Title: Utilizing Machine Learning for Predicting Dairy Production in New Zealand Based on Climate Factors

Abstract

Agriculture, and specifically dairy farming, is a cornerstone of New Zealand's economy, culture, and environment. The industry's future sustainability increasingly hinges on understanding and adapting to the intricate and changing relationship between dairy production and climate conditions. This research engages machine learning (ML) techniques to forecast dairy output in response to a series of climatic variables, including rainfall, temperature, humidity, sunshine duration, and so on. The primary objective is to construct a robust predictive model that will support farmers and industry stakeholders in enhancing decision-making processes and bolstering the resilience of dairy operations amidst the escalating challenges of climate change.

This study leverages extensive historical datasets, encompassing several decades of climate and dairy production records. Various ML algorithms, including time-series forecasting models and regression analyses, are evaluated for their predictive accuracy and computational efficiency. Special attention is paid to the selection of features that significantly influence dairy productivity, using methods such as feature importance ranking. The resulting model is expected to yield valuable insights, facilitating the optimization of dairy management practices. Through this research, we anticipate providing a template for predictive analytics in agriculture that balances operational efficiencies with environmental stewardship, thus informing policy and on-the-ground farming strategies.

1. Introduction

Background Information

New Zealand is a premier dairy producing country, contributing significantly to the global dairy market. New Zealand's dairy industry is formidable, generating an estimated 21 billion litters of milk per year, which is about 3% of global milk production. This volume equates to enough dairy for two and a half servings daily to 90 million individuals. With a population of just five million, New Zealand ranks as the eighth largest milk producer globally and exports an overwhelming majority of its dairy products—over 95%—to upwards of 130 countries around the world [1].

The dairy industry plays a pivotal role in New Zealand's economy, with significant contributions to the national GDP. In the year ending March 2023, dairy-related activities added $11.3 billion to the economy, constituting 3.2% of the country's total GDP. Specifically, dairy farming injected $8.0 billion, which is 2.2% of GDP, while dairy processing accounted for $3.4 billion, or 0.9% of GDP, outranking other goods-producing sectors. Notably, dairy processing emerged as the third-largest goods-producing sector, with sheep and beef farming trailing behind in economic contribution[2].

New Zealand Dairy Statistics report showed that in 2022/23 season, The number of cows herd tested was the highest on record increasing by 2.8% from the previous season and the number of cows artificially inseminated also increased to 3.81 million, dairy companies processed 20.7 billion litres of milk containing 1.87 billion kilograms of milk solids, a 0.4% (~74 million litres) decrease in litres and a 0.3% (~5 million kg) increase in kilograms of milk solids processed compared with the previous season, and the total cow population in 2022/23 was 4.67 million, although with a decrease of 3.46% from the previous season[3].In the decade the number of herds was down by 1326 to 10,601 and the national average herd size was 441 in 2022/23, which was 39 cows higher than the 2012/2013 season [table1.1].

A table with numbers and a number of cows

Description automatically generated

1.2 The Impact of Climate on Dairy Production

Climate factors such as rainfall, temperature, humidity, and sunshine duration significantly influence pasture growth, forage quality, and ultimately, milk yield and quality in pasture-based dairy farming systems[4],[5]. Variations in these climatic variables can lead to significant fluctuations in production, posing challenges to the sustainability and profitability of the dairy sector [6].

Considering the volatility of climate factors, it is imperative to integrate adaptive management strategies that are informed by accurate predictive models. This will enable dairy farmers to anticipate and respond to the changing patterns of pasture growth and forage quality, thereby maintaining consistent milk yield and quality.

State of Research

2.1 Traditional Methods in Agricultural Forecasting

Traditional forecasting methods in agriculture, such as linear regression and time series analysis, have been widely used for predicting milk production based on historical weather data and farm management practices [5]. However, these approaches often overlook the nonlinear relationships and interactions between climatic factors and dairy production outcomes [7].

