

Research of Estrus Detection Models in Dairy Cows by Activity

Daoerji Fan^{*}, Huijuan Wu

College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China

Email address:

fandaoerji@126.com (Daoerji Fan), dream_meteor@126.com (Huijuan Wu)

^{*}Corresponding author

To cite this article:

Daoerji Fan, Huijuan Wu. Research of Estrus Detection Models in Dairy Cows by Activity. *Science Discovery*. Vol. 6, No. 2, 2018, pp. 102-109. doi: 10.11648/j.sd.20180602.15

Received: April 19, 2018; Accepted: May 17, 2018; Published: June 22, 2018

Abstract: The detection of estrus in large-scale ranch is an extremely labor-intensive task. Accurate and efficient detection can shorten the tires spacing of cow and increase economic income of ranch. The change in activity is one of the most important external features of dairy cows' estrus. In this paper, the activity of cows was collected remotely using activity collectors, wireless transmission networks and cloud servers, and the data was framed, labeled and preprocessed. The logistic regression, MLPs and SVMs model were trained by differential activity data of current and historical dates within the same hour. The experimental results show that the detection model by cow activity has a higher accuracy in predicting the estrus of cows. The SVMs in the three models have the best prediction effect. The recall and accuracy rate reach 90.12% and 93.74%, respectively. In the actual ranch test, the estrus detection system using the SVM model has more than double the exposure rate than the manual disclosure.

Keywords: Cows Estrus Prediction, Activity, Logistic Regression, Multilayer Perceptions, Support Vector Machine

基于活动量的奶牛发情预警模型研究

范道尔吉^{*}, 武慧娟

电子信息工程学院, 内蒙古大学, 呼和浩特, 中国

邮箱

fandaoerji@126.com (范道尔吉), dream_meteor@126.com (武慧娟)

摘要: 规模化牧场中发情检测是一项极其耗费人力的工作, 准确高效检测能够缩短胎间距, 增加经济效益。活动量变化是奶牛发情最主要外部特征之一。文中利用活动量采集器、无线传输网络和云服务器远程采集了奶牛活动量数据, 并对数据做了分帧、标注和预处理。用奶牛同时间段活动量差分作为训练数据, 训练了逻辑回归模型、多层感知器和SVMs模型。实验结果表明活动量建模预测奶牛发情有较高的准确率, 三个模型中SVMs有最好的预测效果, 召回率和准确率分别达到90.12%和93.74%。实际牧场使用中, 利用SVMs模型的发情检测系统的揭发率比人工揭发高出了1倍以上。

关键词: 奶牛发情预测, 活动量, 逻辑回归, 多层感知器, SVMs

1. 引言

规模化牧场中发情检测是一项极其耗费人力的工作, 准确高效检测能够缩短胎间距, 是增加经济效益的重要手

段。研究表明, 奶牛发情时活动量和体温有较明显的变化。1977年Kiddy测试发现93%的奶牛发情时活动量有400%的提升。1991年Schofield等人发现奶牛发情时其活动量比平

时有显著提高[1-6]。近几年中国学者也对奶牛运动特征和发情特征做了很多研究^[7-10]。

目前中国绝大多数养牛场仍采用人工观察, 依靠经验判断的方法进行奶牛发情识别。这种判断方法仅能用于小规模奶牛养殖场, 已不适用于较大规模、集约化的牧场。研究表明50%以上的奶牛在傍晚或夜间发情, 且部分隐性发情奶牛无发情的外部表现, 造成人工发情诊断的检出率较低。

信息技术和物联网技术的蓬勃发展为研究新型的奶牛行为检测技术奠定了基础。文献[11]中提出基于融合图像与运动量的奶牛行为识别方法, 并获得了交好的效果。文献[12]中提出了基于加速度传感器、Zigbee网络和二叉决策树支持向量机(Support Vector Machine,SVM)时间序列模型的奶牛个体行为检测系统, 为及早发现奶牛病情和发情提供了有力的支持。文献[13]中提出了基于声音传感器的奶牛反刍行为检测方法。实验结果表明, 该方法性能稳定, 可靠性较高, 精准度达90%, 可以有效的对奶牛反刍行为进行检测。

