百度翻译

标题：利用机器学习根据气候因素预测新西兰的乳制品产量

抽象的

农业，特别是奶牛养殖，是新西兰经济、文化和环境的基石。 该行业未来的可持续性越来越依赖于理解和适应乳制品生产与气候条件之间复杂且不断变化的关系。 本研究采用机器学习 (ML) 技术来预测乳制品产量，以响应一系列气候变量，包括降雨量、温度、湿度、日照时长等。主要目标是建立一个潜在的预测模型来支持农民和行业利益相关者 在气候变化挑战不断升级的情况下，加强决策过程并提高乳制品业务的复原力。 该研究利用了广泛的历史数据集，包括数十年的气候和乳制品生产记录。 评估各种机器学习算法（包括时间序列预测模型和回归分析）的预测准确性和计算效率。 特别注意使用特征重要性排名等方法来选择对乳制品生产率有重大影响的特征。 由此产生的模型预计将产生有价值的见解并促进乳制品管理实践的优化。 通过这项研究，我们希望为农业预测分析提供一个模板，以平衡运营效率与环境管理，为政策和实地农业战略提供信息。

1 简介

1.1 新西兰乳制品

新西兰是领先的乳制品生产国，对全球乳制品市场做出了重大贡献。 新西兰乳制品业非常强大，每年生产约210亿升牛奶，约占全球牛奶产量的3%。 这个数量相当于每天为9000万人提供两份半乳制品。 新西兰人口仅有500万，但却是全球第八大牛奶生产国，其绝大多数乳制品（95%以上）出口到全球130多个国家[1]。

乳制品行业在新西兰经济中发挥着举足轻重的作用，为国家GDP做出了重大贡献。 截至 2023 年 3 月的一年里，乳制品相关活动为经济增加了 113 亿美元，占该国 GDP 总额的 3.2%。 具体而言，奶牛养殖业贡献了 80 亿美元，占 GDP 的 2.2%，乳制品加工贡献了 34 亿美元，占 GDP 的 0.9%，超过了其他商品生产部门。 [2]。

新西兰乳业2022/23产季的牛群检测和人工授精创历史新高，检测奶牛数量增加2.8%，人工授精奶牛增加381万头。 乳制品企业加工了207亿升牛奶，其中含乳固体18.7亿公斤，产量较上一季略有下降，但固体含量略有增加[3]。 尽管奶牛总数下降了3.46%，但全国平均存栏规模增加到441头，比十多年前的39头增加[表1.1]。

1.2 气候对乳制品生产的影响

降雨量、温度、湿度和日照时间等气候因素显着影响牧场生长、饲料质量，并最终影响牧场奶牛养殖系统的牛奶产量和质量[4]、[5]。 这些气候变量的变化可能导致产量大幅波动，对乳制品行业的可持续性和盈利能力构成挑战[6]。

鉴于气候因素的波动性，必须整合由准确预测模型提供信息的适应性管理策略。 这将使奶农能够预测并应对牧场生长和牧场质量模式的变化，保持稳定的牛奶产量和质量。

2、研究现状

2.1 传统农业预报方法

传统的农业预测方法，例如线性回归和时间序列分析，已被广泛用于根据历史天气数据和农场管理实践来预测牛奶产量[5]。 然而，这些方法往往忽略了气候因素与乳制品生产结果之间的非线性关系和相互作用[7]。 尽管传统的预测方法很基础，但它们在处理复杂的多维数据时存在局限性。 为了解决这个问题，我们的研究包括一种新方法，将这些传统模型与机器学习技术叠加，以提供对数据中潜在模式的更细致的理解。

2.2 农业机器学习进展

机器学习（ML）的最新进展，特别是在预测建模和数据分析领域，为提高农业预测的准确性和可靠性提供了新方法[8]。 机器学习算法，例如人工神经网络（ANN）、支持向量机（SVM）和随机森林，能够捕获大数据集中的复杂模式和非线性关系，从而增强对不同气候条件下乳制品产量的预测[9 ]。

