Logistic Regression Modeling to Predict Sarcopenia Frailty among Aging Adults

Sukhminder Kaur[1], Azween Abdullah[2]
Siva Kumar Sivanesan[4]
School of Computer Science and Engineering
Taylor's University, Malaysia

Noran Naqiah Hairi[3]
Centre for Epidemiology and Evidence Based Practice
Department of Social and Preventive Medicine Faculty of
Medicine University of Malaya, Kuala Lumpur

Abstract—Sarcopenia and frailty have been associated with low aging population capacities for exercise and high metabolic instability. To date, the current models merely support one classification with an accuracy of 83%. The models also reflect overfitting dataset complexities in predicting the accuracy and detecting the misclassifications of rare diseases. As multiple classifications led to incongruent data analyses and methods, each evaluation yielded inaccurate results regarding high prediction accuracy. This study intends to contribute to the current medical informatics literature by comparing the most optimal model to identify relevant patterns and parameters for prediction model development. The methods were duly assessed on a real dataset together with the classification model. Meanwhile, the obesity physical frailty (OPF) model was presented as a conceptual study model. A matrix of accuracy, classification, and feature selection was also utilized to compare the computer output and deep learning models against current counterparts. Essentially, the study findings predicted that an individuals' risk of sarcopenia corresponded to physical frailty. Each model was compared with an accuracy matrix to determine the best-fitting model. Resultantly, logistic regression produced the highest results with an accuracy rate of 97.69% compared to the other four study models.

Keywords—Sarcopenia; frailty; logistic regression model; prediction

I. INTRODUCTION

Obesity has been proven to induce frailty in elderly individuals through the most extensively utilized obesity measurement: body mass index (BMI) and age. High risks of coronary heart disease, stroke, and early death have been recently linked to obesity [1, 2] with the perpetually rising rates among aging adults: a drastic 56% increase among individuals between 60 and 69 years old and a 36% increase among individuals over 70 years old in 2020 compared to 2010 [3, 4]. Currently, 37% of adults over 65 years old are obese with a predicted rise in the future [5]. Both skeletal muscle and fat mass would decline between 60 and 70 years of age, thus resulting in a different body distribution [6, 7].

Frailty and obesity are connected to disabilities, high healthcare usage, institutionalization, and early mortality [11-18] with specific symptoms: self-reported tiredness (muscle weakness based on grip strength), slow walking pace (0.8m/s), and inadequate physical exercise [19]. Contrary to popular assumptions, obesity seemingly increased the risk of frailty [4]. According to the Cardiovascular Health Study [20] in Finland, frail patients reflected higher BMI than pre-frail or robust counterparts.

Obesity has posed significant intricacies for public health organizations worldwide. Specifically, the steady and global rise of obesity among aging adults is a substantial phenomenon in both developed and developing countries. Despite the current increase in lifespan, obesity among aging adults is parallelly rising with additional years of disease (cardiovascular illness) susceptibility. In this vein, aging obesity induces significant health complications and a high risk of cancer and death. As a significant contributor to insulin resistance and metabolic syndrome, aging is associated with high cholesterol levels. Knowledge of the primary causes of aging and age-related disorders proved necessary given the seriousness of aging obesity. This study aimed to correlate the fundamental causes of both obesity and aging to indicate that age-oriented changes in fat distribution and metabolism potentially intensified aging and the onset of age-related diseases. The primary elements in a vicious cycle [31] are listed as follows: BMI between 25 and 29,9 kg/m2 is categorized as overweight while BMI over > 30 kg/m2 is obese under the WHO BMI criterion. [32] Nevertheless, the BMI criterion (the most extensively utilized obesity index) overestimates and underestimates obesity in muscular people, including aging adults who have lost body weight. Obesity is also closely associated with other adverse health conditions, such as type II diabetes, heart disease, cancer, and even death [33].