综上所述, 基于信息技术的奶牛行为检测对规模化养殖有重要意义, 提别是奶牛发情检测对提高牧场效益起到

关键作用。本文利用基于加速度传感器的奶牛可穿戴设备实时采集了活动量量化值, 并利用无线网络技术把活动量数据上传到后台服务器中。以活动量数据为基础分别提出了逻辑回归、多层感知器、SVMs等发情预测模型, 并对比了准确率。

2. 活动量采集系统

2.1. 系统框图

整个活动量数据采集系统如图1所示。系统由佩戴在奶牛脖子上的活动量采集器、活动量基站、以及后台服务器构成。活动量采集器和活动量基站间通过无线通信方式进行数据传输, 通信距离约100m(半径)。活动量基站通过GPRS网络与服务器连接。每个活动量采集器通过三轴加速度传感器检测运动加速度, 当加速度达到阈值后活动量计数加1。活动量采集器每隔1小时把计数值发送给基站、然后将计数值清零从新开始计数。基站实时把活动量数据打包后发送服务器, 服务器中的接收进程负责活动量数据的接收并存入数据库。



图1 系统框图。

2.2. 活动量采集

活动量采集器由低功耗单片机、三轴加速度传感器、和无线收发单元组成, 框图如图2所示。其中三轴加速度传感器和单片机通过SPI口进行连接, 无线收发模块通过串口和单片机连接。单片机每隔1/10秒采集一次三轴加速度传感器各个轴的数值 x,y,z , 并把计算结果 $x^2+y^2+z^2$ 存入数据缓存区中。缓存区可容纳30秒的数据。当缓存区满时, 从头到尾扫描缓存区一遍, 寻找加速度变化的波峰和波谷。计算所有波峰和波谷的差值并和固定阈值进行比较, 如果大于阈值则认为检测到一次有效运动, 小于则忽略。其中阈值是通过实验获得一个经验值。详细算法流程图如图3所示。

在很多研究中使用计步器作为奶牛步数的采集设备。但这里所提的活动量采集器和计步器有本质的区别。首先活动

量采集器是佩戴在奶牛脖颈位置, 而计步器通常佩戴在脚踝上。其次因为佩戴位置的不同, 活动量和步数没有相等的关系, 但奶牛行走时会引起活动量的增加。活动量的范围更加广泛, 它可能由行走、爬跨、采食、摇头等动作产生。所以活动量只是表示奶牛运动剧烈程度的量化值, 没有量纲。从设备使用角度来看, 活动量采集器更容易佩戴, 副作用小, 而计步器的佩戴比较繁琐, 容易对奶牛造成伤害。

