Article

MasPA: A Machine Learning Application to Predict Risk of Mastitis in Cattle from AMS Sensor Data

Naeem Abdul Ghafoor [1,2,](ID) and Beata Sitkowska [2](ID)

1   Department of Molecular Biology and Genetics, Faculty of Science, Mugla Sitki Kocman University, 48000 Mugla, Turkey
2   Department of Animal Biotechnology and Genetics, Faculty of Animal Breeding and Biology, University of Science and Technology, 85-084 Bydgoszcz, Poland; beatas@utp.edu.pl
*   Correspondence: merzanaeem007@gmail.com

Abstract: Mastitis is a common disease that prevails in cattle owing mainly to environmental pathogens; they are also the most expensive disease for cattle in dairy farms. Several prevention and treatment methods are available, although most of these options are quite expensive, especially for small farms. In this study, we utilized a dataset of 6600 cattle along with several of their sensory parameters (collected via inexpensive sensors) and their prevalence to mastitis. Supervised machine learning approaches were deployed to determine the most effective parameters that could be utilized to predict the risk of mastitis in cattle. To achieve this goal, 26 classification models were built, among which the best performing model (the highest accuracy in the shortest time) was selected. Hyper parameter tuning and K-fold cross validation were applied to further boost the top model's performance, while at the same time avoiding bias and overfitting of the model. The model was then utilized to build a GUI application that could be used online as a web application. The application can predict the risk of mastitis in cattle from the inhale and exhale limits of their udder and their temperature with an accuracy of 98.1% and sensitivity and specificity of 99.4% and 98.8%, respectively. The full potential of this application can be utilized via the standalone version, which can be easily integrated into an automatic milking system to detect the risk of mastitis in real time.

Keywords: machine learning; dairy science; animal science; mastitis

1. Introduction

The global dairy industry was valued at around 720 billion USD in 2019, contributing to 54% of the global liquid milk share, and it is projected to grow to 1032 billion USD by 2024 [1]. However, the industry is not invincible, as cattle, like any other animal, can develop diseases. Among them is clinical mastitis, which is the single most expensive disease among the dairy industry, resulting in a loss of around 6% of the production value annually as a result of several factors such as reduction of production, treatment expenses, and milk discard, while also being among the top reasons for permanent removal of the cattle from the herd or even cattle mortality [2–4]. While 6% is not a high overall amount, the loss is drastic to small farms as the loss per cow can be significant, around 100–500 kg/cow/lactation or around a 5–7% decrease in milk yield per lactation [4,5]. Aside from lower milk yield, prevalence of mastitis results in a financial burden to the farmers as each clinical mastitis case involves therapeutic expenses, veterinary expenses, labor expenses, premature culling loss, non-saleable milk losses, future reproductive loss, replacement loss, and/or death loss, all of which could add up to 444 USD; this amount could be massive for smaller farms or those in low-income countries [6]. Antibiotics alone or in combination with non-steroidal anti-inflammatory drugs (NSAID) are often used for preventing mastitis, and they do indeed work efficiently in preventing most of the economic loss due to clinical mastitis; however, such a strategy should be strictly used for treatment as long term use of such drugs for prevention rather than treatment results in

antibiotics/drug residues reaching the end consumers, which results in drug/antibiotic resistance-related health issues, which themselves cost the human health care industry around 55 billion USD annually in the United States alone; in other words, the use of antibiotics for preventive care magnifies the burden and transfers it to other sectors rather than solving the core issue [7–9].

Mastitis is often caused by microbial infections (mainly bacterial) from the environment, either directly or through feed, eventually causing pathological lesions and inflammation of the mammary glands that could result in progressive fibrosis or even occurrence of severe toxemia in the cattle. The severity of the symptoms is determined mainly by the type of the pathogen and the resistance of the cattle's mammary gland [10,11]. The cattle's mammary gland is not entirely defenseless against these pathogens; the humoral and acquired immune response of the cattle for the most part could successfully prevent these pathogens from causing any damage; moreover, the lysosome enzyme found in the cattle's milk can also digest the peptidoglycan layer of the Gram-positive and Gram-negative bacteria, causing their death. Another glycoprotein, lactoferrin, found in milk and other secretions of the cattle, can also kill some bacterias by hindering their iron intake pathways; furthermore, animal breeding activities also consider mastitis risk when breeding cattle. One of the genetic traits considered is the somatic cell count (SCC), as SCCs contribute to the cattle's immune system, as low levels of SCC are directly correlated to higher risk of environmental mastitis, breeding programs tend to favor cattle with high SCCs; however, this approach is limited in practice [12].

