SEMARAK ILMU
PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Obesity Predictor Identification: Comparison of Correlation Based Feature Selection Method and Wrapper Method on Nutrition Dataset

Nur'aina Daud [1,*], Nurulhuda Noordin[1,♣] , Anitawati Lokman[1]

[1] College of Computing, Informatics and Mathematics, School of Computing Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 20 June 2023
Received in revised form 14 April 2024
Accepted 15 June 2024
Available online 25 July 2024

ABSTRACT

The prevalence of obesity among Malaysians is estimated by calculating the obesity prevalence percentage using BMI prevalence data from the national health morbidity survey (NHMS). However, the nutrition data from the NHMS has not been used to predict the national obesity prevalence as it was collected solely for the documentation of an analysis report on the food consumption patterns of the base population. To address this gap, this study utilises nutrition data by employing 15 nutrition variables derived from grocery data to predict obesity. This paper seeks to identify the appropriate nutrition variable, which involved exploring 8238 rows of raw grocery data (grocery receipt) collected from 35 households. During the data pre-processing phase, 15 nutrition variables were generated in the data conversion and data transformation phase of the data pre-processing phase of this study. This study predicts the percentage of selected nutrition variables that could lead to obesity in individuals. The purpose of this study is to find alternative data (grocery data) that can be used to predict obesity and to test the relevance of using that alternative data in predicting obesity by evaluating the accuracy performance measurement of the prediction through the use of data mining technology. This study predicts the percentage of macronutrients variables that could lead to obesity in individuals. To simplify the prediction model, the dataset variables were filtered using the automated feature selection method in the WEKA machine learning tool version 3.8. The objective of the feature selection performance of variables from the dataset was to identify the nutrition variables that have the most significant impact on developing accurate prediction models by evaluating the accuracy performance of the model using area under curve score (AUC). The generated nutrition dataset was subjected to the subset method known as correlation-based-feature-selection (CFS) and wrapper methods that included a learning algorithm in the attribute selection process. Several subsets were extracted during the feature selection phase, which served as potential input datasets (predictor) for developing obesity prediction models using different classification algorithms. Based on the feature selection evaluation conducted in this study, the CFS method was found to be the best feature selection method compared to the three wrapper methods conducted, which resulted in the selection of calorie_intake and foodpyramid_level3%

[*] *Corresponding author.*
E-mail address: ainadaud@uitm.edu.my
[♣] *Corresponding author.*
E-mail address: drnurul@uitm.edu.my

https://doi.org/10.37934/araset.49.1.129138

Keywords:
Classification Algorithm; Correlation Based Feature Selection; Feature Selection; Obesity Prediction; Wrapper Method

variables as the appropriate predictors for this study. These results can enhance the reliability of using household grocery data to predict obesity and open new avenues for research into nutrition and health prediction.

1. Introduction

The motivation behind this study is to emphasise the importance of data analytics using data mining tools to capture dietary intake and caloric consumption from grocery data. To achieve this, we have opted to use the WEKA data mining tool to conduct an exploratory study that investigates the suitability of generated nutrition variables from grocery data to predict obesity, and which variables can be used as proxies for obesity prediction. Our approach involves the use of feature selection in WEKA, which enables us to identify the most significant variables or parameters that contribute to predicting the outcome [1]. We acknowledge that not all variables are pertinent to this process. Therefore, we employed feature selection during the model development phase to select a subset of predictors that can be used to construct a simplified model with good predictive power.

In the feature selection process, the effectiveness of all variables was evaluated. The resulting sets of predictors from the outcomes were then utilized as input datasets for subsequent model development. As noted by Li *et al.,* [2], feature selection confers three key advantages, namely

i. Reduces over-fitting: Having less redundant data means there is less opportunity to make decisions based on noise.
ii. Improves accuracy: Fewer misleading data points enhance the accuracy of the model.
iii. Reduces training time: With less data, algorithms are trained at a faster rate.

The objective of assessing the automated feature selection performance of variables from the dataset is to identify those variables that have the most significant impact on constructing reliable prediction models. The selected variables (predictors) were subsequently used as the input dataset in this study for model development, and a summary of all variable descriptions is presented in Table 1.

