NUTRITIONAL DEFICIENCY PREDICTION BY REGION

USING MACHINE LEARNING

ANNE DASHINI KANNAN

UNIVERSITY TEKNOLOGI MALAYSIA

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter discusses into the challenges faced I previous research study related to this topic and provides as in-dept explanations details of their research findings. It writes and overview of the application of machine learning withing the health care department and areas, focussing its functionality and predictive insights. Key factors and datasets are important to developing prediction models are outlined to provide clear view on their development and creating prediction models. Furthermore, this chapter go in details on how innovative strategies to address the prediction of nutritional deficiencies.

2.2 Overview

Nutritional deficiency has become the global challenge especially on those who has low-income regions, remains a consistent issue affecting tons of lives mainly on women and children. The traditional method of collecting surveys was valuable but often left behind and short in efficiency causing unable to reach the stakeholders to overcome this issue as soon as possible (Ahmed & Kamalakkannan, 2022). Everyone focusing on children, women, elderly people are suggested to have a proper dietary plan with enough nutrition on it which needs in daily basis.

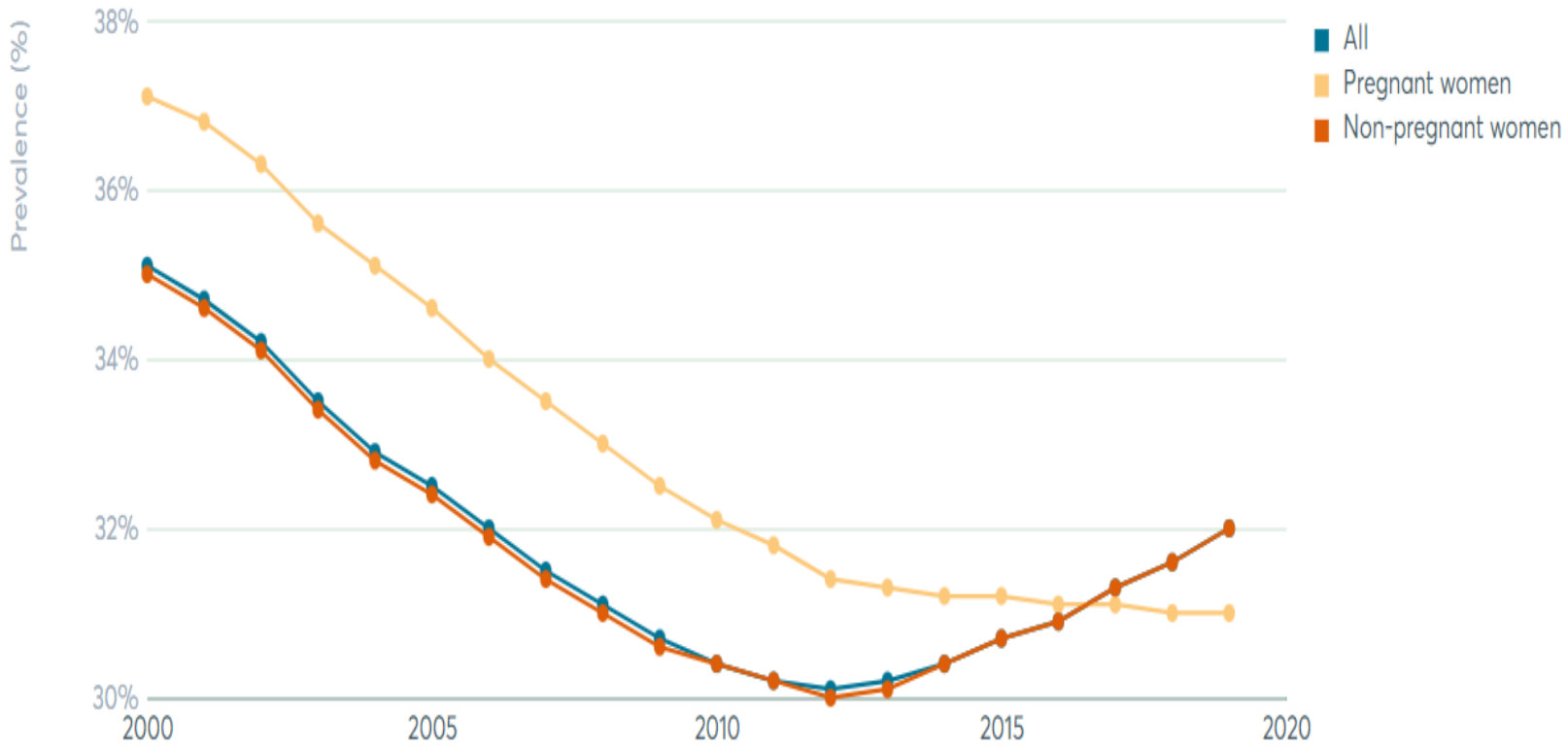

Figure 1 shows the decrease in value of nutrition intake over the year by targeted people according to the research by (Khudri et al., 2023)

This limitation has catalyzed this research into machine learning prediction as a potential step to find out the solution and analyze further on these mitigate nutrient deficiencies. Machine Learning provides a transformative approach to enhance and overcome this prediction analysis and intervention strategies in combating nutritional deficiencies. It is important to all age group to compulsory have their design nutrition pattern to avoid high risk in health issues. Due to some factors such as socioeconomics, income, environmental factors several places were constrained to have proper nutrition diet in their daily intake. The application of machine learning (ML) techniques to predict and analyse nutritional deficiencies in different areas.(Kananura, 2022) By combining various data sets on health conditions, socioeconomic, environmental and dietary factors this paper goals to build model predictive models that can list out regions that are at high risk for deficiencies. Machine learning has the potential to define complex, non-linear patterns within large datasets, which been using often in traditional methods(Zhou et al., 2019). These insights could help improve public health outcomes. Nutritional deficiencies have impacted the rural regions leading to serious health outcomes and changes in human developments. The challenge can predict the risk in nutritional deficiencies across various region by analysing diverse datasets that include health, dietary, socioeconomic and environmental factors. The aim is to develop a model with help of machine learning to identify the high-risk areas and the factors that has been contributing to the nutritional deficiencies.