As eradication of mastitis risk is rather quite difficult, preventive measures have been studied extensively, with antibiotics and probiotics being in the front line of most of the studies [13]. In a recent study on 108 Dutch dairy cattle, it was estimated that the cost of preventive measures against mastitis in cattle was around €120/cow/year, of which €81.6 (or 68%) went towards labor expenses with another average of €301/cow/year in the case of failure or clinical mastitis [14].

Several sensors based on factors such as milk color; temperature; SCC; electrical conductivity; thermal cameras; and/or enzyme based methods such as L-Lactate dehy-drogenase (LDH), N-acetyl-beta-D-glucosaminidase (NAGase), and haptoglobin (Hp), have been developed for commercial use (commonly as biosensors or immunosensors) in automatic milking systems in different farms to detect and alert the cattle at risk of mastitis before its prevalence; however, such single parameters can limit the sensitivity and specificity of the results, not to mention their expenses, cost of specialized labor and equipment, and limitation to only automatic milking systems, all of which could further limit their use in small farms and organic farms that do not utilize such advance systems [15–17].

A promising emerging approach in the early diagnosis of mastitis in cattle is the use biosensors, as such devices are effective at detecting pathogens even at low concentration; however, the pathogens found in cattle milk that could cause mastitis are quite diverse and the research on the development of a multi-pathogen detecting biosensor is yet to de-liver promising results. Among the largest projects aimed at delivering such technology was the Pathomilk project (Grant agreement ID: 30392), which received €1.7 million of funding for the development of a rapid biosensor that could potentially detect multiple pathogens commonly found in milk. The initial technology was based on a DNA hybridization coupled to surface plasmon resonance detection; however, with more than a decade past since the project's initiation, no significant outcomes have been reported. The use of immunoassays for the detection of pathogens in the milk is also limited owing to the heterogenous content of the milk that could hinder the antibody binding mechanisms involved in such assays. Likewise, the use of standard PCR-based methods is limited as the presence of ions (mainly calcium) plasmin, fats, and somatic cells makes it necessary to perform several filtrations before the PCR reaction, eventually increasing the overall cost of the diagnosis [18].

Recent developments in the field of artificial intelligence and machine learning have revolutionized many fields in recent years, with biotechnology not left behind. Especially

with Industry 4.0 initiatives, the internet of things (IoT) has made data acquisitions to perform such analyses more feasible than ever. Recently, the "Sack for Data" approach was proposed, which included four flex sensors and a temperature sensory to collect eight udder parameters along with temperature of the udder using Arduino and Raspberry pi boards only, which are extremely cheap to purchase and easy to use [19]. They have also utilized cloud technology to automate the data maintenance and using K-nearest neighbor (KNN) and support vector machine (SVM) algorithms, where they achieved 73% and 86% mastitis prediction rates, respectively. While these percentages are not perfect, they serve as a proof of concept to a cheap detection method for mastitis with affordable technology and some data analytic approach [19].

By the end of 2009, 8000 dairy farms utilized AMS, with the number growing continuously as AMS provides several benefits such as reduction in the labor cost, more time flexibility, and overall higher milk yield as cattle within AMS can be milking multiple times per day. Reports had also shown that cattle are calmer in such systems as they can be milked whenever they are most comfortable [20–22]. Another advantage of AMS it is an automated system that can collect consistent data, hence different sensors can be integrated into them to collect specific data including the amount milk, heat, milking time, and so on.

The aim of this study is to utilize the latest trends in data science and machine learning to develop semi-automated pipelines that could provide relief to farmers from the cost of mastitis preventive measures, which could add up to €120/cow/year, or at least reduce the contribution of labor expenses to €0 using the affordable Raspberry pi kits to manually collect data and predict the risk of mastitis through an online webserver or by integrating such kits to AMS for a fully automated data collection, which can be integrated into an open-source application to predict the risk of mastitis in real time. Such a solution could provide small farmers or farmers with limited technical background great advantages in monitoring their cattle's mastitis status without any external expenses and allow them to save on revenues [19,23].

