2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications: Track 4: Applications and Business

NB-IoT Estrus Detection System of Dairy Cows Based on LSTM Networks

Nan Ma [*] [†], Lin Pan [*], Shihao Chen [*], and Baoling Liu [*]
[*]State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications, Beijing, China
[†]CETC Key Laboratory of Data Link Technology, Xi'an, China
Email: {manan, panl}@bupt.edu.cn

***Abstract*—To improve the revenue of dairy farms, cow estrus must be accurately monitored to track mating time. Narrow Band Internet of Things (NB-IoT) is considered as a promising technology to realize cost-effective detection system attributing to its wide coverage and low power consumption. To increase the success rate of real-time detection, machine learning based algorithms have been applied to extract patterns from estrus data. However, due to the lack of multivariate time series data, most previous studies do not consider using the time correlation to guide estrus detection. In this paper, we present a NB-IoT based solution framework where multivariate behavioral time series data collected by neck-mounted sensors and then uploaded to a cloud data center for further analysis through NB-IoT network. Based on the collected data, we propose an estrus prediction algorithm which gives estrus alert by exploiting Long-Short Term Memory (LSTM) and Convolution Neural Network (CNN). Through numerical studies conducted using real data set from the pasture, we show that our proposed solution outperforms exiting detection algorithms in terms of accuracy and efficiency.**

***Index Terms*—Estrus Detection, NB-IoT, LSTM, Multivariate Time Series**

I. INTRODUCTION

Accurate and effective detection of estrus is crucial to improve milk yield and reproductive efficiency in the dairy industry. Dairy farms need to track the estrous status of cows and complete the mating process during the estrus period. Traditional estrus detection methods include visual observation, rectal palpation, vaginal measurements [1], tail painting [2], etc. However, manual detection methods frequently misdiagnose dairy cows' estrus status, which would further lead to a declining pregnancy rate and a reduction in milk production. Some pastures try to solve the problem of the missed estrous period by injecting estrous synchronization drugs into dairy cows. Unfortunately, this high-cost solution is harmful to milk quality and dairy cows.

With the advent of large-scale Internet of Things (IoT) technology, this paper build a real-time monitor system of dairy cows' physiologic conditions. These real-time data streams are then automatically processed as the input of a LSTM based estrus prediction model, and dairy cows diagnosed as positive will receive artificial breeding in time. With the help of IoT technology and Machine Learning algorithms, it can effectively shorten the breeding cycle of dairy cows, reduce labor costs, and improve economic benefits.

This work was supported by the National Key Research and Development Project of China under Grant 2018YFB2100202.

978-1-7281-4490-0/20/$31.00 © 2020 IEEE

II. SYSTEM PRINCIPLE AND FRAMEWORK

A. Behavioral Characteristics of Estrus

The whole estrus period of cows can be divided into three stages, namely proestrus, estrus, and metestrus. In the course of proestrus, cows show the symptoms of sensitivity, increased activity index level, restlessness, frequently growl, and sniffing other cows. They would try to mount others and refused to be mounted by other cows. During the estrus period, cows are frequently moving around and willing to be mounted by others. Only a small number of dairy cows still show estrus behaviors during the metestrus period. Behavioral changes are associated with estrus status, which leads us to explore the possibility of detecting dairy cows' estrus by monitoring their behavioral characteristics. In general, the increased activity and the significantly reduced rumination time can be used as a helpful signal for judging the estrus status of cows. Dairy cows will not ovulate until 10 to 14 hours after the end of estrus period. Therefore, artificial insemination should be implemented 12 to 18 hours after the estrus [3]. Proper artificial insemination timing can improve the fertility rate. Therefore, accurate detection of the proestrus and estrus stage is beneficial to improve the success rate of mating.