Age Group	RNI 2005 (µg/Day)	RNI 2005 (IU/Day)	RNI 2017 (µg/Day)	RNI 2017 (IU/Day)	Upper Limit (µg/Day)
Infants					
0–5 months	5	200	10	400	25
6–11 months	5	200	10	400	37.5
Children					
1–3 years	5	200	15	600	
4–6 years	5	200	15	600	100
7–9 years	5	200	15	600	
Boys					
10–18 years	5	200	15	600	100
Girls					
10–18 years	5	200	15	600	100

Figure 2 above shows the findings of a proper nutrition intake that an infant should at least be taking to avoid health issues in early stages.(Jain et al., 2022)

2.3 Application of Machine Learning in Healthcare Industry

Machine Learning has always proven as a transformative element in the healthcare sector, offering robust frameworks to analyze large and complex data. Surveys, papers and studies have visualized the application of machine learning in determining disease outbreaks, medical treatments and optimizing healthcare delivery systems. (Ali et al., 2022) mentioned that few ML models such as supervised learning models, including decision trees and vector machines have been successfully used in predicting anemia in certain populations to showcase the feasibility of such technologies in nutrition focused studies. These methods have always been effective in handling structured data where labeled examples are available enabling accurate prediction of outcomes like malnutrition or any other specific deficiencies. Advanced machine learning techniques such as shown in table 1 below are particularly successful due to their robustness and ability to model complex interactions among different variables.

Supervised Learning Approaches

Studies have used the supervised leaning technique to predict the deficiency model for this nutritional deficiency in each region. Logistic regression and decision trees have been utilized to predict the anemia prevalence using the income list from each household together with their dietary plan. (Mpakairi et al., 2023) proposed random forest and support vector machines have shown robustness in dealing with complex, non-linear relationships between variables in nutritional deficiency predictions.

A notable study by Ling et al. (2020) applies gradient boosting models to predict the vitamin deficiency in achieving high accuracy by studying dietary intake data along with the region datasets taken from and with regional climate variables. The study highlighted the importance of feature selection and preprocessing in improving a prediction model for future references. Singh eth al. (2019) applied these clustering techniques to demographic and health surveys (DHS) data, showing the regional issues in malnutrition across the Sub-Saharan Africa.(Khudri et al., 2023)

Deep Learning Techniques

These model types, especially (CNNs) convolutional neural networks and recurrent neural networks (RNNs), have been explored for malnutrition predictions. This model excels in preprocessing large scale of date sets and complex inputs including imagery and time-series data sets. These been used to predict the productivity of the region according to their socioeconomic directly linking to the regional dietary deficiencies.