2. Materials and Methods

The dataset used to train and build the machine learning model for predicting the risk of mastitis in this study was obtained from recent research lead by Ankitha (2020) [24]. This dataset contains 6600 entries (three entries per cattle) for cattle with 15 attributes; cow ID, date, breed, months since giving birth, previous occurrence of mastitis, front left udder inhale limit (IUFL) front left udder exhale limit (EUFL), front right udder inhale limit (IUFR), front right udder exhale limit (EUFR), rear left udder inhale limit (IURL) rear left udder exhale limit (EURL), rear right udder inhale limit (IURR), rear right udder exhale limit (EURR), temperature of the cow, the hardness of an udder (from user input via a switch), pain due to swelling of the udder (manual user input), photographs of the cow's milk, and a binary class label (healthy or mastitis). Among the attributes, only those with significant variance among the dataset and those that can be measured in a cost-effective manner were selected.

The raw dataset was preprocessed via SciPy tools and Scikit-learn library's feature selection function; unnecessary attributes such as ID, breed, hardness, and pain (which requires labor) were removed; and parameters that constituted less than 50% variance among all the entries were also removed [25,26]. The raw dataset contained 6600 samples; however, it was imbalanced with 3961 healthy cows (60.02%) and 2639 cows with mastitis (39.98%). To overcome the bias that might arise owing to this imbalance, the RandomOver-Sampler function from the imbalanced-learn library (within Scikit-learn) was utilized. This function takes the underrepresented class (here, the cows with mastitis) and generates sample inputs corresponding to it using its AI-based algorithm.

The balanced dataset was divided into training (6337 entries, 80%) and testing (1585 entries, 20%) subsets and a total of 26 classification algorithms were utilized to build 26 classification models (RandomForestClassifier, XGBClassifier, LGBMClassifier, Bagging-Classifier, DecisionTreeClassifier, ExtraTreeClassifier, KNeighborsClassifier, AdaBoostClas-

sifier, LabelPropagation, LabelSpreading, SupportVectorClassifier, QuadraticDiscriminant-Analysis, NuSupportVectorClassifier, SGDClassifier, RidgeClassifier, LogisticRegression, LinearDiscriminantAnalysis, RidgeClassifierCV, CalibratedClassifierCV, LinearSupport-VectorClassifier, GaussianNB, BernoulliNB, PassiveAggressiveClassifier, NearestCentroid, DummyClassifier, and Perceptron). They were built without any hyperparameter tuning (default parameters) and their accuracies were compared. The top performing classifier was selected and hyper parameter tuning via the grid search method was performed. The tuned model was then subjected to a 10-fold cross validation and its average mean accuracy was calculated along with its sensitivity and specificity. The sensitivity and specificity of the model were also calculated using Equation (1) and Equation (2) respectively.

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives } + \text{ False Negatives}} \tag{1}$$

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives } + \text{ False Positives}} \tag{2}$$

The true positives were calculated as the number of healthy cows that were predicted correctly, false positives were calculated as the number of healthy cows predicted to be at risk of mastitis, true negatives were calculated as the number of cows at risk of mastitis that were correctly predicted, and false negatives were calculated as the number of cows at risk of mastitis that were predicted to be healthy.

3. Results

3.1. Data Preprocessing

Following the preprocessing and features' selection steps, the eight-udder parameter (IUFL, EUFL, IUFR, EUFR, IURL, EURL, IURR, and EURR) and temperature attributes of the cattle were sufficient to generate a functional model with significant accuracy. The remaining attributes were dropped as their contributions were insignificant; photographs of the milk and attributes like pain/hardness attributes were also dropped as they are open to bias (by the labors' interpretation) and would contribute to higher labor cost. This dataset was further balanced with RandomOverSampler and the final dataset contained 50% healthy and 50% mastitis samples, totaling 7922 samples. The final curated dataset is provided in Supplementary Material 1 (S1).

3.2. Model Fitting and Hyperparameter Optimization

The accuracy scores and time taken to build the models with the selected algorithms are summarized in Table 1. The best performing model was random forest classifier with an initial accuracy of 99.117%; the suggested parameters from the hyper tuning utilized to build the final model are summarized in Table 2.

