grande volume de dados, ainda há um potencial expressivo e pouco explorado para a caracterização do rebanho e a otimização do manejo (Bonora et al., 2018).

Esse contexto se insere no fenômeno do "*Big Data*", caracterizado por volumes massivos de dados de grande diversidade, passíveis de análise e utilização para apoiar a tomada de decisões (Wolfert et al., 2017). A melhoria na integração desses dados é essencial para explorar esse vasto conjunto de informações e torná-las mais acessíveis aos produtores (Cockburn, 2020). No entanto, as relações complexas e não lineares presentes em tais conjuntos de dados, frequentemente redundantes, apresentam desafios que os modelos estatísticos tradicionais têm dificuldade em resolver (Nayeri et al., 2019).

O aprendizado de máquina é um ramo da inteligência artificial que emprega algoritmos avançados e técnicas de modelagem de dados que complementam os métodos da estatística tradicional (Rebala et al., 2019). A principal vantagem reside na capacidade de identificar padrões e relações relevantes em grandes volumes de dados, frequentemente complexos e redundantes, provenientes de múltiplas fontes. De modo geral, os métodos de aprendizado de máquina consistem em um processo no qual o modelo aprende a executar determinada tarefa a partir da análise e processamento de dados de treinamento. Ao longo do tempo, o desempenho do modelo é aprimorado continuamente por meio do ajuste de seus algoritmos de aprendizado, aumentando sua precisão na interpretação e análise dos dados.

Mais recentemente, técnicas de aprendizado de máquina têm sido aplicadas à pecuária de precisão (García et al., 2020), gestão e produtividade animal (Wang et al., 2021; Brock et al., 2021), monitoramento do comportamento e bem-estar dos animais (Rao et al., 2020; Carslake et al., 2021). No setor leiteiro, esses métodos vêm sendo utilizados para tarefas como detecção de estro (Higaki et al., 2019), avaliação da severidade do estresse térmico (Becker et al., 2021), monitoramento das interações sociais (Ren et al., 2021), previsão do momento do parto em vacas (Miller et al., 2020) e identificação precoce de vacas em risco de doenças, como mastite ou cetose (Bauer e Jagusiak, 2022; Pakrashi et al., 2023).

Os estudos publicados até o momento têm utilizado, em geral, modelos de classificação para distinguir vacas saudáveis daquelas em risco de mastite subclínica, definida pela elevação da CCS no leite. Os valores-limite mais frequentemente adotados são 200.000 células/mL (Hiitiö et al., 2017; Bobbo et al., 2021) ou 250.000 células/mL de leite (Ebrahimi et al., 2019; Ebrahimie et al., 2021). Entre os modelos mais comumente empregados estão árvores de decisão, *Support Vector Machine*, modelos de agrupamento, redes neurais e modelos lineares generalizados, como a

regressão logística (Mammadova e Keskin, 2015; Sitkowska et al., 2017; Ebrahimi et al., 2019; Bobbo et al., 2021).

O grande volume e a diversidade de dados gerados pelos AMS oferecem subsídios valiosos para o desenvolvimento de modelos preditivos capazes de identificar precocemente alterações associadas à mastite subclínica, contribuindo para estratégias mais eficazes de monitoramento da saúde da glândula mamária. No entanto, apesar da crescente adoção na bovinocultura leiteira, ainda persistem desafios relacionados à compreensão dos algoritmos empregados e à aplicação de técnicas de aprendizado de máquina sobre os dados provenientes dos sensores do robô de ordenha e de outros dispositivos integrados às fazendas (Ozella et al., 2023).

A condução desta dissertação teve como propósito aplicar técnicas de aprendizado de máquina para a detecção prévia de mastite subclínica, integrando indicadores provenientes de um sistema de ordenha automatizado com variáveis produtivas, zootécnicas, ambientais e microbiológicas, a fim de avaliar a eficácia dos modelos preditivos na identificação da doença e compreender a influência relativa de cada variável na predição.

CONCLUSION

The RF model demonstrated robust and consistent performance across different classification categories, positioning it as the most promising option for practical application in dairy herds. The classification that grouped udder quarters as healthy (negative microbiological culture and SCC < 200 × 10^3 cells/mL) or affected by

subclinical mastitis (positive culture and/or SCC ≥ 200 × 10^3 cells/mL) yielded the highest F1-score values, highlighting it as the most efficient approach among the evaluated classifications. Among the variables analyzed, the MDi and milk EC stood out as the main predictors, reinforcing their role in the early detection of subclinical mastitis. These findings demonstrate the potential of machine learning as a tool to support intelligent udder health management, contributing to more preventive and efficient practices in dairy production.