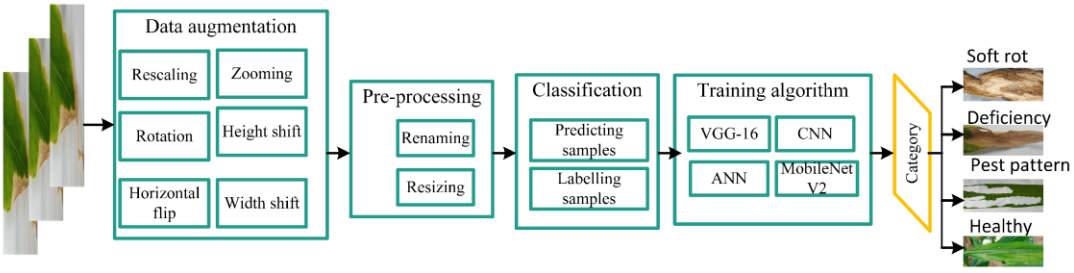

Figure 3 shows the proposed method of deep learning techniques used by (Velásquez et al., 2020) to identify the nutritional deficiencies among plantations.

Unsupervised Learning and Clustering

Unsupervised Learning methods such K-means clustering and hierarchical clustering, have been used to group regions with similar profiles. This approach helps in identifying the high-risk areas and underlying patterns in nutritional deficiency. Moreover, using unsupervised algorithm helps to demonstrate the pattern of the deficiency issue on the region. These patterns enhance the monitoring for the deficiency trend.

Supervised Learning	Unsupervised Learning	Semi-Supervised Learning	Reinforcement Learning
Models are built according to the definitions and findings	Deep Learning is used to arrive at the conclusions and patterns through unlabeled data.	Build model through a mix of labeled and unlabeled data.	Self-interpreting is based on a system learned through trial and error seeking findings.
• Linear Regressions • Support Vector Machines • Decision Trees	• Apriori • K-means clustering • Artificial Neural networks	• Generative networks • Naïve bayes Classifier	• Q-learning • Model based estimation
These algorithms are used to demonstrate risk assessment, predictive analysis, and image classification.	These algorithms are suitable for describing the functions, perform monitoring, data mining and pattern recognition.	These algorithms describe more about manipulation, data visualization and natural language processing.	These algorithms are more self-paced prediction modelling which need linear tasks and estimating parameters.

Table 1 shows the differences between various machine learning algorithms.(Gollapalli, 2022)

2.4 Datasets and Key Factors in Prediction Models

The quality and diversity of a datasets is important and critical in developing a prediction machine learning model. A real time data sets gives more accurate value to the scenarios as it will be taken from real data source. (Meshram et al., 2021) says the traditional way of prediction model has been taken with the use of comprehensive surveys and feedback that provide demographic, health, and nutrition data. Datasets also collected on food supply and consumption organizations to know the amount of food been consumed and distributer to each region. Datasets plays crucial role in finding to justify the findings with proof and evidence.

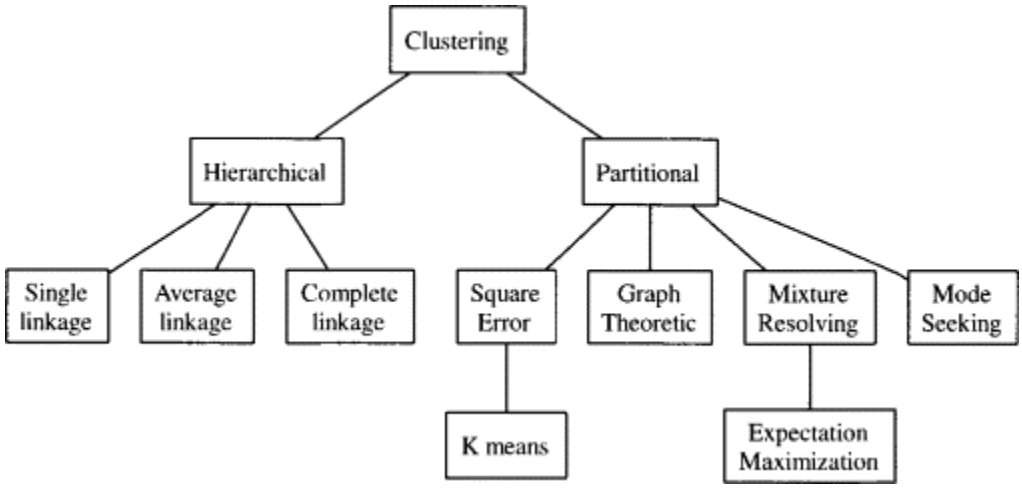

Figure 4 explains how clustering method works in mode prediction using machine learning algorithms.