Table 1
Variables in Nutirition Dataset

#	Variables	Description
1	calorie_intake	The individual estimated calorie intake
2	carbohydrate_intake%	The individual estimated percentage of carbohydrate intake of
3	protein_intake%	The individual estimated percentage of protein intake
4	fat_intake%	The individual estimated percentage of fat intake
5	carbohydrate_intake(g)	The individual estimated carbohydrate intake (gram)
6	protein_intake(g)	The individual estimated protein intake (gram)
7	fat_intake(g)	The individual estimated fat intake (gram)
8	foodpyramid_level1%	The individual estimated percentage of food from level 1 food pyramid intake
9	foodpyramid_level2%	The individual estimated percentage of food from level 2 food pyramid intake
10.	foodpyramid_level3%	The individual estimated percentage of food from level 3 food pyramid intake

Within WEKA, the feature selection method utilizes a specific search method to identify a set of attributes. The goal of this study was to determine the optimal feature selection method, and the performance of each method was measured based on the area under curve (AUC) performance metric, as detailed in the methods section of this paper. The selection of the best feature selection method is vital in identifying the most suitable nutrition variable or predictor to be used in the further development of an obesity prediction model.

2. Literature Review

The ever-growing and complex nature of medical data has led to an increase in data dimensionality, resulting in irrelevant, redundant, and noisy attributes. In data mining (DM), constructing a model with a large number of attributes may lead to a reduction in predictive power [3], referred to as the "curse of dimensionality," as coined by Hughes [4]. The curse of dimensionality refers to the sparse nature of data as the volume of data increases, resulting in inefficient predictive power. However, reducing the number of features or attributes is a viable method of reducing data dimensionality in DM without compromising its objectives. Previous studies have demonstrated that selecting a significant set of attributes aids in data visualization, data understanding, reduces over-fitting, and improves overall prediction performance [5]. Furthermore, reducing attributes results in a faster training time and a simpler model. Fast training time is crucial when dealing with vast datasets, and a simpler model provides a deeper understanding of the underlying processes that generate the data.

In classification modelling, the feature selection method can be classified into three categories: 1) the subset method, and 2) the wrapper method. The correlation-based-feature-selection (CFS) algorithms belong to the subset approach, which identifies a subset of features with high correlation with the class, but low correlation with each feature [6].

On the other hand, the wrapper method incorporates a learning algorithm into evaluating the selection of features [7]. This method offers the advantages of the subset feature selection and the specific classifiers. However, compared to the filter method, the wrapper method has a higher computational cost due to the additional evaluation of the subset of features with the specific learning algorithm. Moreover, the wrapper method tends to over-fit the learning algorithm used to evaluate the subset of features. Therefore, it is recommended to develop the model on other classification algorithms instead of using the algorithm used for feature selection.

3. Methods

3.1 Feature Selection Methods in WEKA

The feature selection methods employed on the generated nutrition dataset were the subset and wrapper methods.

3.1.1 Correlation based feature selection (CFS) method

In this study, the subset method used was the correlation-based-feature-selection (CFS) algorithm, which selects attributes with high correlation with the class and low correlation with each other [6]. In WEKA, this subset method is known as CfsSubsetEval. The study also employed a filter feature selection method, which is a type of filter method that reduces the number of attributes during the pre-processing steps before running the dataset into any classifier algorithm. This method selects subsets of attributes using properties of the data itself, independently of any learning algorithm [8].

Two types of filter methods are: 1) univariate-filter, and 2) multivariate-filter (subset) methods. The study adopted the latter method, which considers the relationship of individual attributes as well as the correlation between attributes towards the outcome. One benefit of applying a filter method is that it reduces the number of features used during the final induction algorithm, leading to improved classification algorithm performance and reduced computer processing time. Filter methods are also independent of the final learning algorithm used and the same features may be used in different learning algorithms for comparative analysis. Unlike wrapper methods, filter methods do not incorporate the final learning algorithm in their process, which is another benefit of using filter methods [9].

3.1.2 Wrapper method

The wrapper method differs from the filter method in that it involves using a learning algorithm for attribute selection. For this study, two classification algorithms, BN, NB, and JRip were selected for the wrapper method. These algorithms were chosen due to their strong predictive performances in a previous task, where their average AUC scores exceeded 0.65 as shown in Table 2.

In WEKA, a specific search method is used to determine the set of attributes for feature selection. This study employed the Greedy search strategy with the forward selection approach. The Greedy search method works by progressively adding attributes to the subset until the best subset is obtained. It is a computationally efficient and robust method that helps prevent over-fitting, which can occur when a model attempts to predict a trend in noisy data. Over-fitting arises due to a complex model with too many parameters [5].