这项研究承认机器学习在农业中的应用正在蓬勃发展。 通过探索一系列复杂的机器学习模型，我们弥合了气候变量的动态与乳制品生产的有形指标之间的差距。

3 模型与方法

3.1 数据收集和预处理

本研究将从气象机构、乳制品合作社和研究机构等信誉良好的来源收集气候变量（例如温度、降水、太阳辐射）和乳制品生产指标（例如产奶量、脂肪含量、蛋白质含量）的历史数据[9]， [10]。 数据将经过严格的预处理步骤，包括数据清洗、标准化和特征工程，以确保其准确性和模型训练的适用性[11]。

数据完整性对于我们的机器学习模型的成功至关重要。 因此，我们的预处理还将包括估算缺失数据的步骤，确保全面覆盖所有数据集中的气候和生产变量。

3.2 型号选择

我们将对多种机器学习算法进行综合评估，以确定最合适的模型来预测乳制品生产结果。 候选模型可能包括支持向量回归（SVR）、梯度增强机（GBM）和长短期记忆（LSTM）网络[12]。 模型选择标准将优先考虑预测准确性、计算效率和可解释性，以确定乳制品实际应用的稳健预测框架[13]。

为了确保精心挑选过程，我们的研究将包括通过绩效指标矩阵进行比较分析。 这不仅会提高准确性，还会提高每个算法的计算效率和可解释性。 在研究的后期阶段，将概念化一个供农民与模型交互的用户友好界面。

3.3 验证与测试

所选机器学习模型的性能将结合训练-验证-测试数据集和交叉验证技术进行严格评估[14]。 将数据集随机分为训练（70%）、验证（15%）和测试（15%）子集，以评估模型在不同气候区域和生产系统中的普遍性和鲁棒性[15]。 此外，将采用 k 折交叉验证来估计模型的预测不确定性并识别潜在的偏差或过度拟合来源 [16]。

在验证和测试阶段，我们将实施模拟各种气候条件（包括极端天气事件）的场景测试，以评估模型的稳健性。 这将极大地有助于其在实际场景中的适用性。

4、研究价值

4.1 对奶牛养殖的影响

开发准确可靠的基于机器学习的乳制品产量预测模型对于奶牛养殖的可持续性和盈利能力具有重要意义[17]。 通过整合实时天气数据和农场管理信息，农民可以主动管理气候相关风险，优化资源配置，提高整体生产力和盈利能力[18]。

该预测模型整合实时气候数据的能力凸显了其在帮助农民采用精准农业技术方面的效用，从而优化他们对气候相关风险的反应。

4.2 对农业研究的贡献

这项研究通过证明机器学习技术在预测复杂生物系统对环境变化的响应方面的功效，为更广泛的农业科学领域做出了贡献[19]。 本研究开发的方法框架可以适应并扩展到其他农业部门，促进循证决策和可持续资源管理[20]。

通过绘制环境因素与农业产出之间的相互作用，我们的研究为数据驱动的政策制定提供了蓝图。 这可以指导可持续实践和激励措施的发展，鼓励在农业中采用预测分析。

5 结论

总之，这项研究概述了机器学习的变革潜力，以增强新西兰乳制品行业抵御气候变化破坏稳定力量的能力。 通过仔细分析历史数据并应用先进的机器学习算法，本研究提出了一个预测模型，展示了农业知识与尖端技术的和谐融合。 该模型不仅有望改善奶农的决策过程，还预示着气候智能型农业的新时代。 所开发模型的固有适应性为其在不同农业部门的应用铺平了道路，增强了整个粮食系统的弹性。 最终，这项研究超越了学术论述，为可持续乳制品生产提供了务实的解决方案，与环境管理和经济可行性的紧迫性产生了共鸣。