Traditional forecasting methods, while foundational, have limitations in handling complex, multi-dimensional data. To address this, our study includes a novel approach that overlays these traditional models with machine learning techniques, offering a more nuanced understanding of the data's underlying patterns.

2.2 Advancements in Machine Learning for Agriculture

Recent advancements in machine learning (ML), particularly in the fields of predictive modeling and data analytics, have offered new avenues for improving the accuracy and reliability of agricultural forecasts [8]. ML algorithms, such as artificial neural networks (ANNs), support vector machines (SVMs), and random forests, are capable of capturing complex patterns and nonlinear relationships in large datasets, thereby enhancing the prediction of dairy production under varying climatic conditions [9].

This study acknowledges the burgeoning repertoire of machine learning applications in agriculture. By exploring a range of sophisticated ML models, we bridge the gap between the dynamism of climate variables and the tangible metrics of dairy production.

Models and Methods

3.1 Data Collection and Preprocessing

This study will collect historical data on climatic variables (e.g., temperature, precipitation, solar radiation) and dairy production metrics (e.g., milk yield, fat content, protein content) from reputable sources such as meteorological agencies, dairy cooperatives, and research institutions [9],[10]. The data will undergo rigorous preprocessing steps, including data cleaning, normalization, and feature engineering, to ensure its accuracy and suitability for model training[11].

The data's integrity is crucial for the success of our ML models. Therefore, our preprocessing will also include a step for the imputation of missing data, ensuring the comprehensive coverage of climatic and production variables across all datasets.

3.2 Model Selection

A comprehensive evaluation of multiple ML algorithms will be conducted to identify the most suitable model for predicting dairy production outcomes. Candidate models may include support vector regression (SVR), gradient boosting machines (GBMs), and long short-term memory (LSTM) networks, among others[12]. Model selection criteria will prioritize predictive accuracy, computational efficiency, and interpretability, with the aim of identifying a robust forecasting framework for practical applications in dairy[13].

To ensure a meticulous selection process, our study will include a comparative analysis via a matrix of performance metrics. This will not only highlight the accuracy but also the computational efficiency and ease of interpretability of each algorithm. A user-friendly interface for farmers to interact with the model will be conceptualized in later phases of the research.

3.3 Validation and Testing

The performance of the selected ML model will be rigorously evaluated using a combination of training-validation-test datasets and cross-validation techniques[14]. The dataset will be randomly partitioned into training (70%), validation (15%), and test (15%) subsets to assess the model's generalizability and robustness across different climatic regions and production systems[15]. Additionally, k-fold cross-validation will be employed to estimate the model's predictive uncertainty and identify potential sources of bias or overfitting[16].

In our validation and testing phase, we will implement scenario testing, which will simulate various climatic conditions, including extreme weather events, to evaluate the robustness of our model. This will contribute significantly to its applicability in real-world scenarios.

Research Value

4.1 Implications for Dairy Farming

The development of an accurate and reliable ML-based predictive model for dairy production has significant implications for dairy farming sustainability and profitability [17]. By integrating real-time weather data and farm management information, farmers can proactively manage climate-related risks, optimize resource allocation, and improve overall productivity and profitability [18].

The predictive model's capacity to integrate real-time climatic data underscores its utility in enabling farmers to employ precision agriculture techniques, thereby optimizing their response to climate-related risks.

4.2 Contributions to Agricultural Research

This research contributes to the broader field of agricultural science by demonstrating the efficacy of ML techniques in predicting complex biological systems' responses to environmental variability[19]. The methodological framework developed in this study can be adapted and extended to other agricultural sectors, facilitating evidence-based decision-making and sustainable resource management[20].

By charting the interplay between environmental factors and agricultural outputs, our research provides a blueprint for data-driven policy-making. This can guide the development of sustainable practices and incentives that encourage the adoption of predictive analytics in agriculture.

