Time series data analysis to predict the status of mastitis in dairy cows by applying machine learning models to automated milking systems data

Muhammad N Dharejo [1*], Olivier Kashongwe [2, 3], Thomas Amon [2, 4], Tina Kabelitz [2], Lukas Minoque [2], Marcus G Doherr [1]

[1] Institute for Veterinary Epidemiology & Biostatistics, School of Veterinary Medicine, Freie Universität Berlin, [2] Department of Sensors and Modelling, Leibniz Institute for Agricultural Engineering and Bioeconomy e.V. (ATB), [3] Joint-Lab Artificial Intelligence and Data Science, University Osnabrück, [4] Institute for Animal Hygiene and Environmental Health, School of Veterinary Medicine, Freie Universität Berlin

1 Introduction & Aims

- Mastitis is a significant disease affecting dairy cows, leading to economic losses and welfare concerns[1].
- Traditional detection methods rely on manual inspections, which are labor-intensive[2].
- Automated Milking Systems (AMS) generate large datasets that can be leveraged using Machine Learning (ML) for early mastitis prediction[3].
- This study applies ML techniques to AMS data to predict mastitis occurrence: (i) one day prior and (ii) on the day of mastitis observation.

3 Results

3.1 Prediction One Day Before Mastitis Observation

- **Performance Metrics:**
- Accuracy: **0.80 - 0.90**
- Sensitivity: **0.64 - 0.78**
- Specificity: **0.80 - 0.90**
- Area under curve: **0.80 - 0.84**
- **Best Performing Model:**
- SVM (highest accuracy & specificity)
- MLP (highest sensitivity)
- LR & RF (highest AUC)

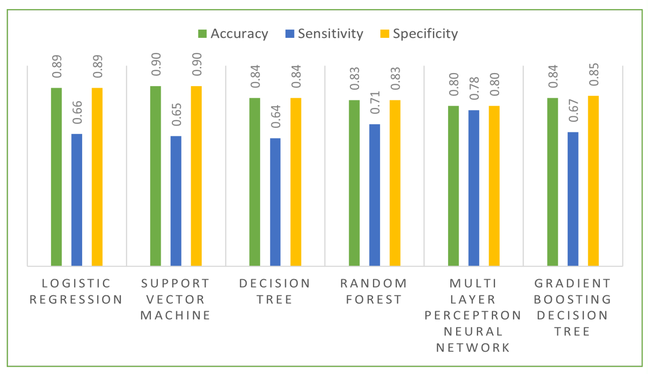

Figure 1: ML scores for predicting mastitis one day prior.

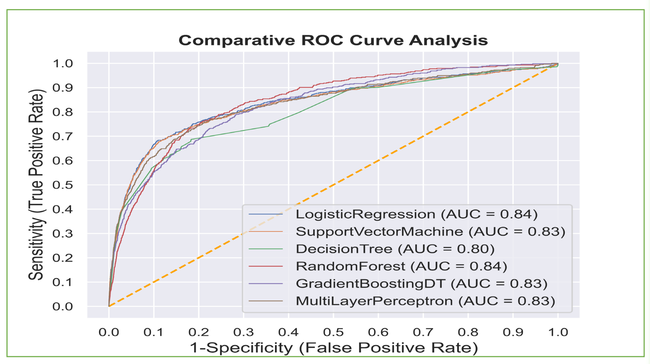

Figure 2: AUC scores for predicting mastitis one day prior.

3.2 Prediction on the Day of Mastitis Observation

- **Performance Metrics:**
- Accuracy: **0.84 - 0.93**
- Sensitivity: **0.76 - 0.91**
- Specificity: **0.84 - 0.93**
- Area under curve: **0.92 – 0.95**
- **Best Performing Model:**
- SVM (highest accuracy & specificity)
- MLP (highest sensitivity)
- RF, GB & MLP (highest AUC)

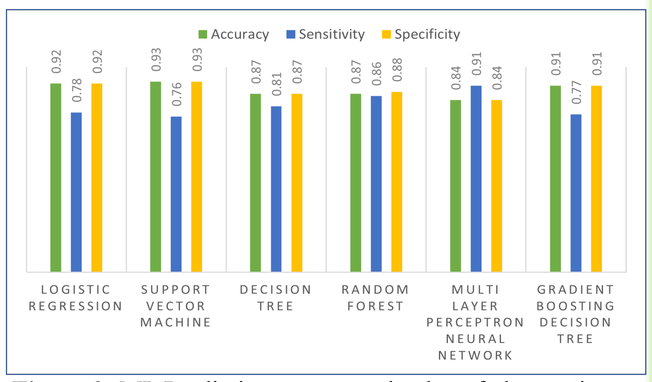

Figure 3: ML Prediction scores on the day of observation.