Title: Using Machine Learning to Predict Dairy Production in New Zealand Based on Climate Factors

Abstract

Agriculture, especially dairy farming, is the cornerstone of New Zealand's economy, culture, and environment. The future sustainability of the industry increasingly depends on understanding and adapting to the complex and constantly changing relationship between dairy production and climate conditions. This study uses machine learning (ML) technology to predict dairy production in response to a range of climate variables, including rainfall, temperature, humidity, and sunshine duration. The main goal is to establish a potential predictive model to support farmers and industry stakeholders in strengthening decision-making processes and improving the resilience of dairy businesses in the face of escalating climate change challenges. This study utilized a wide range of historical datasets, including decades of climate and dairy production records. Evaluate the prediction accuracy and computational efficiency of various machine learning algorithms, including time series prediction models and regression analysis. Pay special attention to using methods such as feature importance ranking to select features that have a significant impact on dairy productivity. The resulting model is expected to generate valuable insights and promote the optimization of dairy management practices. Through this study, we hope to provide a template for agricultural forecasting analysis to balance operational efficiency and environmental management, and to provide information for policies and field agricultural strategies.

1 Introduction

1.1 New Zealand dairy products

New Zealand is a leading producer of dairy products and has made significant contributions to the global dairy market. The dairy industry in New Zealand is very strong, producing approximately 21 billion liters of milk annually, accounting for approximately 3% of global milk production. This quantity is equivalent to providing two servings of semi dairy products to 90 million people per day. New Zealand has a population of only 5 million, but it is the eighth largest milk producing country in the world, with the vast majority of its dairy products (over 95%) exported to over 130 countries worldwide.

The dairy industry plays a crucial role in the New Zealand economy and has made significant contributions to the country's GDP. In the year ending March 2023, dairy related activities added $11.3 billion to the economy, accounting for 3.2% of the country's total GDP. Specifically, the dairy industry contributed $8 billion, accounting for 2.2% of GDP, while dairy processing contributed $3.4 billion, accounting for 0.9% of GDP, surpassing other commodity production sectors. [2] .

The testing and artificial insemination of cattle in New Zealand's dairy industry reached a historic high in the 2022/23 production season, with a 2.8% increase in the number of tested cows and a 3.81 million increase in artificial insemination cows. Dairy companies processed 20.7 billion liters of milk, including 1.87 billion kilograms of milk solids, with a slight decrease in production compared to the previous season, but a slight increase in solid content. Although the total number of cows decreased by 3.46%, the national average inventory size increased to 441, an increase from 39 more than a decade ago [Table 1.1].

1.2 The impact of climate on dairy production

Climate factors such as rainfall, temperature, humidity, and sunshine duration significantly affect the growth of pastures, feed quality, and ultimately affect the milk production and quality of dairy farming systems in pastures [4], [5]. The changes in these climate variables may lead to significant fluctuations in production, posing challenges to the sustainability and profitability of the dairy industry [6].

Given the volatility of climate factors, it is necessary to integrate adaptive management strategies that provide information from accurate predictive models. This will enable dairy farmers to predict and respond to changes in pasture growth and quality patterns, maintaining stable milk production and quality.

2. Research status

2.1 Traditional agricultural forecasting methods

Traditional agricultural forecasting methods, such as linear regression and time series analysis, have been widely used to predict milk production based on historical weather data and farm management practices [5]. However, these methods often overlook the nonlinear relationship and interaction between climate factors and dairy production outcomes [7]. Although traditional prediction methods are fundamental, they have limitations when dealing with complex multidimensional data. To address this issue, our research includes a new approach that overlays these traditional models with machine learning techniques to provide a more detailed understanding of potential patterns in the data.