B. Background

With the development of image processing technology, video monitor systems are developed for estrus detection [4] [5]. However, this method is difficult to implement when cows herd on pasture. Nowadays, many sensor-based wireless detection systems are used in estrus detection. Sensors can be roughly separated into three categories according to the wearing manners, i.e., pedometer, ear-tag, and neck-mounted accelerometer. A comparison of these three types of sensors are shown in Table I. As shown in Table I, pedometer is comparatively accurate in collecting the number of cows' step and activity index level. However, this type of sensor can't gather data of rumination and feeding. The body temperature collected by the ear-tag sensor is relatively perfect, but due to the involuntary movement of the ear, there is strong noise

978-1-7281-4490-0/20/$31.00 ©2020 IEEE

TABLE I
COMPARISON OF DIFFERENT SENSOR TYPES

	Manner of Wearing		
	Pedometer	***Ear-tag***	***Neck-mounted***
Data Type	Number of steps, Activity	Body Temperature, Activity	Number of steps, Activity
Disadvantage	No rumination and feeding data	Big noise in the activity data	No body temperature data

TABLE II
COMPARISON OF DIFFERENT COMMUNICATION METHODS

	ZigBee	***Bluetooth***	***NB-IoT***
Networking	Based on ZigBee gateway	Based on Bluetooth Mesh gateway	Based on existing cellular network
Typical Transmission Distance	100m	10m	10km
Typical Battery Life (AA battery)	Theoretically 2 years	Several days	Theoretically 10 years
Transmission Speed)	Less than 100Kbps	About 1Mbps	Less than 100Kbps

in the collected data. The neck-mounted sensors can obtain the number of steps but can not collect the body temperature data. In this paper, the experimental data is collected by neck-mounted sensors and then converted into multivariate data, including high activity, medium activity, low activity, feeding, and rumination. This data can be used not only for estrus detecting but also for health monitoring, disease warning, and abortion warning.

Collected data needs to be uploaded to the cloud through the communication network. To date, several systems have been designed for the automatic monitoring and transmission of dairy cows' daily status. Afimilk's Afitag cow pedometer adopts a Bluetooth communication approach, where cow's number of steps is only accessible at milking. Therefore, it can not provide real-time estrus monitoring. System developed by Li et al. [6] have realized real-time data collection by transmitting the data to a mesh gateway through short-distance communication method, ZigBee. However, due to limited coverage, a large number of gateways need to be deployed. When applied to large area grazing ranches, the system cost is too high.

Narrow Band Internet of Things (NB-IoT) is an promising technology which has the advantages of large network capacity, high reliability, high network security, and wide-area coverage. NB-IoT enables devices to directly connect to a relatively stable carrier network without any user-end maintenance. Theoretically, the low-power feature of NB-IoT enables 10 years of battery life without charging. Therefore, NB-IoT is suitable for data transmission between sensors and the cloud. A comparison of different communication methods, including Bluetooth, ZigBee, and NB-IoT, are shown in Table II.

C. System Framework

This part provides an overview of the system architecture. As shown in Fig. 1, the estrus monitoring system includes data collection layer, data transmission layer, and cloud processing layer. The neck-mounted smart sensor, as shown in Fig. 2 with 3D accelerometer worn on the cow's neck can capture 100 frames of data per second. The collected data needs to

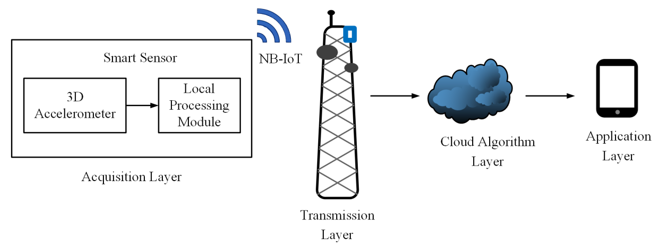

Fig. 1. System Framework.

Fig. 2. Neck-mounted Sensor and Mounting of Cows.

be converted into seconds and then hours level in the local processing module to adapt the transmission rate of NB-IoT.

Data is uploaded and processed, including the number of steps, minutes of dairy cows in different activity index level, feeding, and rumination in one hour. In our system, the cow's activities are divided into three categories, namely high, medium, and low activity. The term high activity here refers to walking or mounting. Medium activity refers to the chewing or shaking head when standing or lying still. And low activity refers to lying or standing without any other movement. To further reduce power consumption, the smart collar uploads the record of one cow to the cloud once every 2 hours via the NB-IoT network. After receiving the data, the cloud will use a LSTM-based estrus detection model to detect dairy cows' estrus status and send estrus alarms.