The current "aging population" reflects a significant rise in the number and proportion of elderly people. Elderly individuals are among the most sedentary community members as aging implies the loss of bodily function and low capacities to sustain physical function and physical autonomy despite a longer lifespan. Consequently, most aging adults could be reduced to rudimentary physical skills that induced physical dependence. In this regard, low life quality and negative social and economic (healthcare) implications were gravely concerning [26].

Prediction models could be employed as a method to screen for physical frailty among older adults. In line with international authors, the aforementioned models denote an explicit and clinically relevant tool that facilitates the systematic utilization of routinely collected data and improves information quality and reliability [11]. This study aimed to forecast individuals with the risk of sarcopenia following a

physical frailty prediction model for the oldest primary health patients in the community through participants' clinical variables.

The following sections in this article are organized as follows: Section II presents an overview of obesity and physical frailty with details on relevant works, Section III outlines the model method and materials, Section IV highlights the critical evaluation, Section V presents the experimentation and result evaluation, and Section VI concludes the study.

II. PROBLEM STATEMENT AND LIMITATION

The rapidly aging global population implicitly affected economic growth and health care. In Malaysia, older age groups contribute to 10.3% of the total population based on the 2019 population projection. On the global scale Malaysian has the fourth-fastest aging nation with a 26% rise between 2008 and 2040 [24], the Malaysian population is anticipated to age by 2040 [20] The age structure was also assumed to change drastically following the paradigm shifts of decreased fertility and increased longevity.

The current circumstance poses a novel challenge to the public health care system due to high medical costs and expenditures. Following statistical evidence, the growing elderly population was primarily caused by a high dependency ratio. Although the increase was suggested as a contributing factor, statistics implied the rise to be a primary cause. [25].

Overfitting datasets significantly lowered prediction accuracy due to multiple classifications, thus causing incongruencies between the data findings and techniques employed [25]. Therefore, each evaluation reflected inaccurate results, low prediction accuracy, overfitting tendencies, and poor performance.

Prediction models that demonstrated imbalanced datasets were more commonly skewed against consensus definitions. The connotations were notably essential given the high costs of misclassifying minority examples, such as rare disease identification [22]. Regardless, the lack of available models hampered the management of obesity frailty and dataset complexity to foresee the obesity implications on public health. The limited parameters utilized to predict outcome accuracy inevitably affected specific dimensionalities [27].

A. Elderly Population Demographics

Regarding the rapidly aging global population, a significant increase in average life expectancy in the 20th century reflected one of the most notable social achievements. The rise resulted in a change of major disease and death factors or "epidemiological transition". The transition followed the decline in infectious and acute diseases and subsequent rise of chronic ailments. The data collected from various studies highlighted that recent life expectancy changes were correlated to high disability rates [28]. The Malaysian population would predictably increase from 32.5 million in 2019 to 32.7 million

by 2020. In 2020, the population of individuals between 15 and 64 years old might decrease from 69.8% in 2019 [29]. The increasing number of obese senior citizens was linked to functional disabilities following muscle loss [30]. Based on the Malaysian Adult Nutrition Study, obesity rates have nearly doubled with an increase of overweight and obese adults by over 60% in the last decade. In 2016, 29.1% of individuals were found to be overweight while 14% were obese [31].

B. Aging Related Diseases

An increase in overweight (4.4%) and obesity (14.6%) cases was identified in Malaysia between 1996 and 2018 [32]. Specifically, obesity was found to be higher in women than men [32]. Adults from 40 to 59 years old reflected the highest rate, followed by Malays, Chinese, and the Aborigines [32]. To date, Malaysia is known as the fattest nation within the Southeast Asian region [33] following Fig. 1. In this vein, obesity poses crucial health, growth, and prosperity-related concerns in many countries, particularly Asian nations.

Recently, Malaysia was ranked the second-most overweight nation in East and Southeast Asia [34].

Inactivity also elevates the risk of heart disease and mortality [35]. As health behaviors, obesity, and chronic diseases relied on various biological mechanisms (glucose control and inflammation) [33], investigating the interrelationships between these factors could disclose disease mechanisms and facilitate clinical study designs.

