

기계학습을 이용한 비육돈의 비율일당증체분석

이용섭* · 황세운* · 김종현*

*국립경상대학교

Analysis on Proportional Daily Weight Increase of Swine Using Machine Learning

Woongsup Lee* · Sewoon Hwang* · Jonghyun Kim*

*Gyeongsang National University

E-mail : wslee@gnu.ac.kr

요 약

최근 기계학습기반 빅데이터 분석이 많은 주목을 받으며 축산분야에도 다양한 기계학습 방안들이 적용되고 있다. 많은 양의 생체데이터와 환경데이터에 기계학습기법을 적용하여 분석함으로써 이전에는 발견하지 못했던 새로운 사실들을 밝혀낼 수 있다. 본 연구에서는 실험돈사에서 수집된 데이터를 기반으로 비육돈의 비율일당증체를 예측하는 방안에 대해서 다루었다. 비율일당증체 예측을 위해서 매일 측정된 온도/습도/이산화탄소/풍속의 일평균/일최소/일최대 데이터(환경데이터)과 비율일당증체(생체데이터)를 이용하였고, 트리기반 알고리즘을 사용하여 비율일당증체 예측수식을 도출하였다. 이를 통해 평균 온도가 비율일당증체에 가장 큰 영향을 미치는 요소인 것을 보였다. 본 연구의 결과는 각 양돈농가에서 비육돈의 성장을 예측하는데 유용하게 활용될 수 있을 것이다.

ABSTRACT

Recently, big data analysis based on machine learning has gained popularity and many machine learning techniques have been applied to the field of agriculture. By using machine learning technique to analyze huge number of samples of biological and environmental data, new observations can be found. In this research, we consider the estimation of proportional daily weight increase (PDWI) based on measurement data from experimental swine farm. In order to derive the exact formulation for PDWI estimation, we have used measured value of mean, daily maximum, daily minimum of temperature, humidity, CO₂, wind speed and measured PDWI values. Based on collected data, we have derived equation for PDWI estimation using tree-based algorithm. In the derived formulation, we have found that the daily average temperature is the most dominant factor that affects PDWI. Our results can be applied to pig farms to estimate the PDWI of swine.

키워드

양돈, 일당증체, 기계학습, 트리 알고리즘

1. 서 론

최근 사물인터넷 기술의 발달로 인해 수집되는 데이터가 많아지면서 기계학습에 기반을 둔 빅데이터 분석이 큰 관심을 받고 있고 특히 농축산

분야를 포함한 다양한 학문분야에서 활용되고 있다 [1-3]. [1]의 연구에서는 젖소와 관련해서 수집된 빅데이터를 기반으로 reproduction rate를 분석하였고, [2]에서는 기계학습분석틀인 WEKA 틀을 이용한 젖소의 reproduction rate분석을 수행

하였다. [3]에서는 머신 러닝을 이용하여 모든의 생산성을 예측하였다. 빅데이터 분석을 통해서 기존에 파악하지 못한 새로운 특이점들을 파악할 수 있고 생산성 예측의 정확도를 높일 수 있다.

본 연구에서는 실험돈사에서 수집된 돼지의 환경데이터에 빅데이터 분석을 적용하여 돼지의 일당증체량을 예측하는 방안에 대해서 다루었다. 이를 위해서 온도/이산화탄소/풍속/습도의 일평균/일최고/일최소 데이터가 활용되었고 기계학습의 M5P 트리를 이용하여[4] 예측방안을 도출하였다. 또한 비율일당증체값을 새로 정의하고 사용하여 예측의 정확도를 높였다. 제안방안의 사용을 통해 각 축산농가에서는 양돈의 생산량을 미리 예측할 수 있고 이를 통해서 더 효율적인 돈사관리를 수행할 수 있을 것이다.

II. 본 론

본 연구에서는 실험돈사에서 사육하는 3개체에 3달동안 측정된 결과를 이용하여 진행되었다. 실험기간동안 체중은 34kg ~ 84kg의 범위에 있었고 온도/습도/이산화탄소/풍속/체중을 매일 측정하여 일최고/일최소/일평균 값을 구하였다. 연구에서 활용된 데이터는 그림 1에 나와 있다.