III. Related Detection Method

In recent years, many studies in animal science have investigated the effective detection of cows' estrus based on Machine Learning (ML) algorithms. The activity index level has been widely used to measure the activity change of dairy cows. Shahriar et al. studied the relationship between the estrus event and activity index level, and they grouped the activities by K-means clustering algorithm, then labeled as high, medium, and low level [7]. Yin et al. used the data generated by K-means clustering algorithm and Support Vector Machine (SVM) to establish an activity classification model, evaluating the onset of estrus by cow's activity variation [8]. Nonetheless, it is prone to get inaccurate estrus detection results when only the activity index level is considered, especially there is noise in the collected data. In order to further improve the accuracy of estrus detection, multivariate data was applied to ML algorithms. Minegishi et al. utilized a logistic regression model for estrus detection using activity and rumination data [9]. Borchers et al. compared Random Forest, Linear Discriminant Analysis, and Neural Network in the estrus detection by using activity, lying, and rumination data [10]. Shahinfar et al. researched different machine learning algorithms such as Naive Bayes, Bayesian Network, Decision Tree, Bootstrap Aggregation, and Random Forest to forecast the insemination result [11]. However, these methods fail to make full use of the time correlation of the cow estrus process.

IV. Proposed Method

A. Problem Formulation

For a given dairy cow feature sequences set $\boldsymbol{A}$, $\boldsymbol{\alpha} = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_T}\}$ is a time sequence of length T and $\boldsymbol{a_t} = [a_t^1, a_t^2, \ldots, a_t^m]$ is a feature vector of length m at time t, where m is the number of feature. In this paper, $m = 6$ and $0 < t \leq n$. $\boldsymbol{h^k(\alpha)} = \{\boldsymbol{a_{T+1-k}}, \boldsymbol{a_{T+2-k}}, \ldots, \boldsymbol{a_{T-1}}, \boldsymbol{a_T}\}$ is the prefix of length k $(0 < k \leq T)$ of time sequence $\boldsymbol{\alpha}$.

Let $\boldsymbol{E}$ be the status domain, i.e., the set of all possible estrus identifiers. We assume whether a dairy cow is estrus or not can be characterized by various features, e.g., the number of steps, the minutes of activity, feeding, and rumination. All six of these features are time series data. A function of $f(\boldsymbol{h^k(\alpha)}) \rightarrow E \in \boldsymbol{E}$ that utilizes a k length time series data is bulit to predict the estrus status at time $T+1$.

B. Recurrent Neural Networks

Recurrent Neural Networks (RNNs) performed well in processing time series data [12]. As shown in Fig. 3-(1), connections between neurons in RNN form a directed cycle [13]. As shown in Fig. 3-(2), each step in unfolded RNN is referred to as a time step. Hidden neurons at time t would receive both the output from the previous round and the current input. Thus the data at time $t-1$ will affect the output at time t. Such a structure is very effective in maintaining dependencies in time sequence data.

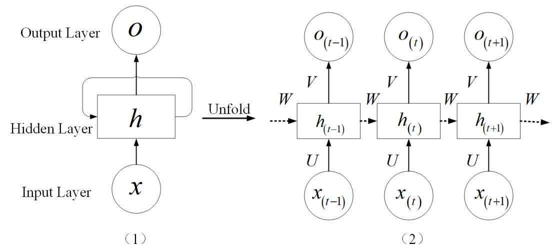

Fig. 3. A recurrent neural network.

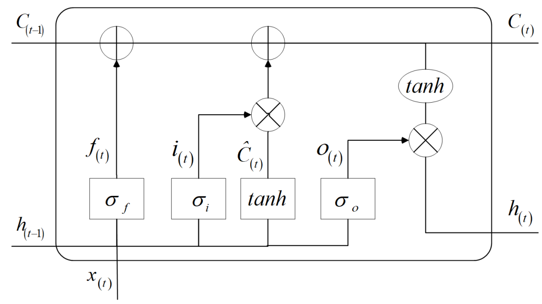

Fig. 4. Standard LSTM unit.

To solve the long-term dependency issue and vanishing-gradient problem in normal RNNs [14], Hochreiter et al. improved RNN and proposed a Long-Short Time Memory (LSTM) network model [15]. The structure of LSTM is basically the same as RNN, and the main difference is that the former has a more complex memory cell. The structure of LSTM unit can control the flow of information, which improves the capability of learning long-term dependencies.

