

融合心理行为的家政服务员专业性分类研究

冯永杰, 刘峰涛

东华大学旭日工商管理学院, 上海

收稿日期: 2022年12月6日; 录用日期: 2023年1月6日; 发布日期: 2023年1月16日

摘要

本文为进行家政服务员专业性分类研究, 设计了心理行为量表进行问卷调查, 并基于Y企业家政服务员的丰富的数据资源, 进行了心理因素与行为属性的相关性分析, 确定了影响家政服务员专业性的心理行为属性, 进而提出了家政服务员专业性分类模型, 并采用过采样方法平衡训练数据集进行分类模型训练, 针对测试集进行家政服务员专业性分类预测。研究结果表明, 心理因素与行为属性显著相关, 确立了影响家政服务员专业性的若干心理行为属性, 基于心理行为属性平衡训练数据集建立的分类算法对测试集中家政服务员专业性分类准确率达到79.8%。

关键词

家政服务, XGBoost算法, 心理行为属性, 大数据分类

Study on the Professional Classification of Domestic Staff Integrating Psychological Behavior

Yongjie Feng, Fengtao Liu

Xuri School of Business Administration, Donghua University, Shanghai

Received: Dec. 6th, 2022; accepted: Jan. 6th, 2023; published: Jan. 16th, 2023

Abstract

In order to study the professional classification of domestic staff, this paper designs a psychological behavior scale to conduct a questionnaire survey. Based on the abundant data resources of domestic staff in Y Enterprise, this paper analyzes the correlation between psychological factors and behavioral attributes, determines the psychological behavior attributes that affect the professionalism of domestic staff, and then proposes a professional classification model for domestic staff.

The over-sampling method is used to balance the training data set for classification model training, and the professional classification prediction of domestic staff is carried out according to the test set. The results show that there is a significant correlation between psychological factors and behavioral attributes, and a number of psychological behavioral attributes affecting the professionalism of domestic staff are established. The classification algorithm based on the balance training data set of psychological behavioral attributes can achieve 79.8% accuracy in the professional classification of domestic staff in the test set.

Keywords

Domestic Service, XGBoost Algorithm, Psychological Behavior Attributes, Big Data Classification

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着经济的发展,人们的生活水平不断提高,经济社会对家政服务员的的需求增大,越来越多的劳动力投身家政行业,但是随之而来的是家政服务员的技能不齐整,整体素质和能力参差不齐,缺乏培训,社会认可度低,家政服务一度陷入价高质次的结构性失衡困境。

解决存在的结构性失衡问题,在于划分专业标准,培养专业的家政服务人员,对家政服务员分级培训,分类型培训[1]。本文通过设计心理量表对家政服务员进行心理因素与行为属性的相关性分析,选取影响家政服务员专业性的心理行为属性,运用大数据分类算法对家政服务员专业性进行分类,对家政服务员的专业性进行定级考量研究,为家政服务领域提供识别家政服务员专业性的新的技术和理论标准。

本文第一节进行心理因素与行为属性的相关性分析,确定相关的心理行为属性,第二节构建基于心理行为属性的大数据算法分类模型,分析家政服务员专业性分类准确率,最后验证并比较分类实验结果。

2. 基于心理行为测度的行为属性确定

2.1. 调研对象

Y 企业在职家政服务人员。

2.2. 实验方法

基于职业倦怠 - 服务行业版(MBI-HSS)心理量表的情绪耗竭、去人性化、个人成就感三个维度共设计了 8 个问题[2]; 基于护士职业认同量表设计了职业认同心理因素下自我认同、家人认同、社会认同的 8 个问题[3]; 基于心理契约量表从店长、发展、薪酬、决策 4 个维度设计了 12 个问题[4]。分析心理因素与行为属性之间的相关关系。