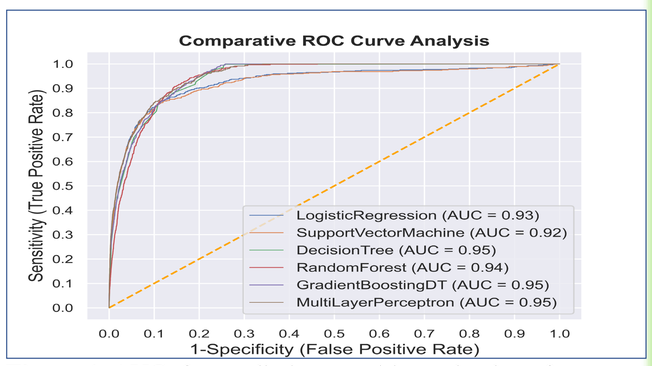

Figure 4: AUC for predicting mastitis on the day of observation.

2 Methods

- **Data Collection:**
  - AMS data from two dairy farms in Germany (Brandenburg & Saxony) covering four years (2019-2022).
  - 2.73 million daily milking records from 1790 cows.
- **Variables:**
  - Seven predictor variables: electrical conductivity, somatic cell count, milk yield, milking flow, milk temperature, fat content, and protein content.
  - Outcome variable: Mastitis events recorded as positive upon veterinary treatment.
- **ML Models Applied:**

  1. Logistic Regression (LR) 2. Support Vector Machine (SVM) 3. Decision Tree (DT)

  4. Random Forest (RF) 5. Gradient Boosting (GB) 6. Multi-Layer Perceptron (MLP)
- **Preprocessing & Model Training:**
  - Data cleaning, feature engineering, and transformation to autoregressive time series format.
  - Synthetic Minority Oversampling Technique (SMOTE) used for class balance.
  - GridSearchCV for hyperparameter tuning and Stratified 5-fold cross-validation.
  - Evaluation based on Accuracy, Sensitivity, Specificity, and Area under curve (AUC).

4 Discussion

- ML models performed better on the day of observation compared to one day before.
- Results aligned with or outperformed similar studies.
- Models demonstrated potential for real-time mastitis detection but require further refinement.
- **Future Recommendations:**
  - Incorporate additional AMS variables (e.g., lactose content, lactation stage).
  - Integrate sensor data from rumen bolus, pedometers, and environmental sensors.
  - Develop farm-specific predictive models for enhanced performance.

5 Conclusion

- ML models effectively predict mastitis using AMS-generated data.
- High accuracy and specificity make them promising for early disease detection
- Further refinement and sensor integration are necessary for real world application

References

1. Petersson-Wolfe, C.S., Leslie, K.E., Swartz, T.H., 2018. An Update on the Effect of Clinical Mastitis on the Welfare of Dairy Cows and Potential Therapies. Vet Clin North Am Food Anim Pract 34, 525–535. https://doi.org/10.1016/j.cvfa.2018.07.006
2. Bar, D., Gröhn, Y.T., Bennett, G., González, R.N., Hertl, J.A., Schulte, H.F., Tauer, L.W., Welcome, F.L., Schukken, Y.H., 2007. Effect of Repeated Episodes of Generic Clinical Mastitis on Milk Yield in Dairy Cows. Journal of Dairy Science 90, 4643–4653. https://doi.org/10.3168/jds.2007-0145
3. Ozella, L., Brotto Rebuli, K., Forte, C., Giacobini, M., 2023. A Literature Review of Modelling Approaches Applied to Data Collected in Automatic Milking Systems. Animals 13, 1916. https://doi.org/10.3390/ani13121916

Funding & Acknowledgments:
- This research was funded by the **German Federal Ministry for Food and Agriculture (BMEL)** under project **MEDICow** (grant number 28N206601).
- Special thanks to **Dairy Farmers** and collaborators.

Gefördert durch Bundesministerium für Ernährung und Landwirtschaft aufgrund eines Beschlusses des Deutschen Bundestages

Projektträger Bundesanstalt für Landwirtschaft und Ernährung

Contact:
House 21, Königsweg 67, 14163, Berlin
Email: m.dharejo@fu-berlin.de