The feature selection method in WEKA utilises a particular search method to determine a set of attributes. For this study, the Greedy search strategy was employed, using the forward selection approach. This Greedy search strategy gradually builds a subset of attributes by adding them one by one until the best subset is identified. The Greedy search method has been noted for its computational efficiency and robustness against over-fitting, which occurs when a model attempts to predict a trend in data that is too noisy. This issue arises due to an overly complex model with too many parameters [6].

3.2 Classification Algorithms Selection

Prior to conducting the feature selection process, an evaluation of classification algorithms was carried out using all variables in the dataset to identify suitable and unsuitable algorithms for the nutrition dataset. Using classification algorithms in prediction model development is crucial because they ensure accurate categorization of data, enable generalization to unseen data, provide interpretable models, handle large datasets efficiently, offer flexibility across different data types and problem domains, identify important features for prediction, and create robust models less sensitive to noise or outliers.

To select the suitable algorithms to be used in model development, this study considers the algorithm that creates the model that achieved AUC > 0.65 using nutrition dataset. The unsuitable classification algorithms for the nutrition dataset generated in this study are eliminated, while the rest are used for model development. Those algorithms are Bayes net (BN), naïve Bayes (NB), simple logistic (SL), decision table and naïve Bayes (DTNB), repeated incremental pruning to produce error reduction (JRip), projective adaptive resonance theory (PART), decision stump (DS) and C4.5 decision tree (J48).

In turn, a total of 19 DM classification algorithms, each with a default parameter setting available in WEKA, were assessed. These algorithms originated from distinct basic learning concepts, including naive Bayes, linear/non-linear, SVM, neural networks, instance-based rules, and tree models [7]. All

19 selected algorithms support the classification task, with some algorithms also supporting description capabilities, enabling patterns to be understood by humans.

The evaluated DM classification algorithms are grouped into six basic classifiers, as shown in Table 2, including naïve Bayes, linear models/non-linear, SVM, neural networks, rules, and tree-based. Algorithms under the NB learner use classical statistical theory, i.e., Bayes theorem from John *et al.,* [8], as the basis of the algorithm. The LG algorithm in WEKA uses regression technique with a ridge estimator [10]. On the other hand, SVM algorithm uses hyper-plane classifiers, simple linear machines on which SVMs are based, to determine the best separation for the classes [11]. Neural network is a learner that uses the basis of human brain interactions in processing and understanding relationships [12]. It can create simulations and predictions for complex systems and relationships, such as in weather forecasting, medical diagnostics or business processes [13].

The last two classifiers, rules and tree-based learners are based on divide-and-conquer approach which normally work on top-down manner. At each stage, the best identified attribute is split into classes, and recursively process the sub problems resulted from the split. Unlike decision tree, rule-based learner comes with a rule in selecting the instances at each stage. Thus, the rule-based learner will lead to a set of rules rather than a decision tree [14]. Different rules, different splitting methods and different pruning strategies (to reduce number of nodes in a tree) differentiate the algorithms under rule and decision tree learners.

Using all nutrition data in the dataset, the algorithms that achieved AUC > 0.65 using nutrition dataset has been identified. Thus, from the evaluation, there are eight algorithms have been found as suitable algorithms to be used in further model development. Those algorithms are Bayes net (BN), naïve Bayes (NB), simple logistic (SL), decision table and naïve Bayes (DTNB), repeated incremental pruning to produce error reduction (JRip), projective adaptive resonance theory (PART), decision stump (DS) and C4.5 decision tree (J48).

Table 2
List of evaluated modelling algorithms

Basic Classifier	WEKA Modelling Algorithms
Naïve Bayes	1. Bayes net (BN)
	2. Naïve Bayes (NB)
Linear models/ non-linear	3. Logistic (LG)
	4. Simple logistics (SL)
SVM	5. SMO (SVM)
Neural networks	6. MultiLayer perceptron
	7. Voted perceptron (VP)
Rules	8. Decision tables and naïve Bayes (DTNB)
	9. Repeated incremental pruning to produce error reduction (JRip)
	10. OneRule (OneR)
	11. Projective adaptive resources theory (PART)
	12. Zero (ZR)
Tree based	13. Decision stump (DS)
	14. Hoefding tree
	15. C4.5 decision tree (J48)
	16. Logistic model trees (LMT)
	17. Random forest (RF)
	18. Random tree (RT)
	19. REPTree (REPT)

The last two classifiers, rules and tree-based learners are based on a divide-and-conquer approach that typically operates in a top-down manner. At each stage, the best attribute is identified and split into classes, and the resulting sub-problems are processed recursively. Unlike a decision tree, a rule-based learner selects instances at each stage based on a specific set of rules. Therefore, the rule-based learner produces a set of rules instead of a decision tree [14]. In contrast, Boruah *et al.,* [15] highlights that to construct rules from the unique set of split conditions, the resulting rule may have combinations of conditions that may not exist in any of the trees. The algorithms under rule and decision tree learners are distinguished by their different rules, splitting methods, and pruning strategies (used to reduce the number of nodes in a tree). The AUC scores for the evaluated algorithms are tabulated in Table 3.