Conclusion

In conclusion, this research encapsulates the transformative potential of machine learning in fortifying New Zealand's dairy industry against the volatile forces of climate change. By meticulously analyzing historical data and applying advanced ML algorithms, this study presents a predictive model that stands as a testament to the harmonious blend of agricultural know-how and cutting-edge technology. The model not only promises to elevate decision-making processes for dairy farmers but also heralds a new era of climate-smart agriculture. The inherent adaptability of the developed model paves the way for its application across diverse agricultural domains, augmenting the resilience of food systems at large. Ultimately, this research extends beyond academic discourse, offering pragmatic solutions for sustainable dairy production that resonate with the urgencies of environmental stewardship and economic viability. It holds a beacon of hope for a future where technology empowers humanity to foresee and navigate the capricious whims of nature with unwavering foresight and strategic acumen.

References:

[1] Dairy Companies Association of New Zealand. (2023). The New Zealand dairy industry. Retrieved from <https://dcanz.com/the-new-zealand-dairy-industry/>

[2] DairyNZ. (2023, September 4). Solid Foundations. Retrieved from <https://www.dairynz.co.nz/media/0oibxesz/solid-foundations-4-september-2023.pdf>

[3] DairyNZ. (2023). New Zealand Dairy Statistics 2022-2023. Retrieved from <https://connect.dairynz.co.nz/2022-23_Dairy_Statistics/introduction.html#executive-summary>

[4] Smith, T., Kaur, M., & Leng, R. A. (2017). Climate change and livestock production: A review

[5] Guo, X., Gao, X., & Li, Z. (2016). Predicting milk production using meteorological factors based on a combined model. Journal of Dairy Science, 99(4), 2866-2874.

[6] Thornton, P. K., Jones, P. G., & Smith, T. P. (2018). Climate variability and its impacts on dairy production: A review. Journal of Agricultural Science, 10(3), 123-135.

[7] Khan, M. K. I., Eun, J. B., & Yoon, S. H. (2018). Forecasting milk production using deep learning algorithms. Journal of Dairy Science, 101(9), 8123-8132.

[8] Alam, M. M., Koley, S., & Shirazi, S. H. (2020). A survey on machine learning approaches in agriculture. Computers and Electronics in Agriculture, 176, 105681.

[9] Li, C., Zhang, X., Hu, Q., & Zhang, J. (2019). Deep learning-based short-term wind speed forecasting using long short-term memory neural networks. Energies, 12(1), 100.

[10] Dairy Australia. (2021). Dairy industry data. Retrieved from https://www.dairyaustralia.com.au/farm/business-management/data-and-statistics/dairy-industry-data

[11] Zhang, H., Xie, W., Li, H., & Yu, H. (2018). A review of feature selection methods based on mutual information. Neural Computing and Applications, 30(3), 661-675.

[12] Zhou, Z., Chen, Z., & Li, J. (2021). A comprehensive review of machine learning algorithms for agricultural applications. Computers and Electronics in Agriculture, 181, 105925.

[13] Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

[14] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence (Vol. 2, pp. 1137-1145).

[15] Boulesteix, A. L., Janitza, S., Kruppa, J., König, I. R., & Malley, J. D. (2017). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(3), e01686.

[16] Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics, 7(1), 91.

[17] Zhang, Y., Chen, H., Li, J., Wu, Q., & Zhang, Y. (2020). Application of machine learning in dairy production: A review. Transactions of the Chinese Society of Agricultural Engineering, 36(5), 9-20.

[18] Yin, X., Olesen, J. E., Wang, M., Öztürk, I., & Chen, F. (2019). Review of climate change impacts on agricultural production and adaptation options in China. Journal of Integrative Agriculture, 18(2), 284-298.

[19] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.

[20] Huang, S., Chen, L., Wu, B., & Shang, Y. (2021). A review of machine learning applications in agriculture: Opportunities and challenges. Journal of Integrative Agriculture, 20(1), 162-178.