1) 抽样方法

本文运用整群抽样的方法抽取 Y 企业 1~3 星的家政服务员 230 人、292 人、135 人。

2) 计分方法

心理契约、职业认同因素问卷采用 Likert-5 分量表法进行计分(1 = 不同意, 2 = 有点不同意, 3 = 一般, 4 = 有点同意, 5 = 同意), 在题目设置上会采用反向统计[5]。职业倦怠因素同样采用 Likert-5 量表进行计分(1 = 从来没有这样, 2 = 很少这样, 3 = 有时候这样, 4 = 经常这样, 5 = 几乎每天都这样),

在题目设置上同样会采用反向统计[6]。最终总分数越高表示家政服务员的心理状态越好。

2.3. 统计分析

运用 SPSS 26.0 软件进行统计分析, 计数占比用百分比(%)表示, 以表示家政服务员在各个心理维度的得分, 进行检验。心理因素与行为属性相关性采用皮尔逊相关系数进行衡量, 在 $p < 0.05$ 或 $p < 0.01$ 上有显著性差异具有统计学意义。

2.4. 调研结果

1) 实验信度分析

整体问卷的克隆巴赫系数为 0.845, 大于 0.8, 说明问卷的可靠性较高; 该问卷修正后的项与总计相关性越高表示该题与其他题项之间的一致性越高, 一般大于 0.4 为佳, 18、20、21、22、24、25 题修正后的项与总计相关性远小于 0.4, 需要进行优化; 对每一项题目进行删除后得到新的克隆巴赫系数, 18、20、21、22、24、25 题项删除后的系数大于 0.845, 问卷的整体信度获得提升。

2) 基本信息分析

问卷调查家政服务人员数为 657 名, 其中男家政服务员 104 人, 女家政服务员 553 人, 各占 15.8%, 84.2%; 1~3 星级的家政服务员分别为 230 人, 292 人, 135 人。依据受教育程度(小学及以下、初中、中专、高中、大专、本科及以上)从低到高的占比分别为 13.1%、67.0%、6.4%、9.7%、2.7%、1.1%, 详情见表 1。

Table 1. Basic information table of respondents

表 1. 调查对象基本情况表

调查目录	人员数量	占比
性别		
男	104	15.8%
女	553	84.2%
星级		
1	230	35.0%
2	292	44.5%
3	135	20.5%
学历		
小学及以下	86	13.1%
初中	440	67.0%
中专	42	6.4%
高中	64	9.7%
大专	18	2.7%
本科及以上	7	1.1%

2.5. 实验分析

2.5.1. 心理因素分析

1) 职业认同心理维度分析

家政服务员在职业认同心理量表上的总平均分为 4.555 ± 0.845 , 其中家人认同、社会认同的得分较

高, 自我认同的得分较低, 详情见表 2。

2) 心理契约心理维度分析

家政服务员在心理契约心理量表上的总平均分为 4.187 ± 1.127 , 其中店长维度、发展维度的得分较高, 薪酬维度、决策维度的得分较低, 详情见表 3。

3) 职业倦怠心理维度分析

家政服务员在职业倦怠心理量表上的总平均分为 2.885 ± 1.145 , 其中个人成就感维度的得分较高, 情绪耗竭、去人性化维度的得分较低, 详情见表 4。

Table 2. Score of psychological dimension of professional identity
表 2. 职业认同心理维度得分

心理维度	得分
自我认同	4.491 ± 0.888
家人认同	4.613 ± 0.846
社会认同	4.561 ± 0.803

Table 3. Score of psychological dimension of psychological contract
表 3. 心理契约心理维度得分

心理维度	得分
店长	4.621 ± 0.824
发展	4.436 ± 1.008
薪酬	3.953 ± 1.268
决策	3.736 ± 1.408

Table 4. Score of psychological dimension of job burnout
表 4. 职业倦怠心理维度得分

心理维度	得分
情绪耗竭	2.649 ± 1.172
去人性化	2.166 ± 1.161
个人成就感	3.839 ± 1.100

2.5.2. 心理行为属性相关性分析

1) 职业认同心理因素与行为属性的相关分析

职业认同与技能数量的相关系数为 0.101, 职业认同与客户是否继续接受服务的相关系数为 0.108, 并在 $p < 0.01$ 的级别上呈现显著的相关性。职业认同与总评分的相关系数为 0.094, 并在 $p < 0.05$ 的级别上呈现显著的相关性。结果表明, 职业认同与行为属性之间呈现显著的正相关关系, 表明职业认同感高的家政服务员会对行为产生正向的影响, 详情见表 5。