In order to select the most suitable algorithms for model development, this study considered those that produced a model with an AUC > 0.65 using the nutrition dataset. Unsuitable classification algorithms for the nutrition dataset were eliminated, while the remaining algorithms were used for model development. Using all nutrition data in the dataset, the study identified the model that achieved an AUC > 0.65. From the evaluation, a total of eight algorithms were found to be suitable for further model development. These algorithms include Bayes net (BN), naïve Bayes (NB), simple logistic (SL), decision table and naïve Bayes (DTNB), repeated incremental pruning to produce error reduction (JRip), projective adaptive resonance theory (PART), decision stump (DS), and C4.5 decision tree (J48), as shown in Table 3. The check symbol in the right column of the table indicates the suitable algorithms.

Based on their AUC scores, five algorithms with a score of 0.65 and below were deemed 'unsuitable'. These include OneRule (OneR), ZeroR (ZR), Hoeffding tree, random tree (RT), and REPTree (REPT) which have obtained AUC < 0.5. Additionally, SMO (SVM) and random forest (RF) were also considered 'unsuitable' as their AUC scores were below 0.6. Finally, VP produced fluctuating AUC scores for the nutrition datasets.

Table 3
AUC Scores for Evaluated Algorithms

Basic Classifier	VEKA Modelling Algorithms	AUC Score	AUC Score > 0.65
Naïve Bayes	1. Bayes net (BN)	0.78	✓
	2. Naïve Bayes (NB)	0.79	✓
Linear models/ non-linear	3. Logistic (LG)	-	
	4. Simple logistic	0.70	✓
SVM	5. SMO (SVM)	0.58	
Neural networks	6. MultiLayer perceptron	-	
	7. Voted perceptron (VP)	0.63	
Rules	8. Decision tables and naïve Bayes (DTNB)	0.74	✓
	9. Repeated incremental pruning to produce error reductio (JRip)	0.76	✓
	10. OneRule (OneR)	0.48	
	11. Projective adaptive resonance theory (PART)	0.71	✓
	12. Zero (ZR)	0.47	
Tree- based	13. Decision stump (DS)	0.71	✓
	14. Hoeffding tree	0.47	
	15. C4.5 decision tree (J48)	0.71	✓
	16. Logistic model trees (LMT)	-	
	17. Random forest (RF)	0.52	
	18. Random tree (RT)	0.50	
	19. REPTree (REPT)	0.47	

According to Muthu and Palaniappan [16], logistic (LG) consistently performs well when tested with different data sets and cross-validation, showing that it's reliable for predicting stomach cancer.

Despite being a promising algorithm, logistic (LG) was deemed 'unsuitable' as it produced an average AUC score below 0.65 when tested using nutrient dataset to predict obesity. Additionally, two algorithms, MultilayerPerceptron and VP, were unable to run the testing on the nutrition dataset, indicated by the '-' symbol in Table 3. As a result, only eight algorithms were found to be suitable for further evaluation and development. These suitable algorithms are listed in Table 4.

Table 4
The selected classification algorithms

Variables	Font size and style
Naïve Bayes	1. Bayes etn (BN)
	2. Naïve Bayes (NB)
Linear regression	3. Simple logistics (SL)
Rules	4. Decision table
	5. Repeated incremental pruning to produce error reduction (JRip)
Tree based	6. Projective adapted resonance theory (PART)
	7. Decision stump (DS)
	8. C4.5 decision tree (J48)

4. Results and Discussion

4.1 Evaluation of Feature Selection Methods

The selection and filtration of variables for each dataset were performed using the CFS and Wrapper feature selection methods. In the Wrapper method, BN, NB, and JRip algorithms were selected for use with wrapper feature selection methods due to their AUC scores exceeding 0.75 shown in the previous analysis. The subsets of predictors chosen by applying CFS FilterSubset, wrapper with BN algorithm (WrapperBN), wrapper with NB algorithm (WrapperNB), and wrapper with JRip algorithm (WrapperJrip) methods on datasets are tabulated in Table 5.