AUTHOR CONTRIBUTIONS

Mariana Massoni Fraga, Maria Laura da Silva Soarde, Larissa Cristina Brassolatti and Mateus Galli da Silva were responsible for data collection and data analysis. Débora Pompilio Leite, Alan Vitor Lima Silva, Arthur Hideio Noguti and Anderson Ara conducted the statistical analyses. Marcos Veiga dos Santos assisted with the microbiological analyses. Teresa Cristina Alves and Simone Cristina Méo Niciura assisted with data organization and data interpretation. Luiz Francisco Zafalon contributed to data interpretation and final manuscript preparation. All authors reviewed and approved the final manuscript.

FUNDING INFORMATION

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Conselho Nacional de Desenvolvimento Científico e Tecnológico.

ACKNOWLEDGEMENTS

The study was financed by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil (Finance Code 001), and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (Grant No. 403108/2023-3).

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The study was conducted following the criteria established by the Ethics Committee on Animal Use of Embrapa Southeastern Livestock under protocol no. 01/2024.

REFERENCES

1. Busanello M, Rossi RS, Cassoli LD, et al. Estimation of prevalence and incidence of subclinical mastitis in a large population of Brazilian dairy herds. *J Dairy Sci* 2017;100(8):6545–53.
2. Steeneveld W, van der Gaag LC, Ouweltjes W, et al. Discriminating between true-positive and false-positive clinical mastitis alerts from automatic milking systems*. J Dairy Sci* 2010;93(6):2559–68.
3. Sun Z, Samarasinghe S, Jago J. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks*. J Dairy Res* 2010;77(2):168–75.
4. Pakrashi A, Ryan C, Guéret C, et al. Early detection of subclinical mastitis in lactating dairy cows using cow-level features. *J Dairy Sci* 2023;106(7):4978–90.
5. Hogeveen H, Ouweltjes W. Sensors and management support in high-technology milking. *J Anim Sci* 2003;81(3):1-10.
6. Ozella L, Brotto Rebuli K, Forte C, et al. A Literature Review of Modeling Approaches Applied to Data Collected in Automatic Milking Systems. *Animals* 2023;13:1916.
7. Rao Y, Jiang M, Wang W, et al. On-farm welfare monitoring system for goats based on Internet of Things and machine learning. *Int J Distrib Sens Netw* 2020;16:155014772094403.
8. Carslake C, Vázquez-Diosdado JA, Kaler J. Machine Learning Algorithms to Classify and Quantify Multiple Behaviours in Dairy Calves Using a Sensor: Moving beyond Classification in Precision Livestock. *Sensors* 2021;21:88.
9. Miller GA, Mitchell M, Barker ZE, et al. Using animal-mounted sensor technology and machine learning to predict time-to-calving in beef and dairy cows. *Animal* 2020;14:1304-1312.
10. Bauer EA, Jagusiak W. The Use of Multilayer Perceptron Artificial Neural Networks to Detect Dairy Cows at Risk of Ketosis. *Animals* 2022;12:332.
11. Mammadova N, Keskin I. Application of the support vector machine to predict subclinical mastitis in dairy cattle. *Sci World J* 2013;2013(1):603897.
12. Hammer JF, Morton JM, Kerrisk KL. Quarter-milking-, quarter-, udder- and lactation-level risk factors and indicators for clinical mastitis during lactation in pasture-fed dairy cows managed in an automatic milking system. *Aust Vet J* 2012;90(5):167–74.
13. Hiitiö H, Vakkamäki J, Simojoki H, et al. Prevalence of subclinical mastitis in Finnish dairy cows: changes during recent decades and impact of cow and herd factors. *Acta Vet Scand* 2017;59:22.

