

帰納論理プログラミングを用いた乳牛の生涯搾乳量に係る潜在的要因の推定に関する研究

佐々木 駿[‡]秦野 亮^{†‡}大和田 勇人^{†‡}西山 裕之^{†‡}
[†] 東京理科大学理工学部 [‡] 東京理科大学理工学研究科

1 序論

近年日本の酪農は危機的な状況である。酪農家戸数は毎年4%程度の廃業率で推移し、平成21年からの10年間で約32%減少している。加えて、乳牛の飼育頭数は同期間で約12%減少し[1]、平均産次度も低下傾向にある[2]。一方、酪農家一戸当たりの飼育頭数は増加しているため、作業負担は増えている。この問題解決のために、一部の酪農家は作業の自動化や効率化を目的として、自動搾乳機等の自動ロボットの導入を行っている。これらのロボットは、稼働時に様々なデータを収集できる。本研究では、自動搾乳機が搾乳時に収集する乳牛に関するデータを機械学習の1手法である帰納論理プログラミング(ILP)に適用し、乳牛の生涯搾乳量に係る潜在的な要因に関する分析を行う。乳牛の生涯搾乳量は酪農経営において経済的に重要で、大きい方がよいことは自明であるが、事前に予測する事は専門家でも難しい。しかし、状態のよくない乳牛を購入または飼育し続ける事は経済的な損失を与える可能性がある。本研究では、このような損失回避と生産基盤強化に貢献する情報として、生涯搾乳量が多い乳牛に共通する傾向をILPによりルール化する手法を提案する。

本研究で用いるILPはS. Muggletonら[3]により提案された一階述語論理に基づいた帰納推論により分類問題を解決する手法である。通常の機械学習とは異なり、モデルが論理式に基づくルールの集合という形式で記述されているため、直接可読という利点がある。

2 関連研究

Shahinfarら[4]は2014年に機械学習によるホルスタイン乳牛の人工授精結果の予測を行った。決定木を始めとする5種類の機械学習手法に乳牛データを適用し、人工授精による妊娠/非妊娠の結果予測を行った。結果、ランダムフォレストが70%という最も高い精度で人工授精の結果を予測できた。

松本ら[5]は、ILPを用いて乳用牛の人工授精に関わる複数のデータを複合的に解析し、乳用牛の人工授精が成功するための最適な条件のルールを導く研究を行った。この結果、プロゲステロン値に加えて採食量や活動量、産次数を複合的に考慮することでより最適な人工授精のタイミングや条件を予測することが可能であることを示した。ILPの手法は、他の機械学習器を用いた手法と異なりルールを論理形式の意味で明示化できる点に利点がある。そのため、本研究では同様にルールを導くための学習器としてILPを用いる。

3 データセット

本研究で扱う乳用牛に関する様々なデータは、酪農の現場に導入されているデラバル Voluntary Milking System(以下デラバル VMS)によって収集された乳用牛の搾乳量やホルモンデータ等である。デラバル VMSは24時間稼働可能な自動搾乳機であり、乳用牛が機内に入ることで自動的に稼働し、乳用牛から搾乳を行う。その際に、搾乳時間や乳量および血乳含有度などのデータ収集され、酪農管理システム[6]のデータベースに蓄積される。本研究では、2015年から2019年5月までの国内の29農場における8329頭の乳牛データに関する搾乳量や採食量、そしてホルモンデータなどのデータをこのデータベースから抽出し、およそ172万行12列、140MBのデータセットを作成した。

4 提案手法

本研究で提案する手法は、以下のステップで行われる。(1) データベースからデータを抽出し、デー

Estimation of potential factors related to the total milk yield in a lifetime of dairy cows by using Inductive Logic Programming

Shun Sasaki[‡], Ryo Hatano^{†‡}, Hayato Ohwada^{†‡} and Hiroyuki Nishiyama^{†‡}

[†]Faculty of Sci. and Tech. Tokyo University of Science

[‡]Graduate School of Sci. and Tech.