Table 5
Subset of predictors selected by DM automated feature selection

Automated feature selection methods	List of Selected Variables (Predictors)
CFS	1. Calorie_intake
	2. Foodpyramid_level3%
WrapperBN	1. Calorie_intake
	2. Fat_intake%
	3. Foodpyramid_level3%
	4. Raw intake%
WrapperNB	1. Calorie_intake
	2. Fat_intake(g)
	3. Foodpyramid_level2%
	4. Raw intake%
WrapperJRip	1. Calorie_intake
	2. Raw_intake%

The results of all DM feature selection methods for the dataset showed that the combination of calorie_intake was a potential predictor. This variable was selected by all four feature selection methods, indicating its importance in predicting obesity. From Table 5, it can be observed that WrapperNB had the most predictors, while the remaining methods selected only two variables as predictors.

The predictors selected by the feature selection methods were used to develop prediction models using the eight chosen learning algorithms. The performance of these models was evaluated using a

10-fold cross-validation technique and a supplied test set (holdout) technique, as explained in the following sub-section. The results of these evaluations for the 5-month dataset are shown in Table 6.

In Table 6, CFS_Nutrition, WR_BN_Nutrition, WR_NB_Nutrition, and WR_JRip_Nutrition represents the models developed using CFS and wrapper methods (WrapperBN, WrapperNB, and WrapperJRip), respectively. It can be observed that NB is the best algorithm as it achieved an AUC > 0.8 in all feature selection methods. For CFS_Nutrition, WR_BN_Nutrition, and WR_JRip_Nutrition, BN, NB, and SL were among the best three algorithms producing the highest AUC scores. CFS_Nutrition exhibited the best performance for the nutrition dataset, obtaining the highest average AUC score (AUC = 0.760). NB produced the highest AUC score for CFS_Nutrition (AUC = 0.840). On the other hand, models developed using predictors selected by WrapperNB had the lowest average AUC score in most algorithms, with an average score of AUC = 0.749. However, the difference in the average score for WR_NB_Nutrition was not significant compared to models developed using predictors selected by the other two wrapper methods (WrapperBN and WrapperJRip), with an average AUC score of 0.752 and 0.759, respectively.

Table 6
The result of AUC score generated from each set of predictors extracted from automated feature selection methods

Algorithms	CFS_ Nutrition	WR_BN_ Nutrition	WR_NB_ Nutrition	WR_JRip_ Nutrition
BN	0.807	0.825	0.821	0.787
NB	0.840	0.833	0.829	0.839
SL	0.811	0.780	0.762	0.838
DT	0.776	0.776	0.776	0.776
JRip	0.723	0.711	0.715	0.740
PART	0.703	0.686	0.685	0.690
DS	0.719	0.719	0.719	0.719
J48	0.703	0.685	0.685	0.690
Average AUC Score	0.760	0.752	0.749	0.759

4. Conclusion

This paper documents the feature selection process conducted to choose a subset of variables from the grocery-generated nutrition variables as predictors in the development of an obesity prediction model. From the feature selection analysis shown in this paper, it can be concluded that, the CFS method proved to be the most effective feature selection approach compared to the three wrapper methods employed in this study, resulting in the selection of calorie_intake and foodpyramid_level3% variables as suitable predictors. While models developed using the same algorithms used by wrapper feature selection methods (i.e., BN, NB, and JRip) were found to exhibit some bias, the differences were not significant. Therefore, the input dataset for further development of obesity prediction models using nutrition data was selected using the CFS feature selection method.

Based on the findings of this study regarding suitable nutrition variables, the next step will be to utilize them as proxy measures or predictors in the development of the grocery to nutrition obesity prediction (G2NOP) Modelling for obesity prediction. A well-defined flow structure is crucial to ensure the relevance and applicability of the proposed model. Furthermore, the model will be integrated into the obesity prediction process framework (G2NOPF), which serves as a conceptual foundation for the implementation of the alternative method of obesity prediction suggested in this study. This framework can also be used as a guide for the Ministry of health (MOH) by illustrating the prediction process flow and the interrelatedness of the framework processes.

This study has limitations and constraints that were considered and outlined based on several factors. Firstly, the nutrition assumption in this study was that everyone in a household consumes the foods and groceries they bought, and individuals with higher BMIs consume more compared to those with lower BMIs. This assumption could affect the accuracy of the results obtained. Another limitation is the use of self-report BMI data collection, where respondents were asked to self-report their body weight and height [17]. While this method was used to collect data, respondents may have under or over-reported their measurements. The study focused only on the correlation between grocery data and obesity, and thus, only nutrition data and anthropometric data (BMI) were considered as manipulated variables. Other obesity contributing factors, such as physical activity, family genetics, and health conditions, were considered as constant variables. Additionally, the study used purposive sampling to select respondents who purchased grocery foods for their home food inventories, making the study not applicable to households where all members frequently eat outside.