Table 1. Accuracy benchmarks and time taken to build each of the selected 26 models.

Model Name [M]	Accuracy (%) [L]	Time Taken (s)
RandomForestClassifier	99.12	0.60
XGBClassifier	98.68	0.38
LGBMClassifier	98.61	0.17
BaggingClassifier	98.55	0.17
DecisionTreeClassifier	97.85	0.07
ExtraTreeClassifier	96.97	0.04
KNeighborsClassifier	93.56	0.13
AdaBoostClassifier	93.19	0.50
LabelPropagation	91.74	1.82
LabelSpreading	91.48	2.65
SupportVectorClassifier	87.89	0.61
QuadraticDiscriminantAnalysis	87.63	0.05

Table 1. *Cont.*

Model Name [M]	Accuracy (%) [L]	Time Taken (s)
NuSupportVectorClassifier	87.44	1.55
SGDClassifier	87.13	0.09
RidgeClassifier	86.69	0.05
LinearDiscriminantAnalysis	86.62	0.07
RidgeClassifierCV	86.62	0.10
CalibratedClassifierCV	86.50	1.25
LinearSVC	86.44	0.32
LogisticRegression	86.06	0.11
Perceptron	83.41	0.05
PassiveAggressiveClassifier	71.67	0.05
GaussianNB	69.34	0.04
BernoulliNB	67.89	0.04
NearestCentroid	62.08	0.05
DummyClassifier	49.21	0.03

All models were generated with default parameters using their respective Scikit-learn classifier/algorithm (via lazypredict library). [M] All models were build using the same training and testing sets (same random 80–20 split). [L] Accuracy percentages are from the respective model's performance on the testing set.

Table 2. Parameters used to build the best performing model.

Model name	Parameters
Random Forest Classifier	bootstrap=True, ccp_alpha=0.0, class_weight=None, criterion='entropy', max_depth=None, max_features='auto', max_leaf_nodes=None, max_samples=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=100, n_jobs=None, oob_score=False, random_state=1000, verbose=0, warm_start=False

Figure 1 depicts the confusion matrix of the best performing model (i.e., hyper parametrized random forest model). The sensitivity (the percentage of cows in risk of mastitis that were correctly predicted) and specificity (the percentage of cows possessing no risk to mastitis that were correctly predicted) of the model calculated from the confusion matrix were calculated via Equations (1) and (2), and were found to be 99.36% and 98.77%, respectively. The mean accuracy of the model following the 10-fold cross validation was calculated to be 98.10%, with a standard deviation of 0.043.
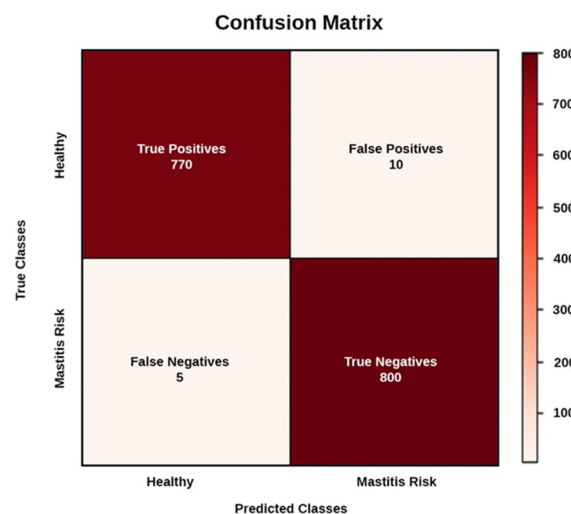

Figure 1. Confusion matrix of the top model (based on the random forest algorithm). The matrix is based on the performance of the model on the test set comprising 1585 samples (780 healthy and 805 mastitis).

3.3. Web App Usage and Local Deployment

The web application developed based on the top model build can be accessed at https://share.streamlit.io/naeemmrz/maspa.py/main/MasPA.py (accessed on 2 August 2021); the user needs to input the eight-udder inhale and exhale limits and the temperature of the cattle (with an optional identifier for each row); a sample input file is provided in Supplementary Material 2 (S2) and can be downloaded from the web interface as well, and the general interface of the web app is explained in Figure 2. For real-time usage or integration with automatic milking systems, both the model "RndmForest_mastistis.pkl" and the application source code "MasPA.py" are available as open source at the author's GitHub page along with a step-by-step guide https://github.com/naeemmrz/MasPA.py (accessed on 2 August 2021).