As shown in Fig. 4, there are three kinds of gate in the standard LSTM units: input gate i_t, forgetting gate f_t, and output gate o_t. These units could be formulated as follows:

$$\begin{aligned} f_t &= \sigma_f\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \\ i_t &= \sigma_i\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \\ \tilde{C}_t &= \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t &= \sigma_o\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \\ h_t &= o_t \odot \tanh\left(C_t\right) \end{aligned} \quad (1)$$

$\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function, $\odot$ is the element-wise product, W and b represent the weights matrix and the biases, respectively. x_t is the input at time t.

The proposed Network for estrus detection is presented in Fig. 5. The first layers of the estrus detection network is a simple convolutional network, which aims to extract local dependencies between variables. The output of the convolutional layer is fed into the Recurrent component. The Recurrent component is a recurrent layer with the LSTM unit. Then the outputs of the Recurrent component are combined by a dense layer for the actual prediction. We use a many-to-one architecture for the one-step prediction, which means that the estrus detection network will take advantage of the prefix

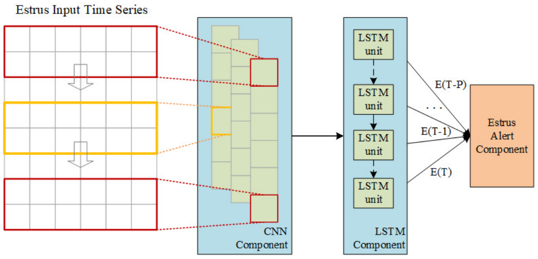

Fig. 5. Estrus Detection Network.

subsequence of length k to predict the estrus state in the next time. Lastly, the continuous estrus factors will be used to determine whether the estrus alert should be sent in the Estrus Alert component.

V. Data set And Results

A. Data set

Data is collected from 40 dairy cows wearing neck-mounted sensors. During the data collection process, 6 cows shows estrus. Our method uses six different behavior characteristics of dairy cows, including the number of steps, minutes of dairy cows in high activity, medium activity, low activity, feeding, and rumination. Fig. 6 depicts the curve of the minutes in high activity, feeding, and rumination of a cow in the week before and after estrus. Fig. 7 shows the number of steps during the estrus period.

For the evaluation of the efficiency of the estrus detection model, the positive (estrus) and negative (not estrus) results were compared to the actual status. The performance of the network is measured by calculating the sensitivity towards the correct estrus alert, the specificity towards the precision of the estrus alert, and positive predictive value (PPV) towards the right estrus alert. These three evaluation criteria are formulated by the following formula: Sensitivity = TP/(TP + FN); specificity = TN/(TN + FP); PPV = TP/(TP + FP); where TP = true positive, TN stands for true negative, FP stands for false positive, and FN stands for false negative.

B. Results and Analysis

In this part, the results of one-step estrus prediction with different Lengths of Prefix Sequences (LoPE) are presented. It is observed that the sensitivity, specificity, and PPV of the estrus alert change with the different Number of Continuous Estrus Factors (NoCEF). The NoCEF represents how many consecutive positive prediction results are required to generate an estrus alert. For example, when the NoCEF is 3, an estrus alert is issued when there are two consecutive predictions result is positive, and the third detection result still needs to be positive.

C. Evaluation

Fig. 8 shows the result of the estrus alert sensitivity under four different NoCEF. Since they are influenced by the LoPE, the Prefix Size in the figure represents the length of prefix subsequence. It can be seen that as the LoPE increases,

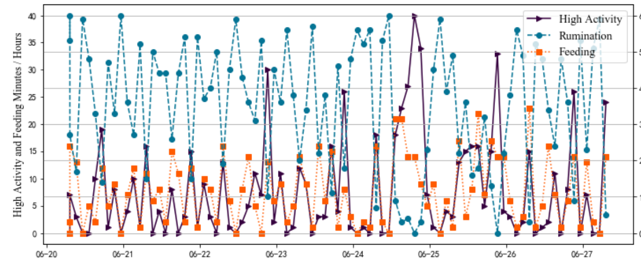

Fig. 6. Characteristics Curve.