Acknowledgement

The registration fee is funded by Pembiayaan Yuran Penerbitan Artikel (PYPA), Tabung Dana Kecemerlangan Pendidikan (DKP), Universiti Teknologi MARA (UiTM), Malaysia.

References

[1] Sheena, K. Kumar, and Gulshan Kumar. "Analysis of feature selection techniques: A data mining approach." In *International Conference on Engineering & Technology*, vol. 4, pp. 17-21. 2016.

[2] Li, Jundong, Kewei Cheng, Suhang Wang, Fred Morstatter, Robert P. Trevino, Jiliang Tang, and Huan Liu. "Feature Selection: A Data Perspective." *arXiv e-prints* (2016): arXiv-1601. https://doi.org/10.1145/3136625

[3] Nisbet, Elizabeth K., John M. Zelenski, and Steven A. Murphy. "The nature relatedness scale: Linking individuals' connection with nature to environmental concern and behavior." *Environment and behavior* 41, no. 5 (2009): 715-740. https://doi.org/10.1177/0013916508318748

[4] Hughes, Gordon. "On the mean accuracy of statistical pattern recognizers." *IEEE transactions on information theory* 14, no. 1 (1968): 55-63. https://doi.org/10.1109/TIT.1968.1054102

[5] Guyon, Isabelle, and André Elisseeff. "An introduction to variable and feature selection." *Journal of machine learning research* 3, no. Mar (2003): 1157-1182. https://doi.org/10.1016/j.aca.2011.07.027

[6] Hall, Mark A., and Lloyd A. Smith. "Practical feature subset selection for machine learning." (1998): 181-191.

[7] Kohavi, Ron, and George H. John. "Wrappers for feature subset selection." *Artificial intelligence* 97, no. 1-2 (1997): 273-324. https://doi.org/10.1016/S0004-3702(97)00043-X

[8] John, George H., Ron Kohavi, and Karl Pfleger. "Irrelevant features and the subset selection problem." In *Machine learning proceedings 1994*, pp. 121-129. Morgan Kaufmann, 1994. https://doi.org/10.1016/B978-1-55860-335-6.50023-4

[9] Ladha, L., and T. Deepa. "Feature selection methods and algorithms." *International journal on computer science and engineering* 3, no. 5 (2011): 1787-1797.

[10] Landwehr, Niels, Mark Hall, and Eibe Frank. "Logistic Model Trees." (2004). https://doi.org/10.1007/978-3-540-39857-8_23

[11] Girma, Henok. "A tutorial on support vector machine." *Center of Experimental Mechanics, University of Ljubljana* (2009).

[12] Zupan, Jure. "Introduction to artificial neural network (ANN) methods: what they are and how to use them." *Acta Chimica Slovenica* 41, no. 3 (1994): 327.

[13] Mijwel, Maad M., Adam Esen, and Aysar Shamil. "Overview of Neural Networks." (2019).

[14] Apté, Chidanand, and Sholom Weiss. "Data mining with decision trees and decision rules." *Future generation computer systems* 13, no. 2-3 (1997): 197-210. https://doi.org/10.1016/S0167-739X(97)00021-6

[15] Boruah, Arpita Nath, Saroj Kr Biswas, and Sivaji Bandyopadhyay. "Rule extraction from decision tree: Transparent expert system of rules." *Concurrency and Computation: Practice and Experience* 34, no. 15 (2022): e6935. https://doi.org/10.1002/cpe.6935

[16] Muthu, Shanmuga Pillai Murutha, and Sellapan Palaniappan. "Categorization of Early Detection Classifiers for Gastric Carcinoma through Data Mining Approaches." *Journal of Advanced Research in Computing and Applications* 32, no. 1 (2023): 1-12

[17] Daud, Nuraina, Nurulhuda Noordin, and Nur Islami Mohd Fahmi Teng. "Data Conversion Process Framework to Generate Individual-Level Nutrition Data from Household-Level Grocery Data." In *2022 IEEE International Conference on Computing (ICOCO)*, pp. 425-430. IEEE, 2022. https://doi.org/10.1109/ICOCO56118.2022.10031274