Figure 2. The MasPA interface. (**a**) Link to input example/template, (**b**) options to add input file, (**c**) authors and affiliations, (**d**) preview of the user input, and (**e**) prediction results and link to download the results as a comma separated file (.csv).

4. Discussion

Mastitis single-handedly costs the dairy industry around 6% of its production value, and this contribution is expected to grow as the demand for more milk production per cattle increases [4]. While several preventive measures are available for early diagnosis of mastitis in cattle, most of these measurements are either too expensive or impractical and inaccessible for small farmers or farmers from low-income countries. The cost of preventive measures for early diagnosis of mastitis in cattle for farmers is up to €120/cow/year, which could contribute a significant amount to the budget [14].

Different farms around the world opt for different preventive measures depending on their geographic region, size of herd, and their revenue; these measures could range from classical inexpensive methods such as temperature monitoring and forestripping that can be performed by any labor, to more sophisticated methods such as LDH, NAGase, immunoassays, and biosensors that provide much higher accuracy, but have higher costs higher, and access could be limited by region [17,18].

Recent developments in the field of data science and artificial intelligence have opened a lot of opportunities in developing new methods of diagnosis and detection of diseases by deploying sophisticated algorithms to these problems. Several attempts have been made in diseases affecting humans, including many other species of plants and animals [19,27–30].

The aim of this study was to integrate these recent developments in the field of data science to derive a solution in predicting the risk of mastitis in cattle before it occurred so as to reduce the high cost of treatment, encourage farmers to avoid using antibiotics as a preventive measure, and reduce unnecessary veterinary expenses by providing an open-source tool accessible online free-of-cost.

The integration of machine learning and deep learning technologies for the prediction of clinical and sub-clinical mastitis in cattle is not a novel approach. Indeed, quite recently, Ebrahimi (2019) analyzed parameters such as milk volume, lactose concentration, electrical conductivity, protein concentration, peak flow, and milking time for 364,249 milking instances in cattle, and applied several deep learning and machine learning algorithms to determine the best statistical model that could predict the risk of sub-clinical mastitis. Their study concluded that the gradient-boosted tree algorithm provided the best accuracy of 84.9% from the former parameters, with the random forest algorithm ranking as the worst performing algorithm, with an accuracy of 82.3% [31]. Following a similar path, Fadul-Pacheco (2021) investigated the efficiency of naïve Bayes, random forest, and extreme gradient boosting on the dataset from the Dairy Brain project for early prediction of clinical mastitis. Their study, however, concluded with random forest being the best performing algorithm, with an accuracy of 71% for the first lactation and 85% for the continuous (follow up) model, respectively [32]. These results indicate that different algorithms perform differently when applied to different parameters/attributes. As shown in Table 1, the random forest algorithm performed the best on the attributes in the dataset used for this study.

MasPA is an ML-based solution that can predict the risk of mastitis in the cattle from the inhale and exhale limits for each of the cattle's four udders and the cattle's temperature, which can be collected via highly affordable sensors either manually (for farms with conventional milking systems) or by integrating such sensors into AMS. As shown in Figure 1, MasPA is based on the random forest algorithm, and can predict the risk of mastitis in cattle with a near-perfect accuracy of 98.10%. The application is available as a web application, free of any cost and/or limitations.

To address the inaccessibility to internet for some farms/farmers, we also provided a standalone package MasPA.py (Supplementary Material 3 (S3)) that can run almost on any computer locally while providing the same web interface. Considering the potential lack of technical literacy in some farmers, the interface of the package was made to be as simple as possible (Figure 2); furthermore, the source code of the application is also available open source, so anyone could modify, optimize, or integrate it into their local system (such as AMS) and/or modify and apply it to different datasets to predict different diseases.