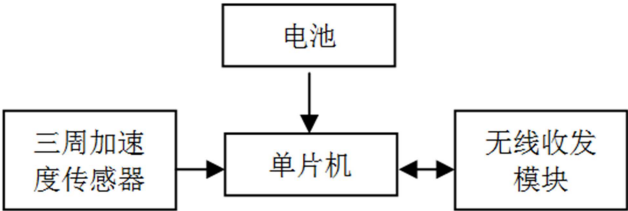


图2 活动量采集器框图。

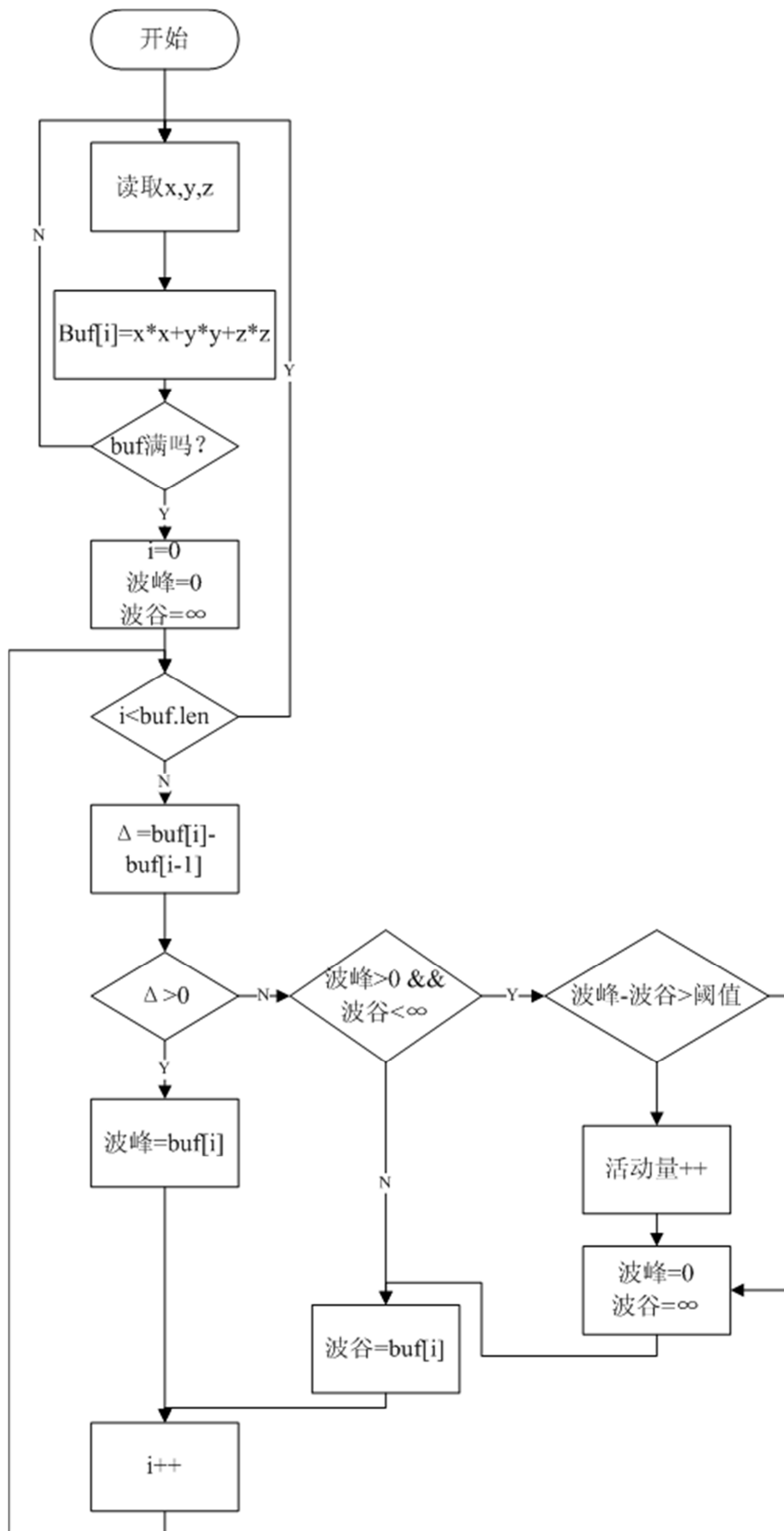


图3 活动量采集流程图。

为了获得充分的测试数, 从4个牧场选取60头奶牛佩戴活动量传感器, 连续检测2~6个月。其中发情状态通过两个途径获取: 一是牧场人工揭发报告, 二是活动量数据观察判断。奶牛的实时活动量和日平均活动量如图4、5所示, 其中纵坐标是活动量、横坐标是时、天。通过观察

日平均活动量可以很清晰的看到奶牛发情规律: 即发情间隔大概在21天左右。因此结合奶牛的发情规律和日平均活动量曲线可以判断出发情状态, 但该方法只适用于数据准备阶段, 不适合作为实时判断方法。结合图4和图5可以看出05:00~20:00可以看做是发情状态。