날짜	온도				습도				이산화탄소				풍속				체중
	Max	Min	Avg	Std	Max	Min	Avg	Std	Max	Min	Avg	Std	Max	Min	Avg	Std	
2015.4.17	27.510	22.949	25.995	0.2247	75.3	42.21	59.796	7.84	1120.07	0	0	0	0	0	0	0	37.1
2015.4.18	27.121	26.828	27.027	0.2095	61.2	13.99	53.037	6.61	481	0	0	0	0	0	0	0	37.1
2015.4.19	27.148	23.828	25.707	0.2084	71.1	1.29	54.242	6.91	439	0	0	0	0	0	0	0	37.1
2015.4.20	27.189	25.538	26.3	0.1888	60.15	0.71	57.332	7.96	489	1.6787	4	0	0	0	0	0	38.5
2015.4.21	27.046	22.888	25.005	0.1742	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.6
2015.4.22	27.688	27.004	27.349	0.1712	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.7
2015.4.23	28.026	28.207	28.122	0.1712	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.7
2015.4.24	28.026	28.207	28.122	0.1712	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.7
2015.4.25	28.026	28.207	28.122	0.1712	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.7
2015.4.26	28.026	28.207	28.122	0.1712	38.7	0	54.461	8.23	461	0	0	0	0	0	0	0	38.7

그림 1. 연구에서 활용된 데이터의 예

빅데이터 분석을 위한 기계학습 알고리즘으로는 트리 기반 알고리즘은 M5P 트리를 이용하였다. 이는 본 빅데이터 분석에서 예측하는 내용이 True/False의 바이너리 결과 값이 아닌 각 개체의 일당증체량으로서 SVD와 같이 두 그룹을 구별하는 알고리즘을 사용할 수 없었기 때문이다. 또한 다양한 사용가능한 빅데이터 알고리즘 중에서 M5P 트리가 가장 정확한 결과를 보였다.

본 연구에서는 빅데이터 분석을 통해 온도와 같은 환경정보를 기반으로 각 개체의 일당증체량(하루에 증가하는 체중량)을 예측하는 방안을 개발하였다. 특정일의 일당증체량을 예측하기 위해서 하루 전의 환경정보를 활용하였다. 또한 일당증체량이 개체의 체중에 따라서 변할 수 있으므로(표준 성장곡선을 참조해보면 개체의 체중이 많이 나갈 때의 일당증체량이 더 큼) 일당증체량을 체중으로 나눈 비율일당증체량을 예측하게 하였다. 본 연구에서는 첨부하지 않았지만 실제 일당증체량과 비율일당증체를 예측하는 알고리즘을 비교하였을 때 비율일당증체 예측 알고리즘의 정확도가 더 높은 것을 확인할 수 있었다.

성능분석을 위해서 10-fold validation을 활용하

였다. 또한 측정에 오차가 크다고 판단되는 경우의 데이터 샘플들을 제거하여 예측의 정확도를 높였다. 빅데이터 분석을 위해서는 WEKA 툴을 이용하였다 [2].

M5P 트리를 이용한 분석을 통해 비율일당증체 예측식을 다음과 같이 선형회귀식으로 예상할 수 있었다.

$$\text{비율일당증체} = -0.0014 \times (\text{일평균기온}) + 0.0002 \times (\text{일최대기온}) - 0.0001 \times (\text{일평균습도}) + 0.0003 \times (\text{일최대습도}) - 0.0053 \times (\text{일최대풍속}) + 0.0225$$

비율일당증체 예측식에서 볼 수 있듯이 각 개체의 비율일당증체에 가장 큰 영향을 주는 요소는 일평균 기온이고 일평균 기온이 높아짐에 따라 비율일당증체가 줄어드는 것을 확인할 수 있다. 또한 이산화탄소와 환경요소의 일최소데이터는 비율일당증체에 큰 영향을 주지 않는 것을 확인할 수 있다. 이를 통해서 실제 스마트돈사를 구축할 때 설치하는 센서의 종류와 수집되는 데이터를 조절할 수 있다.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient          0.9064
Mean absolute error            0.0011
Root mean squared error        0.0014
Relative absolute error        42.0673 %
Root relative squared error    41.8342 %
Total Number of Instances      77
    