5. Conclusions

The proposed web application, MasPA, which is based on the random forest algorithm, can predict the risk of mastitis in cattle from the inhale and exhale limits of their udder and their temperature with an accuracy of 98.10% and sensitivity and specificity of 99.4% and 98.8%, respectively. The full potential of this application can be utilized via the standalone version (available as open source), which can be easily integrated into AMS to detect the risk of mastitis in real time.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/agriengineering3030037/s1, S1: Curated data used for model building (.csv); S2: Sample input file (.csv); and S3: MasPA.py source code (python code).

References

1. Shahbandeh, M. Statista. Global Dairy Industry—Statistics & Facts. 2021. Available online: https://www.statista.com/topics/4649/dairy-industry/ (accessed on 14 June 2021).
2. Gill, R.; Howard, W.H.; Leslie, K.E.; Lissemore, K. Economics of Mastitis Control. *J. Dairy Sci.* **1990**, *73*, 3340–3348. [CrossRef]
3. Miller, R.; Paape, M.; Fulton, L.; Schutz, M. The Relationship of Milk Somatic Cell Count to Milk Yields for Holstein Heifers After First Calving. *J. Dairy Sci.* **1993**, *76*, 728–733. [CrossRef]
4. Shim, E.; Shanks, R.; Morin, D. Milk Loss and Treatment Costs Associated with Two Treatment Protocols for Clinical Mastitis in Dairy Cows. *J. Dairy Sci.* **2004**, *87*, 2702–2708. [CrossRef]
5. Hortet, P.; Seegers, H. Loss in milk yield and related composition changes resulting from clinical mastitis in dairy cows. *Prev. Vet. Med.* **1998**, *37*, 1–20. [CrossRef]
6. Rollin, E.; Dhuyvetter, K.C.; Overton, M.W. The cost of clinical mastitis in the first 30 days of lactation: An economic modeling tool. *Prev. Vet. Med.* **2015**, *122*, 257–264. [CrossRef] [PubMed]
7. Dadgostar, P. Antimicrobial Resistance: Implications and Costs. *Infect. Drug Resist.* **2019**, *12*, 3903–3910. [CrossRef] [PubMed]
8. Rana, S.; Lee, S.Y.; Kang, H.J.; Hur, S.J. Reducing Veterinary Drug Residues in Animal Products: A Review. *Food Sci. Anim. Resour.* **2019**, *39*, 687–703. [CrossRef]
9. Jayalakshmi, K.; Paramasivam, M.; Sasikala, M.; Tamilam, T.; Sumithra, A. Review on antibiotic residues in animal products and its impact on environments and human health. *J. Entomol. Zool. Stud.* **2017**, *5*, 1446–1451.
10. Bianchi, R.M.; Schwertz, C.I.; De Cecco, B.S.; Panziera, W.; De Lorenzo, C.; Heck, L.C.; Snel, G.G.M.; Lopes, B.C.; Da Silva, F.S.; Pavarini, S.P.; et al. Pathological and microbiological characterization of mastitis in dairy cows. *Trop. Anim. Health Prod.* **2019**, *51*, 2057–2066. [CrossRef]
11. Benites, N.R.; Guerra, J.L.; Melville, P.A.; Costa, E.O. Aetiology and Histopathology of Bovine Mastitis of Espontaneous Occurrence. *J. Veter. Med. Ser. B* **2002**, *49*, 366–370. [CrossRef]
12. Pyörälä, S. New Strategies to Prevent Mastitis. *Reprod. Domest. Anim.* **2002**, *37*, 211–216. [CrossRef]
13. Koba, I.S.; Lysenko, A.A.; Koshchaev, A.G.; Shantyz, A.K.; Donnik, I.M.; Dorozhkin, V.I.; Shabunin, S.V. Prevention of Mastitis in Dairy Cows on Industrial Farms. *J. Pharm. Sci. Res.* **2018**, *10*, 2582–2585. Available online: https://www.jpsr.pharmainfo.in/Documents/Volumes/vol10Issue10/jpsr10101840.pdf (accessed on 14 July 2021).
14. Van Soest, F.; Santman-Berends, I.M.; Lam, T.J.; Hogeveen, H. Failure and preventive costs of mastitis on Dutch dairy farms. *J. Dairy Sci.* **2016**, *99*, 8365–8374. [CrossRef] [PubMed]
15. Hogeveen, H.; Kamphuis, C.; Steeneveld, W.; Mollenhorst, H. Sensors and Clinical Mastitis—The Quest for the Perfect Alert. *Sensors* **2010**, *10*, 7991. [CrossRef]
16. Hovinen, M.; Siivonen, J.; Taponen, S.; Hänninen, L.; Pastell, M.; Aisla, A.-M.; Pyörälä, S. Detection of Clinical Mastitis with the Help of a Thermal Camera. *J. Dairy Sci.* **2008**, *91*, 4592–4598. [CrossRef] [PubMed]
17. Chagunda, M.G.; Larsen, T.; Bjerring, M.; Ingvartsen, K.L. L-lactate dehydrogenase and N-acetyl-β-D-glucosaminidase activities in bovine milk as indicators of non-specific mastitis. *J. Dairy Res.* **2006**, *73*, 431–440. [CrossRef]
18. Martins, S.; Martins, V.C.; Cardoso, F.A.; Germano, J.; Rodrigues, M.; Duarte, C.; Bexiga, R.; Cardoso, S.; Freitas, P.P. Biosensors for On-Farm Diagnosis of Mastitis. *Front. Bioeng. Biotechnol.* **2019**, *7*, 186. [CrossRef]
19. Ankitha, K.; Manjaiah, D.H. Comparison of KNN and SVM Algorithms to Detect Clinical Mastitis in Cows Using Internet of Animal Health Things. *Adv. Intell. Syst. Comput.* **2020**, 51–60. [CrossRef]