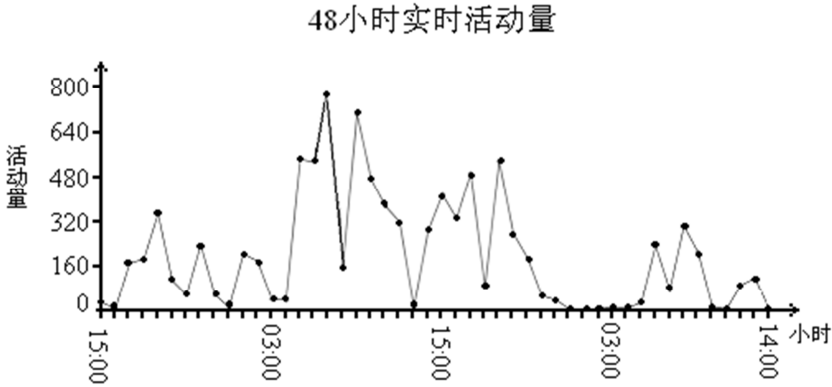


图4 实时活动量。

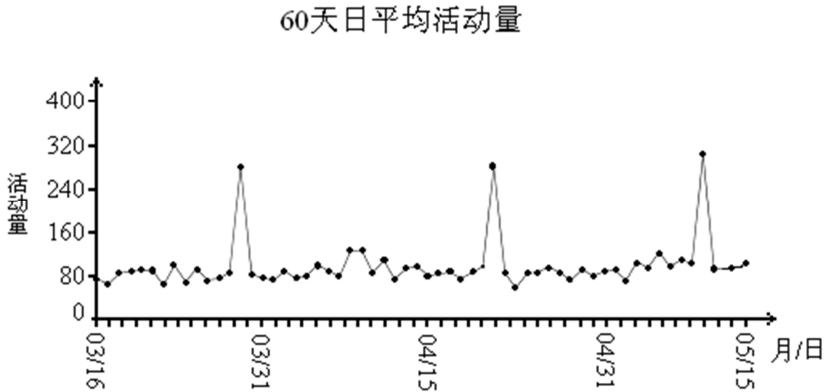


图5 日平均活动量。

通过上述方式对每个测试奶牛的数据做人工标记处理。把每小时数据用三元组 $data=(t, x, f)$ 表示, 式中 t —时间, 精确到小时; x —活动量; f —标记, “0”代表未发情、“1”代表发情。所有三元组按时间 t 的顺序排序, 并把每个奶牛数据保存成一个原始数据文件。

2.3. 特征向量及预处理

由于奶牛的个体差异和传感器的个体差异, 每头牛的活动量范围并不统一, 所以对每个奶牛的活动量各自按公式(1)做归一化处理。

$$\bar{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

用滑动窗口对活动量数据进行分帧处理, 其中窗口长度 L (单位: 小时), 帧移 ΔL (单位: 小时)。 L 和 ΔL 都是系统参数。通过实验 $L=8$ 、 $\Delta L=3$ 时获得较好效果。其中每一帧也用三元组 $frame=(t, x, f)$ 表示, 式中 t —记录每一帧的起始

时间; x —长度为 L 的向量; f —表示这一帧是不是发情状态, 用下面的方式计算:

$$\begin{aligned} \text{If } \text{count}(data.f==1) > L/2 + 1 \text{ then } & f=1 \\ \text{Else } & f=0 \end{aligned}$$

通过观察活动量变化发现奶牛发情时活动量的增高是相对于同一个时间段的活动量而言变化比较明显。用 $fram_0$ 代表当前帧, $fram_1$ 代表昨天同时间段的帧。最后分类用特征向量数据 X 和标记 y 定义为:

$$X = \begin{Bmatrix} fram_0.x \\ \Delta fram_{-1}.x \\ \Delta fram_{-2}.x \\ \vdots \\ \Delta fram_{-D}.x \end{Bmatrix} \text{ 其中 } \Delta fram_{-i}.x = fram_{-i}.x - fram_0.x \tag{2}$$

$$y = \{ fram_0.f \} \tag{3}$$

其中 D 也是系统参数, 代表差值天数。本文中选用 $D=5$, 所以 X 的每一行是48维向量。

3. 发情预测模型

3.1. 逻辑回归模型

统计学家D.R.Cox在1958年提出并发展了逻辑回归模型。二值逻辑回归模型中利用一个或多个特征数据预测“0”或“1”^[14,15]。设逻辑回归模型的输入数据 $x \in \mathbb{R}^{n+1}$, $x^{(i)}$ 是 $n+1$ 维向量(n 是原始数据维度, 前面固定加入一列‘1’), 输出数据 $y \in (0, 1)$ 。参数空间 $\theta \in \mathbb{R}^{n+1}$ 。定义预测函数、代价函数如公式(4), (5):

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}} \quad (4)$$

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)})) + \lambda \sum_{j=2}^n \theta_j^2 \quad (5)$$

其中 m 是训练数据样本个数, λ 是正则项参数。 θ 的估值问题就是最小化 $J(\theta)$ 的问题, 其中 x , y 是已知数据、 θ 看做变量, 如公式(6)所示。

$$\theta_{\text{estimate}} = \underset{\theta}{\operatorname{argmin}} J(\theta) \quad (6)$$