Tokyo University of Science

タセットを作成する。(2) 作成したデータセットの値をもとに様々な計算を行い、特徴量および論理形式データを生成する。(3) 学習ターゲットとする乳牛のデータを後述する一定の条件の下で決定し、閾値によって正事例と負事例を分割、ラベル付けを行う。(4) 特徴量のデータセットを用いて ILP のモデル (一階述語論理の形式で記述されたルールの集合) を学習し、評価・分析する。

4.1 前処理

ILP で機械学習を行うためには、デルプロTM から抽出し作成したデータセットを ILP に適用可能な形に変換しなければならない。このため、前処理として特徴量の生成および一階述語論理の形式のデータへの変換などの前処理を行う。本研究では、産次数や人工授精の失敗回数、酪農場のエリア、また各産次数で搾乳量ピークまでおよびピーク後における搾乳量および採食量、ホルモンデータに関する基本統計量など、全部で 26 種類の特徴量を生成した。また、搾乳量ピークの計算には以下の式 (1) で表される、指数平滑移動平均 (EMA) を用いた。

$$EMA = \begin{cases} 0 & (i < n) \\ \frac{\sum_{k=1}^n V_k}{n} & (i = n) \\ (V_i - \text{前日の EMA}) \times \alpha \\ \quad + \text{前日の EMA} & (i > n) \end{cases} \quad (1)$$

V_i は i 日目の搾乳量を表し、 α は $\frac{2}{n+1}$ で表される。本研究では、 n を 3 ($\therefore \alpha = 0.5$) に設定した。

続いて、生成したこれらの特徴量を一階述語論理の形式に変換する。本研究では、「特徴量を表す述語」、「特徴量の期間を表す述語」、「値の大小を表す述語」の 3 種類を考慮し、それぞれ定義した。

4.2 モデル学習

本研究では、ILP の学習を並列分散処理として実行可能なシステムである Parallel GKS[7] を用いた。なお、GKS には表 1 のようなパラメータを設定す

表 1. GKS で用いるパラメータ一覧

パラメータ名	説明
vlimit	ルールに用いられる変数の上限
slimit	ルールに含まれる述語の上限
plimit	ルールに含まれる正事例の下限
nlimit	ルールに含まれる負事例の上限
width	探索の幅

る必要がある。そのため本研究では、 $vlimit = 20$,

$slimit = 10$, $plimit = 3$, $nlimit = 10$, $width = 2$ に設定して学習を行った。

5 結果

学習の結果、正事例の乳牛に対する 50 種類のルールを得られた。この中には、「北部の酪農場では 1 産次の搾乳ピーク前採食量平均がやや少なく、2 産次および 3 産次の搾乳ピーク後の BHB 分散が小さい乳牛の生涯搾乳量は多い傾向がある」と言った意味合いの自明ではないルールを得る事ができた。これは、酪農家が乳牛を飼養する際や意思決定を行う際に有用と考えられる。

6 まとめ

本稿では、乳牛の搾乳量やホルモンデータを ILP を用いて分析することで、搾乳量の多い乳牛に共通する傾向をルール化する手法を提案した。結果、酪農家が乳牛を飼養する際に有用と考えられるルールが得られた。今後は、エリア別分析や時系列データを述語として扱う事でより有用なルールの抽出を目指す。

参考文献

- [1] 農林水産省. 畜産統計調査結果概要. Available: <http://www.maff.go.jp/j/tokei/kouhyou/tikusan/>. [Accessed: 5-Jan-2020].
- [2] 家畜改良事業団. 平成 30 年度乳用牛群能力検定成績速報. Available: <http://liaj.lin.gr.jp/japanese/newmilkset.html>. [Accessed: 5-Jan-2020].
- [3] Stephen Muggleton and Luc De Raedt. Inductive logic programming: Theory and methods. *The Journal of Logic Programming*, Vol. 19, pp. 629–679, 1994.
- [4] Saleh Shahinfar, David Page, Jerry Guenther, Victor Cabrera, Paul Fricke, and Kent 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.
- [5] Atsushi Matsumoto, Chikara Kubota, and Hayato Ohwada. Extracting rules for successful conditions for artificial insemination in dairy cattle using inductive logic programming. In *Proceedings of the 9th International Conference on Machine Learning and Computing*, pp. 6–10. ACM, 2017.
- [6] DeLaval - Japan. デルプロTM. Available: <http://www.delvalcorporat.com/our-products-and-services/farm-support/DelPro/>.
- [7] Hiroyuki Nishiyama and Hayato Ohwada. Parallel inductive logic programming system for superlinear speedup. In Nicolas Lachiche and Christel Vrain, editors, *Inductive Logic Programming*, pp. 112–123, Cham, 2018. Springer International Publishing.