2) 心理契约心理因素与行为属性的相关分析

心理契约与培训次数的相关系数为 0.106, 并在 $p < 0.01$ 的级别上呈现显著的相关性。职业认同与月度合作费的相关系数为 0.090, 并在 $p < 0.05$ 的级别上呈现显著的相关性。结果表明, 心理契约与行为属性之间呈现显著的正相关关系。心理契约感强的家政服务员更愿意参加培训提升自己的专业能力, 获得更多的报酬, 详情见表 6。

3) 职业倦怠心理因素与行为属性的相关分析

职业倦怠与请假次数的相关系数为 0.096, 并在 $p < 0.05$ 的级别上呈现显著的相关性。职业倦怠与是否准时的相关系数为-0.116, 职业倦怠与订单量的相关系数为-0.117, 并在 $p < 0.01$ 的级别上呈现显著的相关性。职业倦怠与准时到达、订单量呈现显著的负相关, 结果表明, 强职业倦怠感会对家政服务员的行为表现产生消极的影响, 详情见表 7。

Table 5. Correlation analysis between professional identity and behavioral attributes

表 5. 职业认同与行为属性的相关分析

	技能数量	是否被推荐	招聘来源	总评分	服务质量评分	服务态度评分	客户是否继续接受服务
职业认同	0.101**	0.093*	0.108**	0.094*	0.098*	0.092*	0.108**

** : 在 0.01 级别(双尾), 相关性显著。* : 在 0.05 级别(双尾), 相关性显著。

Table 6. Correlation analysis between psychological contract and behavioral attributes

表 6. 心理契约与行为属性的相关分析

	店长信息	培训次数	月度合作费
心理契约	0.190**	0.106**	0.090*

** : 在 0.01 级别(双尾), 相关性显著。* : 在 0.05 级别(双尾), 相关性显著。

Table 7. Correlation analysis between job burnout and behavioral attributes

表 7. 职业倦怠与行为属性的相关分析

	请假次数	是否准时	订单量
职业倦怠	0.096*	-0.116**	-0.117**

** : 在 0.01 级别(双尾), 相关性显著。* : 在 0.05 级别(双尾), 相关性显著。

基于以上心理行为实验研究可知, 确定影响家政服务员专业性的心理行为属性为: 技能数量、是否被推荐、招聘来源、总评分、服务质量评分、服务态度评分、客户是否继续接受服务、店长信息、培训次数、月度合作费、请假次数、是否准时、订单量。

3. 融合心理行为的家政服务员专业性分类模型构建

3.1. 大数据分类算法流程构建

针对有监督类型的 XGBoost 分类算法的分类特征[7], 将家政服务员的星级作为标签, 心理行为属性以及基础属性作为分类属性基准点, 以 7:3 的比例将数据库数据集分为训练集和测试集, 平衡训练数据集建立家政服务员专业性分类模型。

1) XGBoost 模型定义[8]

XGBoost 是一种有监督的学习算法, 由多棵决策树构成。假设存在数据集 $T = \{(x_i, y_i)\}$, 其中 $x_i \in D^m$, $y_i \in D$, $i = 1, 2, \dots, n$ 。 x_i 代表属性向量, 维度为 m , y_i 代表样本的标签值。假设模型由 T 棵决策树组成, 则模型的目标函数定义为:

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f(x_i) \quad (1)$$

其中, y_i 表示经第 t 轮迭代后的模型预测值, y_i 表示已知 $t-1$ 个基础模型的预测值, $f_t(x_i)$ 表示第 t 个决策树。每棵决策树都根据属性划分样本类别, 使得各样本被分到决策树的叶子节点上, 某个叶子节点的权