经过实验正则项参数 $\lambda=0.79$ 时获得了加好的效果。

3.2. 神经网络模型

采用含有一个隐含层的多层感知器神经网络模型对发情数据做分类处理。其中输入层包含48个神经元、隐含层5个、输出层2个。神经元的激发函数采用sigmoid函数, 如公式7所示。

$$\operatorname{sigmoid}(z) = \frac{1}{1 + e^{-z}} \quad (7)$$

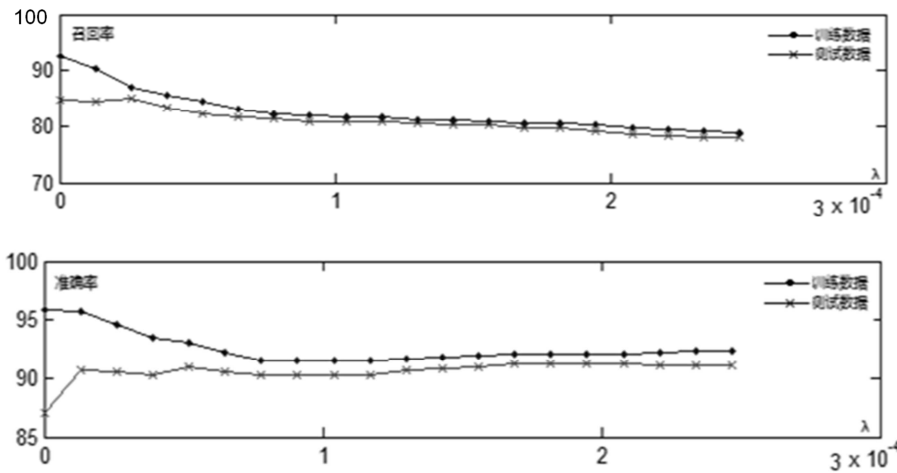


图6 λ 与召回率和准确率的关系。

3.3. SVMs模型

支持向量机SVMs(Support Vector Machines)1995年由Vapnik与其领导的贝尔实验室的研究小组一起开发出来的一种机器学习技术^[17]。训练数据集 D 定义为:

$$D = \{(x^{(i)}, y^{(i)}) | x^{(i)} \in \mathbb{R}^p, y^{(i)} \in \{-1, 1\}\}_{i=1}^n \quad (14)$$

对于线性SVMs分类问题可以转换为公式(15)所示的优化的问题:

定义 ℓ 层的神经元个数为 $d^{(\ell)}$ 、第 i 个神经元的输入为 $z_i^{(\ell)}$, 输出为 $a_i^{(\ell)}$, ℓ 层到 $\ell+1$ 层的权值矩阵定义为 $\theta^{(\ell)}$, 它是 $d^{(\ell+1)} \times d^{(\ell)} + 1$ 矩阵。输入为 x 时逐层计算输出为:

$$a^{(1)} = z^{(1)} = x \quad (8)$$

$$z^{(2)} = a^{(1)} \theta^{(1)T} a^{(2)} = \operatorname{sigmoid}(z^{(2)}) \quad (9)$$

$$z^{(3)} = a^{(2)} \theta^{(2)T} a^{(3)} = \operatorname{sigmoid}(z^{(3)}) \quad (10)$$

神经网络的代价函数定义为:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(a^{(3)}) + (1 - y^{(i)}) \log(1 - a^{(3)}) + \lambda \sum_{i=1}^{d^{(2)}} \sum_{j=2}^{d^{(1)}+1} \theta_{ij}^{(1)} + \lambda \sum_{i=1}^{d^{(3)}} \sum_{j=2}^{d^{(2)}+1} \theta_{ij}^{(2)} \quad (11)$$