In Table I, the comparative studies on initial accuracy models with various parameter types were utilized for obesity and frailty. DeGregory et al. [36] stated that obesity-related complexities and implications could be identified, recognized, and forecasted with machine-learning algorithms. Notably, the key component analysis (PCA) accuracy level was 83% (extremely high). Meanwhile, Bassam et al. [37] defined logistic regression with KNN and SVM for modeling. Resultantly, SVM demonstrated the highest accuracy level (73%) which involved specific attributes (age, BMI, gender, waist circumference, physical activity, diet, pre-existing hypertension, family history of hypertension, and diabetes) as modeling accuracy contributors.
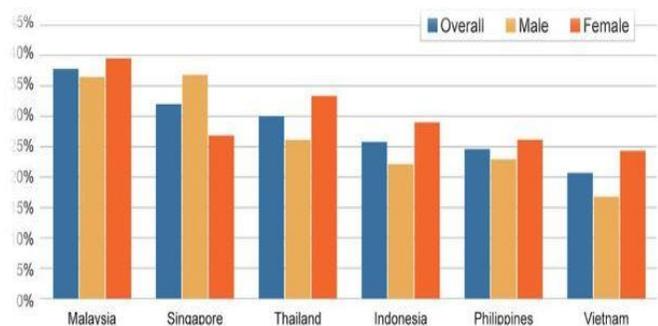
Fig. 1. OBESE Population with BMI >25.

TABLE I. COMPARATIVE STUDIES ON EARLY MODEL ACCURACY

Author	Model	Predictors used	AUC
DeGregory et al	Linear and logistic regression, artificial neural networks, deep learning, decision tree analysis, cluster analysis, principal component analysis (PCA), network science, and topological data analysis	Age, BMI, family history of diabetes, hypertensive status, family history of hypertension, sex, inactivity, and improper diet	83%
Bassam et al.	Logistic regression, k-nearest neighbor, KNN, and support vector machine (SVM)	Age, BMI, sex, waist circumference, physical activity, diet, pre-existing hypertension, family history of hypertension, and diabetes (type II)	SVM: 73 %
Jindal et al.	Random forest	Age, BMI, weight, and height	89.68
Seyla et al.	SVM	Age, BMI, BP, HDL, and LDL cholesterol A1C, and smoking	66.2%
Zheng et al.	Improved decision tree (IDT), KNN, and artificial neural network (ANN)	Inactivity, improper diet	KNN – 88.82% IDT – 80.23% ANN – 84.22%
Brandilyn et al.	Random forest model, and performed cross	Age, BMI, and physical activity	70%

Jindal et al. [38] deployed random forest using the attributes of age, BMI, weight, and height with 89.68% accuracy. Following Seyla et al. [38], the sample employed SVM with the attributes of age, BMI, BP, cholesterol, and smoking with 66.2% accuracy. Zheng et al. [39] elaborated on the decision tree, KNN, and ANN models with 88.82%, 80.23%, and 84.22% accuracy, respectively. Regarding inactivity and poor diet, Brandilyn et al. [40] assessed the random forest model using cross-sectional research based on age, BMI, and physical activity as the modeling parameters with 70% accuracy. The accuracy corresponding to the linear model categorization was merely 90.5%.

Limited feature and dimension selection were identified in driving the result plots under the curve. Carlos Rodriquez's model could not provide a clear depiction of attribute selection following the small number of respondents while other model limitations were primarily related to low model dependability. The model accuracy was also found to be low. Consequently, the approach proved unreliable following sample size insufficiency to support data-derived conclusions.

This study encountered several limitations. First, the model comparisons lacked obese individuals who reflected a higher frequency of pre-frailty and frailty in a 22-year follow-up study in Finland. As such, obesity could be a primary factor regarding frailty progression [21, 22]. As small sample sizes instigated overfitting following past studies, the samples complemented large datasets compared to data from a single-site experimental clinical trial. Regarding computing complexities, addressing the decision tree application with appropriate thresholds denoted the most intricate class of algorithms. Thus, deep learning algorithms were typically most effective on relatively large training datasets. Notwithstanding, dataset pre-processing and standardization could be time-consuming [8].