重 w 即为样本分到此叶子节点上时在此决策树的预测值。将样本在所有叶子节点的权重加和, 即为在 XGBoost 模型上的预测值。

为避免模型复杂度增加, 防止出现过拟合等问题, XGBoost 模型的目标函数为:

$$Obj^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t)}) + \sum_{j=1}^t \Omega(f_j) \quad (2)$$

第一项为损失函数, 用于表示样本真实值与模型预测值之差。 $\Omega(f_t)$ 为第 j 个基础模型的正则项。

$\Omega(f_t)$ 的定义如下:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (3)$$

其中, T 表示叶子节点的个数, $\frac{1}{2} \lambda \sum_{j=1}^T w_j^2$ 表示 L_2 正则项。来控制叶子节点的权重, 保证各个叶子节点的重要性不会相差太大。 CART 树生长得越复杂, 对应的 T 越大, $\Omega(f_t)$ 也越大。

2) 目标函数近似估计

由于很难使用传统的方法对目标函数进行优化, XGBoost 使用近似估计的方法优化目标函数:

$$Obj^{(t)} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \sum_{j=1}^t \Omega(f_j) \quad (4)$$

$\hat{y}_i^{(t-1)}$ 表示样本 x_i 在 $t-1$ 个基础模型上的预测值, $f_t(x_i)$ 表示第 t 个训练的决策树。

XGBoost 模型用泰勒公式对损失函数做近似估计, 泰勒公式可以用一个点的函数值表示一个这个点附近的函数值, 它用该点的各阶导数构建一个多项式来对该点附近的函数值做近似。

泰勒展开式为:

$$f(x + \Delta x) \approx f(x) + f'(x) \Delta x + \frac{1}{2} f''(x) \Delta x^2 \quad (5)$$

其中, $f(x)$ 是一个具有二阶可导的函数, $f'(x)$ 为 $f(x)$ 的一阶导函数, $f''(x)$ 为 $f(x)$ 的二阶导函数, Δx 为 $f(x)$ 在某点处的变化量。假设令损失函数 L 为泰勒公式中的 f , 令损失函数中 $\hat{y}_i^{(t-1)}$ 项为泰勒公式中的 x , 令损失函数中 $f_t(x_i)$ 项为泰勒公式中的 Δx , 则目标函数 $Obj^{(t)}$ 可以近似表示为:

$$Obj^{(t)} \approx \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 + \Omega(f_t) \quad (6)$$

g_i 和 h_i 分别是损失函数 $L(y_i, \hat{y}_i^{(t-1)})$ 关于 $\hat{y}_i^{(t-1)}$ 的一阶导函数值和二阶导函数值, 即他们可以表示为:

$$\begin{cases} g_i = \frac{\partial L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}} \\ h_i = \frac{\partial^2 L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)^2}} \end{cases} \quad (7)$$

常数项不会对结果产生影响, 去掉常数项 $L(y_i, \hat{y}_i^{(t-1)})$, 并将 $\Omega(f_t)$ 的表达式带入目标函数, 则目标函数为:

$$Obj^{(t)} \approx \sum_{i=1}^n \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (8)$$

3) 最优目标函数和叶子权重

$f_t(x_i)$ 表示第 t 棵树, 是 XGBoost 模型的一个基础模型, w_j 表示叶子节点的权重, 进一步优化目标函数可得:

$$Obj^{(t)} \approx \sum_{i=1}^n \left(g_i w_{q(x_i)} + \frac{1}{2} h_i w_{q(x_i)}^2 \right) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (9)$$

$w_{q(x_i)}$ 是第 t 棵树中 1 到 T 中的某个叶子节点的权重值。

采用叶子节点遍历, 计算损失函数, 可得损失函数为:

$$Obj^{(t)} \approx \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} (h_i + \lambda) \right) w_j^2 \right] + \gamma T \quad (10)$$

其中, I_j 是第 j 个叶子节点的样本集, 第 t 个树模型 $f_t(x_i)$ 将样本分到叶子节点, 则其在该叶子节点的权重为 w , 所以当 i 在 I_j 中时, 可以用 w_j 代替 $f_t(x_i)$ 。