20. Vijayakumar, M.; Park, J.H.; Ki, K.S.; Lim, D.H.; Kim, S.B.; Park, S.M.; Jeong, H.Y.; Park, B.Y.; Kim, T.I. The effect of lactation number, stage, length, and milking frequency on milk yield in Korean Holstein dairy cows using automatic milking system. *Asian Australas. J. Anim. Sci.* **2017**, *30*, 1093–1098. [CrossRef]

21. Steeneveld, W.; Tauer, L.; Hogeveen, H.; Lansink, A.O. Comparing technical efficiency of farms with an automatic milking system and a conventional milking system. *J. Dairy Sci.* **2012**, *95*, 7391–7398. [CrossRef] [PubMed]

22. Castro, A.; Pereira, J.M.; Amiama, C.; Bueno, J. Estimating efficiency in automatic milking systems. *J. Dairy Sci.* **2012**, *95*, 929–936. [CrossRef]

23. Gay, W. SD Card Storage. In *Raspberry Pi Hardware Reference*; Apress: Berkeley, CA, USA, 2014; pp. 81–88.

24. Ankitha, K.; Manjaiah, M.; Kartik, M. Data for: Clinical mastitis in cows based on udder parameter using Internet of Things (IoT). *Mendeley Data* **2020**, V2. [CrossRef]

25. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef] [PubMed]

26. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

27. Mansour, R.F.; El Amraoui, A.; Nouaouri, I.; Diaz, V.G.; Gupta, D.; Kumar, S. Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems. *IEEE Access* **2021**, *9*, 45137–45146. [CrossRef]

28. Prabhakar, B.; Singh, R.K.; Yadav, K.S. Artificial intelligence (AI) impacting diagnosis of glaucoma and understanding the regulatory aspects of AI-based software as medical device. *Comput. Med. Imaging Graph.* **2021**, *87*, 101818. [CrossRef]

29. Sujatha, R.; Chatterjee, J.M.; Jhanjhi, N.; Brohi, S.N. Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocess. Microsyst.* **2021**, *80*, 103615. [CrossRef]

30. McKinney, S.M.; Sieniek, M.; Godbole, V.; Godwin, J.; Antropova, N.; Ashrafian, H.; Back, T.; Chesus, M.; Corrado, G.S.; Darzi, A.; et al. International evaluation of an AI system for breast cancer screening. *Nat. Cell Biol.* **2020**, *577*, 89–94. [CrossRef]

31. Ebrahimi, M.; Mohammadi-Dehcheshmeh, M.; Ebrahimie, E.; Petrovski, K.R. 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. [CrossRef]

32. Fadul-Pacheco, L.; Delgado, H.; Cabrera, V.E. Exploring machine learning algorithms for early prediction of clinical mastitis. *Int. Dairy J.* **2021**, *119*, 105051. [CrossRef]