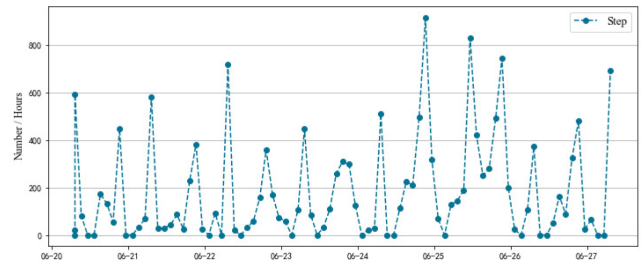

Fig. 7. Number of Steps Curve.

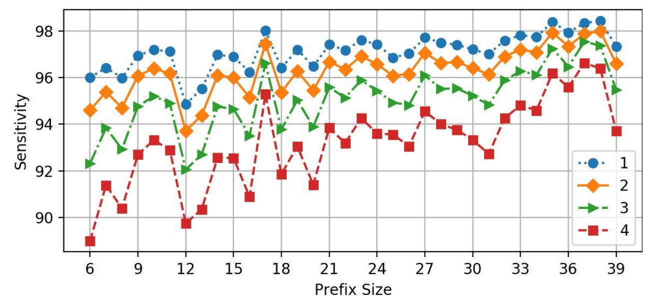

Fig. 8. Sensitivity Curve.

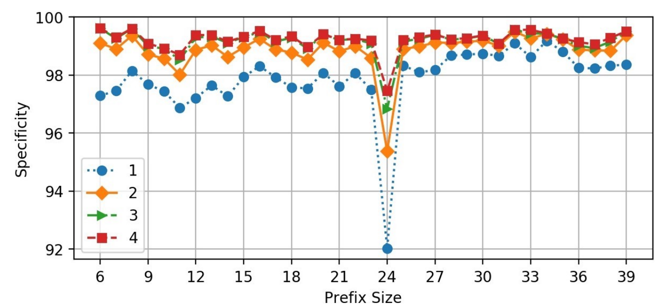

Fig. 9. Specificity Curve.

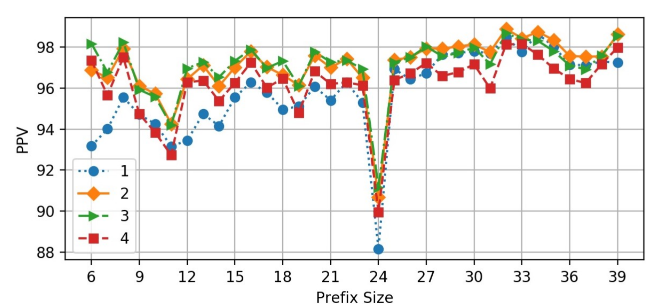

Fig. 10. PPV Curve.

TABLE III
COMPARISON OF DIFFERENT ALGORITHMS

	Sensitivity	*Specificity*	*PPV*
LSTM	94%	98%	95%
K-menas	85%	90%	75%

the sensitivity of estrus alert gradually increases with the oscillatory behavior. However, when LoPE reaches 38, the sensitivity decreases. At the same time, when the LoPE equals 17, the sensitivity reaches 98%, which is similar to the effect of LoPE equals 36. It shows that as the NoCEF increases, the sensitivity progressively decreases.

The results of the specificity of the estrus alert are compared in Fig. 9. As the LoPE increases, the specificity of the estrus alert does not change significantly, except at the point where LoPE equals 24. The figure shows that when LoPE equals 24, the prediction model generates faulty results that misdiagnose estrus negative as positive, resulting in a sharp decline in specificity. It implies that when choosing an appropriate LoPE to train the estrus detection network, LoPE 24 should be avoided.

Fig. 10 shows the PPV of estrus alert, which reflects the proportion of estrus alerts that are likely to be correct. Similar to Fig. 9, when the value of LoPE is 24, the PPV drops sharply in the case of the four different NoCEF values for the same reason mentioned above. Fig. 9 also shows that different NoCEF selections eventually achieve similar PPV as prefix size increases. As can be seen from the three figures, when NoCFE is 1, although the sensitivity of estrus alert is the highest, the specificity and PPV of the other two evaluation criteria are significantly lower than other NoCFE values. At the same time, when the LoPE reaches 35, the disparity in sensitivity under different NoCFE values becomes smaller. Therefore, in combination with the graphs of these three evaluation standards, NoCFE equals to 2 or 3, and the LoPE value between 35 and 38 is the most appropriate value for dairy cows estrus detection.