令 $G_j = \sum_{i \in I_j} g_i$; $H_j = \sum_{i \in I_j} h_i$ 。它们分别表示所有属于叶子节点 j 的样本点对应的 g_i 之和以及 h_i 之和。所以目标函数是寻找一个合理的 f_i , 使得式子 $\sum_{j=1}^T \left(G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right) + \gamma T$ 尽可能大地减小。对其求偏导数, 令导函数为 0, 可得:

$$w_j = -\frac{G_j}{H_j + \lambda} \quad (11)$$

将 w_j 的值导入到目标函数 $Obj^{(t)}$ 中, 可得:

$$J(f_i) = -\frac{1}{2} \sum_{j=1}^T \left(\frac{G_j^2}{H_j + \lambda} \right) + \gamma T \quad (12)$$

容易计算出分裂前后损失函数的差值为:

$$Gain = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} - \gamma \quad (13)$$

其中, G 和 H 分别代表当前节点的一阶导函数值和二阶导函数值的和, G_R 和 G_L 代表在当前节点划分出的右子节点和左子节点样本集的一阶导函数值的和; H_R 和 H_L 分别表示在当前节点划分出的右子节点和左子节点样本集的二阶导函数值的和。

XGBoost 依据分裂前后损失函数相差最大的那个特征进行首次分裂, 依次进行。而分裂前后的损失函数差值一定是正的, 这也起到了预剪枝的作用。

3.2. 实验分析

本文基于学者研究的学历、婚姻、性别、年龄、户籍的基础上融合心理行为属性进行家政服务员专业性分类预测。

1) 数据预处理

通过对 Y 企业数据进行清洗和分析, 对缺失值和异常值进行处理之后, 符合家政服务员专业性的相关属性和星级数据共 1239 个样本。按照 7:3 的比例将样本分为训练数据集和测试数据集。

婚姻状况未婚、已婚、离异分别表示为 1、2、3。性别 1 代表男, 0 代表女。学历划分为小学及以下为 1、初中为 2、中专为 3、高中为 4、大专为 5、本科及以上为 6。籍贯取身份证前 2 位数字。星级 1~3 代表家政服务员的 3 个星级, 1 为最低星级, 3 为最高星级, 详情见表 8。

Table 8. Examples of partial sample processing results
表 8. 部分样本处理结果示例

id	性别	学历	来源	婚姻	户籍	推荐人	门店	技能数量	请假次数
1	0	0	4	1	34	0	62	19	2
2	0	0	4	1	34	1	62	19	2
3	0	1	9	1	41	0	73	22	4
4	0	1	9	1	41	0	73	22	4
5	0	1	9	1	41	1	73	22	4
6	1	4	4	1	13	0	25	22	0
7	0	0	4	1	34	0	60	10	11
8	0	0	1	1	34	1	25	13	0
9	1	2	12	1	34	0	27	21	0
10	1	2	12	1	34	0	27	21	0

2) 分类结果分析

家政服务员数据集中 1 星级、2 星级、3 星级 3 个类别数据分布不平衡, 在分类时会造成模型偏向占比较大的类别, 为此要进行数据集的平衡化处理。首先输入未经平衡化的训练集, 采用默认参数建立融合心理行为属性的 XGBoost 算法的家政服务员专业性分类模型作为参照, 然后引入 SMOTE 算法对训练集数据进行平衡后, 采用调优后的参数建立融合心理行为属性的 XGBoost 算法的家政服务员专业性分类模型。对比测试数据集在这两种分类模型上的预测结果, 详情见表 9。

Table 9. Comparison of the effect of classification models of balanced data and unbalanced data
表 9. 平衡数据与非平衡数据分类模型效果对比

分类模型	星级	Precision	Recall	F1-score	准确率
平衡数据分类模型	1	0.84	0.86	0.85	79.8%
	2	0.80	0.75	0.77	
	3	0.42	0.50	0.45	
非平衡数据分类模型	1	0.84	0.87	0.86	80.4%
	2	0.78	0.76	0.77	
	3	0.47	0.40	0.43	