和逻辑回归一样把 $J(\theta)$ 看做是 θ 的函数 x , y 是已知量, 参数估计问题就变成公式(6)中求 θ_{estimate} 的问题。为了获得函数最小值需要求解 $\partial J(\theta)/\partial \theta$, 这个问题可以用BP算法从输出层往回逐层计算梯度^[16]。 $\theta^{(1)}$, $\theta^{(2)}$ 的偏导数求解过程如公式(12), (13)所示。实验证明正则项参数 $\lambda=6.1558e-6$ 时发情帧的召回率(Recall)和准确率(Precision)比较理想。召回率、准确率和 λ 的关系如图6所示。

$$\frac{\partial J(\theta)}{\partial \theta^{(2)}} = \frac{\partial J(\theta)}{\partial a^{(3)}} \times \frac{\partial a^{(3)}}{\partial z^{(3)}} \times \frac{\partial z^{(3)}}{\partial \theta^{(2)}} \quad (12)$$

$$\frac{\partial J(\theta)}{\partial \theta^{(1)}} = \frac{\partial J(\theta)}{\partial a^{(3)}} \times \frac{\partial a^{(3)}}{\partial z^{(3)}} \times \frac{\partial z^{(3)}}{\partial a^{(2)}} \times \frac{\partial a^{(2)}}{\partial z^{(2)}} \times \frac{\partial z^{(2)}}{\partial \theta^{(1)}} \quad (13)$$

$$\theta = \arg \min_{\theta} C \sum_{i=1}^n \xi^{(i)} + \frac{1}{2} \sum_{j=2}^p \theta_j^2 \quad s.t. \{y^{(i)}(\theta^T x^{(i)}) \geq 1\}_{i=1}^n \quad (15)$$

其中 $\xi^{(i)}$ 是代价函数, 由 $y^{(i)}$ 和 $y^{(i)}_{predict}$ 计算; θ 是分类面的法向量也是要优化的参数; C 是可调参数, 调节代价函数和正则项所占比重[18]。非线性分类问题可以利用核函数。本研究中使用公式(16)所示径向基函数(RBF)获得了较

好的效果, 其中 σ 是核函数参数, 也是可调参数。通过实验 $C=4.26$ 和 $\sigma=2.35$ 时获得较好效果, C 、 σ 和召回率准确率的关系如图7、8所示。