Another element potentially limiting the study validity was data collection. According to Jacy Aurelia, the cross-sectional study design deterred the causal relationships between clinical variables and study outcomes. The findings could not be generalized as the sample represented a specific community. Alternatively, longitudinal and multi-center studies should be performed to better understand the relationships and verify the transition between frailty levels based on case severity and reversibility in the medium and long terms [9].

The integration of both measures could elevate precision in identifying aging adults who were more vulnerable to adverse health events. For example, a conceptual development that involved scholarly collaboration from various disciplines could alleviate some of the aforementioned shortcomings. Notably, a hypothetical future model of frailty among elderly individuals should not be as restrictive as the physical phenotype or as broad as the index derived from multiple domains [10].

III. METHOD AND MATERIAL

This study aimed to determine the use of specific parameters and predict obesity and frailty with logistic regression. Primarily, the suggested methodology strived to develop the prediction categorization criteria for the most optimal model (Naïve Bayes, logistic regression, random forest, decision tree, or KNN). This section explained the proposed research methodology to compare the performance of all five study models and employ the most accurate predictive model with specific parameters.

A. Data Pre-processing

Data pre-processing is a critical stage in data management before being utilized in data-mining algorithms. The procedure encompasses several steps: cleaning, normalization, and feature selection transformation. Data transformation significantly influenced this study analysis as most of the features were categorical with a combination of string and special characters. The values needed to be converted into numeric-categorical values for improved prediction performance by the predictive models. The cleaned data would then be split into two sets (train and test sets) after data pre-processing. The ratio between both sets (80:20) proved ideal for data-splitting. Specifically, predictive model accuracy was higher if the model trained on 80% of the data and assessed the model performance on the remaining 20%. In model development, five supervised learning classification models were implemented for prediction. Supervised learning models were chosen in line with the target attributes for known outcome prediction. The data was fed into the predictive models while prediction accuracy was employed to assess the model performance post-test.

B. Feature Selection

As feature engineering involves the process of selecting a subset of relevant attributes for model inclusion, various cutoff points were established for the most influential traits to be incorporated into the tests. The accuracy of each algorithm was compared with a dataset using the selected features. The process was then repeated with multiple thresholds for optimal results.

C. Classification using Data Mining

Several algorithms could be run on a dataset to determine the most accurate counterpart in predicting obesity and frailty impacts. Clustering and association rule-mining denoted some of the tasks that could be accomplished through data-mining apart from classification and prediction. The data were assigned with a pre-defined class label using a two-step process for dataset classification. Specifically, training data were analyzed for classification model development (classification rules) that described a set of pre-defined classes. Meanwhile, the model was classified post-test data as part of the second phase to determine model accuracy.

Pseudocode: Identify Response and Explanatory Variables

1. Original Dataset ← contains all variables
2. Target ← Obese and Frail, Non-obese and Frail, Obese and Non-Frail, Pre-frailty
3. Response Variable ← Target
4. Correlation Matrix ← Generate a correlation plot to find highly correlated attributes
5. If Correlation => 0.50 THEN
   a. Explanatory Variable ← Feature == Chosen
6. Else
   b. Feature == Dropped
7. Data ← Explanatory Variable, Response Variable
8. Data.to_csv()
9. Logistic Regression : Input Test X, Target Y
10. Import Logistic Regression from Linear Model
11. LogisticRegr <– LogisticRegression ( )
12. If LogisticRegr.fit( ) == NULL Then
13. Return "Error"
14. Else LogisticRegr.fit (x_train, y_train)
15. Train the model on the data
16. Accuracy ← LogisticRegr.score (Test Data)
17. Return Accuracy
18. Confusion Matrix ← LogisticRegr.predict (Test Data)
19. Return Confusion Matrix
20. Combined Model : Input ← Features == LogisticRegr
21. If LogisticRegr.fit ( ) == TRUE Then
22. Create New Model OPF (Obesity Physical Frailty)
23. Else
24. Return.