2.5 Innovation in Nutritional Deficiency Prediction

Developing standardized and comprehensive datasets for nutritional deficiency prediction helps to monitor more in details which could leave the stakeholders to a proper decision making. Integrating explainable data sets along with machine learning techniques will enhance the model interpretability. Incorporating real time data streams improve the accuracy in the findings which can justify the prediction model. Exploring hybrid models that combines both the statical and machine learning techniques approaches will lead to a better generalization. In this paper the provided data driven predictions of nutritional deficiencies across regions enables the healthcare providers and stakeholders to identify the risk in the populations efficiently. It's also able to highlight the main key factors that been affecting to the nutrition deficiencies and help to reduce it. This project contributes by providing evidence based on live data to guide and develop for long term strategies to combat malnutrition. All these contributions are aim to create lasting impact and provide practical solutions for improving health.

References	Experiment	Strength	Limitations
(Ahmed & Kamalakkannan, 2022)	Micronutrient classification in IoT based agriculture using machine learning (ML) Algorithm	• Machine learning algorithms are used to predict the pattern for the micronutrient	• Prediction pattern only shows the future effects.
(Ali et al., 2022)	Machine Learning Approaches for Prediction of Nutrition Deficiency among Women of Different Age Groups	• Able to find the root cause of deficiency mostly among women with prediction models	• Unable to visualize the findings to the providers
(Gollapalli, 2022)	Ensemble machine learning model to predict the waterborne syndrome	• Elaborates the factors and demonstrate the solutions with a prediction model	• No accurate data shown to justify the prediction model.
(Jain et al., 2022)	Efficient Machine Learning for Malnutrition Prediction among under-five children in India	• Stakeholders manage to visualize the future effects from the	• Visualizations are well delivered.

		predictions model created	
(Kananura, 2022)	Machine learning predictive modelling for identification of predictors of acute respiratory infection and diarrhoea in Uganda's rural and urban settings	• Algorithms are used to identify the accurate cause of infections	• No real time data are shown to have a clear view of the accuracy
(Khudri et al., 2023)	Predicting nutritional status for women of childbearing age from their economic, health, and demographic features: A supervised machine learning approach	• Random forests are used to gather the data from various variable and telecast the outcome in one prediction model	• Noisy data should have been cleaned.
(Qasrawi et al., 2024)	Machine Learning Approach for Predicting the Impact of	• Demonstrate risk assessment and perform	• Risk assessment are predicted without

	Food Insecurity on Nutrient Consumption and Malnutrition in Children Aged 6 Months to 5 Years	monitoring form the prediction theory	solutions to overcome it.
(Velásquez et al., 2020)	A method for detecting coffee leaf rust through wireless sensor networks, remote sensing, and deep learning	• Integrates the data finely with the algorithms and transfer into machine learning concepts to visualize it	• Data are finely integrated and but not real time data are used to prove the findings
(Zhou et al., 2019)	Combining optical, fluorescence, thermal satellite, and environmental data to predict county-level maize yield in China using machine learning approaches	• Provides image classification • Predictive analysis demonstrates the clear view of classification.	• Still exist irrelevant image with not accurate to the findings

(Zhang et al., 2019)	Application of deep learning in food	• Deep learning techniques are used to filter out the main cause of food deficiency	• Fewer number of clusters

Table 2 shows the summary of findings from various resources and research paper s used in this literature review

2.6 Research Gap

Despite significant advancements, several research gaps are to be remain in the application of machine learning for nutritional deficiency prediction. The region data will be limited according to several factors. Including many regions even though will enlarge the findings there will be lack in producing high quality predictions especially involving low in come Ares(Zhang et al., 2019). Limiting model development and validation can enhance the precision in the findings(Qasrawi et al., 2024). This research will be focussing on few deficiencies only limiting to the main five important nutrition deficiencies to helps cover the overall nutrition deficiencies. Most of the studies concentrate on one type of deficiency, while comprehensive models addressing multiple deficiencies are scarce. While multimodal approaches exist it is important to verify that diverse datasets are remain underexplored to interpret and study while encountering this research. Most of the datasets from different regions lead to fail vales due to differences in their culture and economic state or dietary plans. Addressing these gaps earlier before carrying out the research will require collaborative efforts between researchers, policymakers and data providers to enhance the utility of machine learning based solution.