平衡训练数据集之后建立的家政服务员专业性分类模型, 命名为模型 1。测试集的分类准确率为 79.8%。非平衡训练数据集建立的家政服务员专业性分类模型, 命名为模型 2, 测试集的分类准确率为 80.4%。模型 1 与模型 2 相比, 在 1 星级上的召回率降低了 1 个百分点, 在 2 星级上的召回率降低了 1 个百分点; 但是在 3 星级上的召回率提高了 10 个百分点。而且 3 星级的精确率降低了 5 个百分点, 说明平衡数据之后, 模型可以更好地识别样本量占比较小的 3 星级样本, 但是也把其他星级的样本识别为 3 星级样本。由此可见, 模型分类召回率的提升是以降低分类精确率为代价的。模型 1 的 F1-score 高于模型 2, 表明模型 1 的分类性能优于模型 2。

4. 研究结论

本文设计心理量表进行调查实验, 并依据互联网家政企业的丰富的数据资源, 进行心理因素与行为

属性的相关性分析, 确定影响家政服务员专业性的心理行为属性, 在对家政服务员不平衡级别的数据基础上, 平衡训练数据集建立家政服务员专业性分类模型, 并且利用测试集数据对比平衡数据前后建立的家政服务员专业性分类模型的准确率, 以下是研究结论:

1) 对心理因素与行为属性进行相关性分析, 确定影响家政服务员专业性的心理行为属性包括技能数量、是否被推荐、招聘来源、总评分、服务质量评分、服务态度评分、客户是否继续接受服务、店长信息、培训次数、月度合作费、请假次数、是否准时、订单量。通过访谈发现, 基层资深管理者依据经验进行家政服务员专业性判别, 本文依据心理量表实验, 从统计学角度确定影响家政服务员专业性的心理行为属性。

2) 在融合心理行为属性的 XGBoost 算法中, 平衡训练数据集建立的分类模型中测试集数据的分类准确率达到 79.8%, 并且在占比较小的类别上有很好的分类表现, 大数据分类算法在家政服务员专业性判别上有较好的应用效果, 为企业建立专业标准、培养专业人才提供了新的思路。

参考文献

- [1] 苏国明, 王超超, 徐红, 吴秀荣, 张景春. 多元需求背景下家政服务业人才队伍建设探究[J]. 职业技术, 2022, 21(10): 53-59.
- [2] 于丽玲, 吕兆丰, 王晓燕, 等. 乡村医生的医疗卫生服务行为现状及心理因素分析[J]. 中国全科医学, 2011, 14(8): 881-883.
- [3] 徐启慧, 陈翠萍, 贺学敏, 杜康泰. 基于 CiteSpace 的我国护理职业认同相关研究的可视化分析[J]. 职业与健康, 2022, 38(23): 3282-3287.
- [4] 蔡敏, 季子凌, 俞国桥. 心理契约与制造业新生代员工离职倾向——职业前景感知的调节作用[J]. 杭州电子科技大学学报(社会科学版), 2022, 18(5): 18-24+40.
- [5] 陈肇第. 社会服务传递视角下社会工作专业性研究——以武汉市 H 区青少年暑期托管工作为例[D]: [硕士学位论文]. 武汉: 华中师范大学, 2017.
- [6] 李园园, 鄢超云, 韦嘉. 幼儿园教师认知共情对职业倦怠的影响[J]. 中国健康教育, 2022, 38(7): 613-616+626.
- [7] Patnaik, B., Mishra, M., Bansal, R.C. and Jena, R.K. (2021) MODWT-XGBoost Based Smart Energy Solution for Fault Detection and Classification in a Smart Microgrid. *Applied Energy*, **285**, Article ID: 116457. <https://doi.org/10.1016/j.apenergy.2021.116457>
- [8] Gertz, M., et al. (2020) Using the XGBoost Algorithm to Classify Neck and Leg Activity Sensor Data Using On-Farm Health Recordings for Locomotor-Associated Diseases. *Computers and Electronics in Agriculture*, **173**, Article ID: 105404. <https://doi.org/10.1016/j.compag.2020.105404>