2.2 Progress in Agricultural Machine Learning

The latest advances in machine learning (ML), particularly in predictive modeling and data analysis, provide new methods for improving the accuracy and reliability of agricultural forecasting [8]. Machine learning algorithms, such as artificial neural networks (ANN), support vector machines (SVM), and random forests, can capture complex patterns and nonlinear relationships in big datasets, thereby enhancing the prediction of dairy production under different climate conditions [9].

This study acknowledges that the application of machine learning in agriculture is flourishing. By exploring a series of complex machine learning models, we have bridged the gap between the dynamics of climate variables and tangible indicators of dairy production.

3 Models and Methods

3.1 Data collection and preprocessing

This study will collect historical data on climate variables (such as temperature, precipitation, solar radiation) and dairy production indicators (such as milk production, fat content, protein content) from reputable sources such as meteorological agencies, dairy cooperatives, and research institutions [9], [10]. The data will undergo strict preprocessing steps, including data cleaning, standardization, and feature engineering, to ensure its accuracy and the applicability of model training [11].

Data integrity is crucial for the success of our machine learning models. Therefore, our preprocessing will also include steps to estimate missing data, ensuring comprehensive coverage of climate and production variables in all datasets.

3.2 Model Selection

We will conduct a comprehensive evaluation of multiple machine learning algorithms to determine the most suitable model to predict dairy production results. Candidate models may include Support Vector Regression (SVR), Gradient Booster Machine (GBM), and Long Short Term Memory (LSTM) networks [12]. The model selection criteria will prioritize prediction accuracy, computational efficiency, and interpretability to determine a robust prediction framework for practical applications in dairy products [13].

To ensure a careful selection process, our research will include comparative analysis through a performance indicator matrix. This not only improves accuracy, but also enhances the computational efficiency and interpretability of each algorithm. In the later stage of the research, conceptualize a user-friendly interface for farmers to interact with the model.

3.3 Verification and Testing

The performance of the selected machine learning model will be rigorously evaluated by combining training validation testing datasets and cross validation techniques [14]. Randomly divide the dataset into training (70%), validation (15%), and testing (15%) subsets to evaluate the universality and robustness of the model in different climate regions and production systems [15]. In addition, k-fold cross validation will be used to estimate the predictive uncertainty of the model and identify potential sources of bias or overfitting [16].

During the validation and testing phase, we will conduct scenario tests simulating various climate conditions (including extreme weather events) to evaluate the robustness of the model. This will greatly contribute to its applicability in practical scenarios.

4. Research value

4.1 Impact on dairy farming

Developing accurate and reliable machine learning based dairy yield prediction models is of great significance for the sustainability and profitability of dairy farming [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 ability of this prediction model to integrate real-time climate data highlights its effectiveness in helping farmers adopt precision agriculture technologies, thereby optimizing their response to climate related risks.

4.2 Contribution to Agricultural Research

This study has contributed to the broader field of agricultural science by demonstrating the effectiveness of machine learning techniques in predicting the response of complex biological systems to environmental changes [19]. The methodological framework developed in this study can be adapted and extended to other agricultural sectors, promoting evidence-based decision-making and sustainable resource management [20].

Our research provides a blueprint for data-driven policy-making by mapping the interaction between environmental factors and agricultural output. This can guide the development of sustainable practices and incentive measures, encouraging the adoption of predictive analysis in agriculture.

5 Conclusion

In summary, this study outlines the transformative potential of machine learning to enhance the ability of the New Zealand dairy industry to withstand the destabilizing forces of climate change. By carefully analyzing historical data and applying advanced machine learning algorithms, this study proposes a predictive model that demonstrates the harmonious integration of agricultural knowledge and cutting-edge technology. This model is not only expected to improve the decision-making process of dairy farmers, but also heralds a new era of climate intelligent agriculture. The inherent adaptability of the developed model paves the way for its application in different agricultural sectors and enhances the resilience of the entire food system. Ultimately, this study goes beyond academic discourse and provides practical solutions for sustainable dairy production, resonating with the urgency of environmental management and economic feasibility.