$$K(x, x^{(i)}) = e^{-\frac{\|x-x^{(i)}\|^2}{2\sigma}} \quad (16)$$

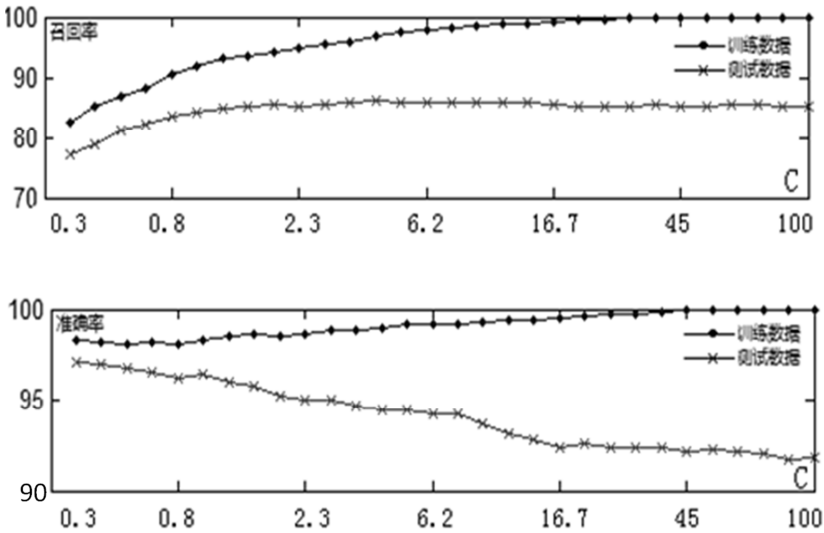


图7 C与召回率和准确率的关系。

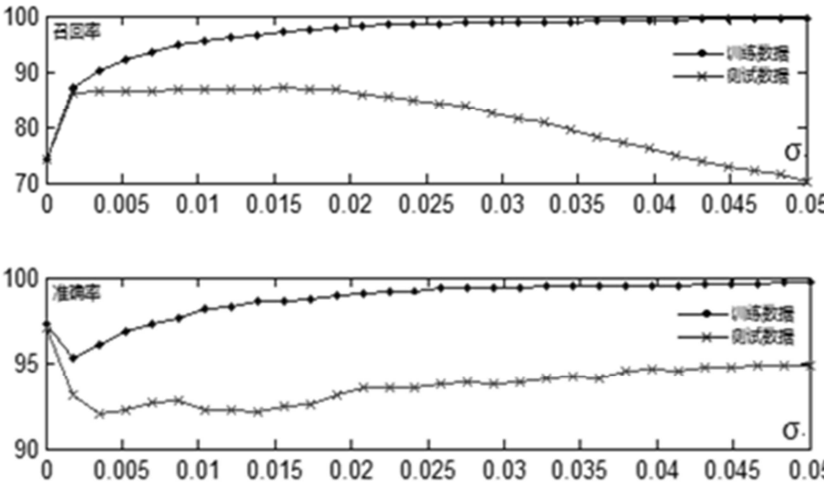


图8 sigma与召回率和准确率的关系。

4. 实验结果与分析

从四个合作牧场总共选取了60头奶牛作为实验牛群, 并佩戴活动量采集器。其中包括青年牛(未生牛犊)20头, 泌乳牛40头。泌乳牛中涵盖了2~4胎奶牛。实验连续采集了2~6个月的活动量数据, 并进行了人工标记, 其中共碰到122次发情, 每次发情大概持续8~16个小时。对发情数据用1小时帧移进行分帧处理后获得991个有效帧、未发情数据用3小时帧移进行分帧处理后获得37348个有效帧。用991个发情帧和10000个未发情帧(随机抽取1万个帧)构

成了原始数据, 发情用标记“1”表示, 未发情用标记“0”表示。

实验过程中, 使用了pmtk3工具在matlab环境中实现并验证了各个模型。利用K=4的交叉验证方法, 把原始数据分成4等份。其中用3份做为训练数据, 另一份做测试数据, matlab代码如下:

```
Kfold=4;
indices=crossvalind('Kfold',length(X),Kfold);
训练和测试数据准备:
train_data=X(indices~=1,:);
train_labels=y(indices~=1,:);
test_data=X(indices==1,:);
```



```
test_labels=y(indices==1,:);
逻辑回归模型训练:
lambda=0.001;
model=logregFit(train_data,train_labels,'lambda',0.7931);
[pred, prob]=logregPredict(model, train_data);
BP 神经网络训练:
lambda=6.1558e-006;
hidden_layer_size=5;
options.maxIter = 100;
options.display = 'off';
options.CostType = 1;
```

```
net=MPL(train_data,train_labels,hidden_layer_size,lambda,options);
pred=getForward_MPL(net,train_data);
[~,pred]=max(pred,[],2);
SVM 模型训练:
model=svmFit(train_data,train_labels,'kernelParam',0.0215,'C', 6.1054);
pred = svmPredict(model, train_data);
```

实验1中直接使用活动量数据, 实验结果如表1所示。
实验2中采用活动量差分, 实验结果如表2所示。

表1 直接利用活动量实验结果。

模型	训练集				测试集			
	帧个数	发情帧个数	召回率	准确率	帧个数	发情帧个数	召回率	准确率
SVMs	8243	743	95.23%	97.52%	2748	248	88.13%	93.25%
MLPs	8243	743	95.13%	98.16%	2748	248	88.99%	90.84%
LogRegr	8243	743	85.03%	92.08%	2748	248	84.23%	92.24%

表2 利用活动量差分实验结果。

模型	训练集				测试集			
	帧个数	发情帧个数	召回率	准确率	帧个数	发情帧个数	召回率	准确率
SVMs	8243	743	98.90%	99.60%	2748	248	90.12%	93.74%
MLPs	8243	743	95.12%	98.12%	2748	248	89.96%	92.81%
LogRegr	8243	743	86.72%	94.23%	2748	248	86.19%	93.34%

从实验结果看SVMs和神经网络模型要远优于逻辑回归模型, SVMs的召回率比神经网络稍微高一些、准确率高1%。活动量差分建模的效果普遍要比未差分的高1%~2%, 证明采用活动量差分建模是较好的选择。各个模型中召回率和准确率都是互相制约的参数。由于发情数据和未发情数据比例悬殊, 实验中准确率普遍高于召回率, 所以模型参数调节时按召回率最大化原则进行的。