2.7 Summary

Machine learning holds crucial role for prediction nutritional deficiencies by region, offering a data driven approach to addressing global health disparities. By interpreting diverse data sets and advanced algorithms, machine learning leads the understanding on how important the nutrition is related to our human organizations and helps to support the target people by region. However, addressing challenges related to data quality, interpretability, and ethical concerns is significant for a complete implementation of these technologies in health policies. This chapter has included the literature review of this ongoing research regarding the prediction model based on nutrition deficiency with the findings of previous work. Apart from it, it also explains the machine learning algorithm will be used throughout this research.

References

Ahmed, G. N., & Kamalakkannan, S. (2022). Micronutrient classification in IoT based agriculture using machine learning (ML) Algorithm. *2022 4th International …*. https://ieeexplore.ieee.org/abstract/document/9716293/

Ali, J., Khan, M. A., & Khan, N. A. (2022). Machine Learning Approaches for Prediction of Nutrition Deficiency among Women of Different Age Groups. *2022 3rd International Conference …*. https://ieeexplore.ieee.org/abstract/document/10100340/

Gollapalli, M. (2022). Ensemble machine learning model to predict the waterborne syndrome. In *Algorithms*. mdpi.com. https://www.mdpi.com/1999-4893/15/3/93

Jain, S., Khanam, T., Abedi, A. J., & ... (2022). Efficient Machine Learning for Malnutrition Prediction among under-five children in India. *2022 IEEE Delhi Section …*. https://ieeexplore.ieee.org/abstract/document/9753080/

Kananura, R. M. (2022). Machine learning predictive modelling for identification of predictors of acute respiratory infection and diarrhoea in Uganda's rural and urban settings. In *PLOS Global Public Health*. journals.plos.org. https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0000430

Khudri, M. M., Rhee, K. K., Hasan, M. S., & Ahsan, K. Z. (2023). Predicting nutritional status for women of childbearing age from their economic, health, and demographic features: A supervised machine learning approach. In *Plos one*. journals.plos.org. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0277738

Meshram, V., Patil, K., Meshram, V., Hanchate, D., & ... (2021). Machine learning in agriculture domain: A state-of-art survey. In *… Intelligence in the Life …*. Elsevier. https://www.sciencedirect.com/science/article/pii/S2667318521000106

Mpakairi, K. S., Dube, T., Sibanda, M., & Mutanga, O. (2023). Fine-scale characterization of irrigated and rainfed croplands at national scale using multi-source data, random forest, and deep learning algorithms. In *ISPRS Journal of …*. Elsevier. https://www.sciencedirect.com/science/article/pii/S0924271623002460

Qasrawi, R., Sgahir, S., Nemer, M., Halaikah, M., & ... (2024). Machine Learning Approach for Predicting the Impact of Food Insecurity on Nutrient Consumption and Malnutrition in Children Aged 6 Months to 5 Years. In *Children*. mdpi.com. https://www.mdpi.com/2227-9067/11/7/810

Velásquez, D., Sánchez, A., Sarmiento, S., Toro, M., & ... (2020). A method for detecting coffee leaf rust through wireless sensor networks, remote sensing, and deep learning: Case study of the Caturra variety in Colombia. In *Applied Sciences*. mdpi.com. https://www.mdpi.com/2076-3417/10/2/697

Zhang, L., Zhang, Z., Luo, Y., Cao, J., & Tao, F. (2019). Combining optical, fluorescence, thermal satellite, and environmental data to predict county-level maize yield in China using machine learning approaches. In *Remote Sensing*. mdpi.com. https://www.mdpi.com/2072-4292/12/1/21

Zhou, L., Zhang, C., Liu, F., Qiu, Z., & ... (2019). Application of deep learning in food: A review. *… in Food Science and Food …*. https://doi.org/10.1111/1541-4337.12492