14. Ebrahimie E, Mohammadi-Dehcheshmeh M, Laven R, et al. Rule Discovery in Milk Content towards Mastitis Diagnosis: Dealing with Farm Heterogeneity over Multiple Years through Classification Based on Associations. *Animals* 2021;11:1638.
15. Coatrini-Soares A, Coatrini-Soares J, Popolin Neto M, et al. Microfluidic E-tongue to diagnose bovine mastitis with milk samples using Machine learning with Decision Tree models. *Chem Eng J* 2023;451:138523.
16. Tian H, Zhou X, Wang H, et al. The Prediction of Clinical Mastitis in Dairy Cows Based on Milk Yield, Rumination Time, and Milk Electrical Conductivity Using Machine Learning Algorithms. *Animals* 2024;14:427.
17. Faris D, Gad A, El-Tarabany M, et al. Machine learning algorithms for clinical mastitis prediction in a dairy herd using automated milking system data. *J Adv Vet Res 2024*;14:975–981.
18. Satoła A, Satoła K. Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking, and super-learner ensembles versus single machine learning models. *J Dairy Sci* 2024;107(6):3959–3972.
19. Schreiner DA, Ruegg PL. Relationship between udder and leg hygiene scores and subclinical mastitis. *J Dairy Sci* 2003;86(11):3460–5.
20. Oliver SP, Lewis MJ, Gillespie BE. et al. Microbiological Procedures for the Diagnosis of Bovine Udder Infection and Determination of Milk Quality. 4th ed. Verona: National Mastitis Council, 2004.
21. Koneman EW, Allen SD, Janda WM. et al. Diagnóstico Microbiológico: Texto e atlas colorido. 5th ed. Rio de Janeiro: Medsi, 2001.
22. Burton GRW, Engelkirk, PG. Microbiologia para Ciências da Saúde. 7th ed. Rio de Janeiro: Guanabara Koogan, 2005.
23. Barcelos MM, Martins L, Grenfell RC. et al. Comparision of standard and on-plate extraction protocols for identification of mastitis-causing bacteria by MALDI-TOF MS. *Braz J Microbiol* 2019;50(3):849–857.
24. El-Sayed A, Awad W, Abdou NE. et al. Molecular biological tools applied for identification of mastitis causing pathogens. *Int J Vet Sci Med* 2017;5:89–97.
25. Damm M, Holm C, Blaabjerg M. et al. Differential Somatic Cell Count—A Novel Method for Routine Mastitis Screening in the Frame of Dairy Herd Improvement Testing Programs. *J. Dairy Sci* 2017;100:4926–4940.
26. National Mastitis Council. Guidelines on Normal and Abnormal Raw Milk Based on SCC and Signs of Clinical Mastitis. Madison: NMC; 2001.
27. German Veterinary Society. Leitlinien zur Bekämpfung der Mastitis als Bestandsproblem. Giessen: Verlag der Deutschen Veterinärmedizinischen Gesellschaft; 2002.
28. Bansal BK, Hamann J, Grabowski NT. et al. Variation in the composition of selected milk fraction samples from healthy and mastitic quarters, and its significance for mastitis diagnosis. *J Dairy Res* 2005;72(2):144–52.
29. Gonçalves JL, Lyman RL, Hockett M. et al. Using milk leukocyte differentials for diagnosis of subclinical bovine mastitis. *J Dairy Res* 2017;84(3):309–17.
30. Yu W, Liu T, Valdez R. et al. Application of support vector machine modeling for prediction of common diseases: The case of diabetes and pre-diabetes. *BMC Med Inform Decis Mak* 2010;10(1):16.

31. Miekley B, Traulsen I, Krieter J. Mastitis detection in dairy cows: The application of support vector machines. *J Agric Sci* 2013;151(6):889–97.
32. Kleinbaum DG, Dietz K, Gail M. et al. Logistic regression. 2nd ed. New York: Springer, 2002.
33. Louzada F, Ara A, Fernandes GB. Classification methods applied to credit scoring: Systematic review and overall comparison. *Surv Oper Res Manag Sci* 2016;21:117–34.
34. Breiman L, Friedman JH, Olshen RA. et al. Classification and regression trees. Monterey (CA): Wadsworth & Brooks; 1984.
35. Jena M, Dehuri S. DecisionTree for Classification and Regression: A State-of-the-Art Review. *Informatica* 2020;44(4):541-552.
36. Yap BW, Ong SH, Husain NHM. Using data mining to improve assessment of credit worthiness via credit scoring models. *Expert Syst Appl* 2011;38(10):13274–83.
37. Breiman L. Random Forests. *Mach Learn* 2001;45(1):5–32.
38. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13–17; San Francisco, USA. New York: ACM; 2016. p. 785–794.
39. Raghu S, Sriraam N. Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures. *Expert Syst Appl* 2017;89:205–21.
40. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 2006;7:91.
41. Shalev-Shwartz S, Ben-David S. Understanding machine learning: from theory to algorithms. Cambridge (UK): Cambridge University Press; 2014.
42. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017;30:4765–4774.
43. Yalçın H. Interpretable Machine Learning Models for Early Detection of Subclinical Mastitis using Routine Milk Composition Data. ISPEC Tarım Bilimleri Dergisi [ISPEC *J Agric Sci*] 2025;9(3):830–843.
44. Ebrahimi M, Mohammadi-Dehcheshmeh M, Ebrahimie E. et al. Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep Learning and Gradient-Boosted Trees outperform other models. *Comput Biol Med* 2019;114:103456.
45. Bobbo T, Biffani S, Taccioli C. et al Comparison of machine learning methods to predict udder health status based on somatic cell counts in dairy cows. *Sci Rep* 2021;11(1):13642.
46. Raspanti CG, Bonetto CC, Vissio C. et al. Prevalence and antibiotic susceptibility of coagulase-negative *Staphylococcus* species from bovine subclinical mastitis in dairy herds in the central region of Argentina. *Revista Argentina de Microbiología* 2016;48(1):50–6.
47. Rychshanova R, Mendybayeva A, Miciński B. et al. Antibiotic resistance and biofilm formation in Staphylococcus aureus isolated from dairy cows at the stage