```

그림 2. 비율일당증체 예측식의 정확도

그림 2는 M5P 트리를 이용해서 예측한 비율일당증체 예측식의 정확도를 보여준다. 실험 결과에서 확인할 수 있듯이 본 연구에서 개발한 예측식은 0.9의 높은 correlation을 보이고 42%의 상대적으로 낮은 비율오차를 보이는 것을 확인할 수 있다.

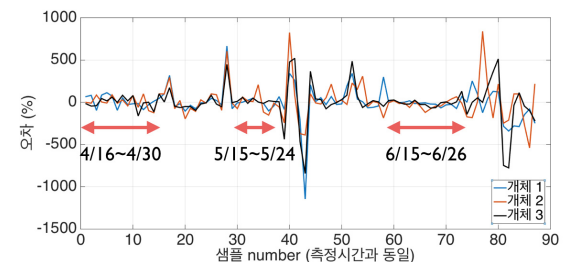


그림 3. 비율일당증체 예측식의 오차

1) 비율일당증체가 음수 값을 가지는 경우가 이에 해당한다. 이러한 데이터를 제거함으로써 정확도를 50% 이상 향상시킬 수 있다.

마지막으로 그림 3은 M5P 트리로 예측된 비육 일당증체값과 실제 값의 오차를 보여준다. 그림에서 볼 수 있듯이 3개의 개체에 대해서 비슷한 오차동향을 보이는 것을 확인할 수 있고 특히 일정 기간의 날짜에 오차율이 급증하는 것을 확인할 수 있다.

III. 결 론

본 연구에서는 실험돈사에서 측정된 돼지들의 생체데이터와 환경데이터에 빅데이터 분석방안을 적용하여 돼지의 일당증체량을 예측하는 알고리즘을 개발하였다. 이를 위해서 온도/이산화탄소/풍속/습도의 일평균/일최대/일최소 환경데이터를 활용하였다. 또한 높은 정확도를 갖는 비육일당증체량 예측방안을 M5P 트리를 이용하여 개발하였고 이를 통해 돈사의 평균온도가 돼지의 일당증체에 가장 큰 영향을 미치는 환경요소임을 확인하였다. 본 연구에서 개발한 비육일당증체 예측식은 일반 양돈농가에서 각 개체의 생산성을 예측하는데 활용할 수 있고 생산량이 낮을 때 어떠한 환경요소를 조절해야 하는지 돈사환경조절 방안에 활용될 수 있을 것이다. 또한 스마트돈사를 구축할 때 어떠한 환경요소를 측정해야 하는지 결정하는데 활용될 수 있을 것이다.

ACKNOWLEDGMENTS

본 성과물(논문, 산업재산권, 품종보호권 등)은 농촌진흥청 연구사업 (세부과제명: 가축관리 및 돈사환경 관측 빅데이터 활용에 관한 연구, 세부과제번호: PJ0105412015)의 지원에 의해 이루어진 것임. 또한 이 연구는 2014년도 경상대학교 학술진흥지원사업 연구회지원 연구비에 의하여 수행하였음

참고문헌

- [1] DZ. Caraviello, KA Weigel, M. Craven, D. Gianola, NB. Cook, KV. Nordlund, PM. Fricke, MC. Wiltbank, "Analysis of reproductive performance of lactating cows on large dairy farms using machine learning algorithms," *J Dairy Sci.* 89(12) 2006.
- [2] R. Scott Mitchell, Robert A. Sherlock, Lloyd A. Smith, "An investigation into the use of machine learning for determining oestrus in cows," *Computers and Electronics in Agriculture*, 15(3), 1996.

- [3] 이민수, 최영찬, "머신러닝을 활용한 모돈의 생산성 예측모델," *한국농촌지도학회*, 16(4), 2009
- [4] P. Flach, "Machine learning: the art and science of algorithms that make sense of data," *Cambridge University Press*, 2012.