For a more comprehensive evaluation of performance, the K-means clustering algorithm [7] is also investigated. As shown in Table III, LSTM approach outperforms K-means model in sensitivity, specificity, and PPV. Since the increase in the amount of data helps to improve the accuracy of the model, it can be predicted that its performance will be better when used in larger dairy farms.

VI. CONCLUSION

In this paper, we presented an efficient cow estrus detection framework based on NB-IoT network, which can be applied to both captive breeding and graze breeding thanks to the stable performance of NB-IoT network in harsh outdoor environments. To support accurate and timely estrus detection, we proposed a learning-aided estrus data analysis strategy. First, the raw data collected by neck-mounted sensors are converted locally into multivariate behavioral time series data. Then the real time statistics are uploaded to cloud data center, where LSTM and CNN are exploited to capture the time correlation to improve accuracy of estrus detection. Simulation results showed that the performance of our proposed solution was better than most existing schemes in terms of accuracy and efficiency. In addition to estrus detection, the proposed solution can also be extended to realize efficient monitoring of cow's health, disease, and abortion.

VII. ACKNOWLEDGMENT

The authors would like to thank Mr. J. Peng with Beijing Ninty-Nine Harmony Technology Co. Ltd. for the help of collecting data from the pasture.

REFERENCES

[1] R. Foote, “Estrus detection and estrus detection aids,” *Journal of Dairy Science*, vol. 58, no. 2, pp. 248–256, 1975.
[2] Z. Xu, D. McKnight, R. Vishwanath, C. Pitt, and L. Burton, “Estrus detection using radiotelemetry or visual observation and tail painting for dairy cows on pasture,” *Journal of Dairy Science*, vol. 81, no. 11, pp. 2890–2896, 1998.
[3] D. N. Hossain, M. Talukder, M. K. Begum, and A. K. Paul, “Determination of factors that affect the pregnancy rate of cows after artificial insemination at Monirampur Upazila of Jessore District of Bangladesh,” *Journal of Embryo Transfer*, vol. 31, no. 4, pp. 349–353, 2016.
[4] D. M. Tsai and C. Y. Huang, “A motion and image analysis method for automatic detection of estrus and mating behavior in cattle,” *Computers and electronics in agriculture*, vol. 104, pp. 25–31, 2014.
[5] C. J. Yang, Y. H. Lin, and S. Y. Peng, “Develop a video monitoring system for dairy estrus detection at night,” in *2017 International Conference on Applied System Innovation (ICASI)*. IEEE, 2017, pp. 1900–1903.
[6] J. Li, J. Fang, Y. Fan, and C. Zhang, “Design on the monitoring system of physical characteristics of dairy cattle based on zigbee technology,” in *2010 World Automation Congress*. IEEE, 2010, pp. 63–66.
[7] M. S. Shahriar, D. Smith, A. Rahman, D. Henry, G. Bishop-Hurley, R. Rawnsley, M. Freeman, and J. Hills, “Heat event detection in dairy cows with collar sensors: An unsupervised machine learning approach,” in *2015 IEEE SENSORS*. IEEE, 2015, pp. 1–4.
[8] L. Yin, T. Hong, and C. Liu, “Estrus Detection in Dairy Cows from Acceleration Data using Self-learning Classification Models,” *JCP*, vol. 8, no. 10, pp. 2590–2597, 2013.
[9] K. Minegishi, B. J. Heins, and G. Pereira, “Peri-estrus activity and rumination time and its application to estrus prediction: Evidence from dairy herds under organic grazing and low-input conventional production,” *Livestock Science*, vol. 221, pp. 144–154, 2019.
[10] M. Borchers, Y. Chang, K. Proudfoot, B. Wadsworth, A. Stone, and J. Bewley, “Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle,” *Journal of dairy science*, vol. 100, no. 7, pp. 5664–5674, 2017.
[11] S. Shahinfar, D. Page, J. Guenther, V. Cabrera, P. Fricke, and K. Weigel, “Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms,” *Journal of dairy science*, vol. 97, no. 2, pp. 731–742, 2014.
[12] Z. C. Lipton, J. Berkowitz, and C. Elkan, “A critical review of recurrent neural networks for sequence learning,” *arXiv preprint arXiv:1506.00019*, 2015.
[13] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[14] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
[15] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.