of subclinical mastitis in northern Kazakhstan. *Arch Anim Breed* 2022;65(4):439-448.
48. Israel LFS, Rabello RF, Domingos SCB. et al. Produção de biofilme por *Staphylococcus chromogenes* isolados de amostras de leite provenientes de rebanhos bovinos com mastite. *Arq Bras Med Vet Zootec* 2018;70(6):1943-1949.
49. Nogara KF, Busanello M, Zopollatto M. Relationship between subclinical mastitis occurrence and pathogen prevalence in two different automatic milking systems. *Animals* 2025;15(6):776.
50. Ebrahimie E, Ebrahimi F, Ebrahimi M. et al. Hierarchical pattern recognition in milking parameters predicts mastitis prevalence. *Comput Electron Agric* 2018;147:6–11.
51. Zecconi A, Zaghen F, Meroni G. et al. Machine learning approach for early lactation mastitis diagnosis using total and differential somatic cell counts. *Animals* 2025;15(8):1125.
52. Motohashi H, Ohwada H, Kubota C. Early detection method for subclinical mastitis in auto milking systems using machine learning. In: International Conference on Computational Intelligence and Knowledge Economy (ICCIKE); 16–17 Dec 2020, Dubai, United Arab Emirates. IEEE, 2020:389–93.
53. Guo R, Dai Y, Hu J. Research on the prediction model of mastitis in dairy cows based on time series characteristics. *Front Vet Sci* 2025;12:1575525.
54. Špakauskas V, Klimienė I, Matusevičius A. A comparison of indirect methods for diagnosis of subclinical mastitis in lactating dairy cows. *Vet Arhiv* 2006;76(2):101–109.
55. Lusis I, Antane V, Laurs A. Effectiveness of Mastitis Detection index for cow monitoring and abnormal milk detection in milking robots. In: Proceedings of the 16th International Scientific Conference: Engineering for Rural Development Conference; 24-26 May 2017, Jelgava, Latvia. Jelgava: Latvia University of Life Sciences and Technologies, 2017:1383–7.
56. Juozaitienė V, Juozaitis A, Brazauskas A. et al. Investigation of electrical conductivity of milk in robotic milking system and its relationship with milk somatic cell count and other quality traits. *J Meas Eng* 2015;3(3):55-92.
57. Bobbo T, Matera R, Pedota G. et al. Exploiting machine learning methods with monthly routine milk recording data and climatic information to predict subclinical mastitis in Italian Mediterranean buffaloes. *J Dairy Sci* 2023;106(3):1942–1952.
58. Sutera A, Louppe G, Huynh-Thu VA. et al. From global to local MDI variable importances for random forests and when they are Shapley values. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 2021.
59. Pegolo S, Tessari R, Bisutti V. et al. Quarter-level analyses of the associations among subclinical intramammary infection and milk quality, udder health, and cheesemaking traits in Holstein cows. *J Dairy Sci* 2022;105:3490–507.
60. Krömker V, Leimbach S. Mastitis detection: current trends and future perspectives. *J Dairy Res* 2020;87(2):114-121.

61. Halasa T, Huijps K, Østerås O. et al. Economic effects of bovine mastitis and mastitis management: a review. *Vet Q* 2007;29:18–31.
62. Forsbäck L, Lindmark-Månsson H, Andrén A. et al. Day-to-day variation in milk yield and milk composition at the udder-quarter level. *J Dairy Sci* 2010;93:1458–68.
63. Berry DP, Lee JM, Macdonald KA. et al. Associations among body condition score, body weight, somatic cell count, and clinical mastitis in seasonally calving dairy cattle. *J Dairy Sci* 2007;90:637–648.
64. Rahularaj R, Deshapriya RMC, Ranasinghe RMSBK. Influence of bovine sub-clinical mastitis and associated risk factors on calving interval in a population of crossbred lactating cows in Sri Lanka. *Trop Anim Health Prod* 2019;51:2413–2419.
65. Dharejo MN, Minogue L, Kabelitz T et al. Time series data analysis to predict the status of mastitis in dairy cows by applying machine learning models to automated milking systems data. *Prev Vet Med* 2025;242:106575.
66. Dharejo MN, Kashongwe O, Amon T. et al. Farm-Specific Effects in Predicting Mastitis by Applying Machine Learning Models to Automated Milking System and Other Farm Management Data. *Animals* 2025;15(19):2825.