为了验证本方案在实际牧场中的应用效果, 对合作牧场未孕的50头奶牛进行了3个月的测试。其中系统总共预报了73次发情预报(SVMs模型预测), 牧场配种员人工观察揭发了37次发情, 且全包含在系统预测结果中。对系统预报的每次发情牧场采用直肠检查法确定是否真实发情, 其中69次都是正确预报。从实际应用看利用活动量的发情揭发方式要比传统的人工揭发高出1倍以上。

5. 结论

本文中设计了一种基于加速度传感器的奶牛活动量采集设备。该设备本身不直接做发情判断, 而是把数据按一小时间隔上传到服务器端。服务器端利用强大的存储能力和计算能力进行奶牛发情预警判断。实验证明根据活动量变化预测奶牛发情是一种简单有效的方法。利用同时时间段活动量差分做发情预测效果要优于直接使用活动量的预测方式。各个模型中采用8小时帧长时获得最好的预测效果。逻辑回归、神经网络、支持向量机三个模型中支持向量机的性能最好。本文中提出的活动量采集器安装在奶牛脖颈位置, 便于安装和维护, 对奶牛健康影响较小。本文的研究成果为研发实用的基于活动量的发情检测系统提供了坚实的基础。

致谢

本研究由内蒙古高等学校科学研究项目(NJZC16021, NJZY17010)资助, 特此感谢!

参考文献

- [1] Kiddy C A. Variation in physical activity as an indicator of oestrus in dairy cows[J]. Dairy Sci, 1977, 60(4): 235~243.
- [2] Schofield S A, Phillips C J C, Owens A R. Variation in milk production activity rate and electrical impedance of cervical mucus over the oestrus period of dairy cows [J]. Anim. Reprod. Sci, 1991, 24(2): 231~248.
- [3] Firk R, Stamer E, Junge W et al. Automation of oestrus detection in dairy cows: a review [J]. Livestock Production Science, 2002, 75(3): 219~232.
- [4] Homer EM, Derecka K, Webb R et al. Mutations in genes involved in oestrous cycle associated expression of oestrus [J]. Animal Reproduction Science, 2013, 142(3): 106~112.
- [5] Audrey C, Stéphanie C, Julie G et al. Comparison of three devices for the automated detection of estrus in dairy cows [J]. Theriogenology, 2014, 82(5): 734~741.
- [6] Reith S, Brandt H, Hoy S. Simultaneous analysis of activity and rumination time, based on collar-mounted sensor technology, of dairy cows over the peri-estrus period [J]. Livestock Science, 2014, 170(1): 219~227.

- [7] 田富洋, 王冉冉, 刘莫尘, 等. 基于神经网络的奶牛发情行为辨识与预测研究[J]. 农业机械学报, 2013, 44(z1):277~281。
- [8] 尹令, 洪添胜, 刘迎湖, 等. 基于无线传感器网络支持向量机奶牛行为特征识别[J]. 传感技术学报, 2011, 24(3):458~462。
- [9] 尹令, 刘财兴, 洪添胜, 等. 基于无线传感器网络的奶牛行为特征监测系统的设计[J]. 农业工程学报, 2010, 26(3):203~208。
- [10] 赵凯旋, 何东健, 王恩泽. 基于视频分析的奶牛呼吸频率与异常检测[J]. 农业机械学报, 2014, 45(10):258~263。
- [11] 顾静秋, 王志海, 高荣华, 等. 基于融合图像与运动量的奶牛行为识别方法[J]. 农业机械学报, 2017, 48(6):145-151。
- [12] 侯云涛, 蔡晓华, 吴泽全, 等. 奶牛行为特征识别方法的研究与实现—基于支持向量机[J]. 农机化研究, 2018, 40(8):36-41。
- [13] 张爽. 奶牛个体反刍行为监测技术研究[D]. 东北农业大学, 2017。
- [14] Cox D R. The regression analysis of binary sequences [J]. J Roy Stat Soc B, 1958, 20(3): 215~242.
- [15] Walker S H, Duncan D B. Estimation of the probability of an event as a function of several independent variables [J]. Biometrika, 1967, 54(1):167~178.
- [16] Werbos P J. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences [D]. Phd thesis, Harvard University, 1974.
- [17] Cortes C, Vapnik V. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273~275.
- [18] Collobert R, Bengio S. Links between Perceptrons, MLPs and SVMs [C]. New York: ICML, 2004:23~30.