The aforementioned logistic regression pseudocode indicated the steps to define the coded independent variables X and Y. A logistic regression package was subsequently imported from the "sklearn linear model" library. A logistic regression model with all the unspecified parameters (set to default) was generated as the third stage. Otherwise, the function progressed to the data training phase following step 7.

A model would learn the relationship between "x train" and "y train" through split-data training. The model accuracy was then measured for performance assessment. The logistic regression function was scored for improved model accuracy. Lastly, a confusion matrix table described the classification model performance on a test dataset. Specifically, a function was predicted to forecast the value labels on the testing data. The particular dataset element must be accurately selected for a good forecast to significantly impact the prediction process and result. In this vein, the variable composition within the dataset must be carefully examined for accurate prediction.

IV. CRITICAL EVALUATION

The approach towards analytical design development was utilized to evaluate aging adults with notably high BMIs. Based on a finding that paralleled multiple modeling, the currently-utilized model only supported a single classification with accuracy as low as 83%. The models reflected overfitting dataset issues in predicting the accuracy and misclassification of rare disease detection. Additionally, the data findings and methods became unbalanced with multiple classifications. A significant correlation was identified between obesity and frailty regardless of the classification. This study primarily aimed to examine how obesity affected physical frailty, diseases, and aging men and women's health through the sarcopenia-physical frailty link. This research did not define the differences between Obesity Frailty (OF), Non-Obesity Frailty (NOF), Obesity Non-Frailty (ONF), and Pre-Frailty (PF) but the research explored the expansion of differences between OF, NOF, ONF, and PF. The study parameter types with specifications only included age, BMI, physical activity, protein and meat intake, body composition, fat mass, and disease types.

V. ANALYTICAL APPROACH

The study data were pre-processed and prepared for the following steps using Excel and RapidMiner. Role-oriented attribution ("label" or target variable) informed the predictive models of attribute prediction. The "set role" operator sets the role attribute to identify the key determinants of obesity with the highest accuracy. The dataset was split into two components (train and test) post-role-definition. Dataset splitting proved necessary as the predictive model must learn from the training set to be applied to the test for performance evaluation. The train and test datasets were split into 70:30. Specifically, 70% represented the train datasets while 30% reflected the test counterpart. The training set typically encompassed more data than the test counterpart as the model could learn from the data for improved accuracy. Fewer training datasets implied lower opportunities for data learning and exploration.

Phyton programming was utilized to split and feed data into the selected predictive model once the role (decision tree) was set. The decision tree model was selected as the predictive model to demonstrate the chain relationship between attributes and final result for a clear depiction of obesity contributors and the individuals at risk of becoming obese in the future. The "performance" operator was utilized to assess model performance with accuracy as a criterion. Although the accuracy model demonstrated precision, the manipulated

variable reflected the attributes that increased obesity risks. The predictive model accuracy on the obesity dataset was 92.95% for "breathing difficulties", 96.81% for "heart attack", 86.45% for "hyperlipidemia", and 99.47% for "psychological stress". All four attributes were selected as the most accurate obesity predictors. Consequently, each data was incorporated into modeling and ultimately defined each characteristic table regarding the accuracy outcome among classifiers.

VI. EXPERIMENTATION AND RESULT EVALUATION

This section discusses the experiments and results for all five study classifiers (Naïve Bayes, logistic regression, random forest, decision tree, and KNN) Different comparisons and analyses are also discussed in this section. The high-performance approaches predicting obesity frailty are also highlighted. Notably, accuracy, precision, recall, and FI measures were utilized in the comparison.

A. Dataset Splitting

The pre-processing steps for the study dataset were described in the preceding section. The dataset must be divided into training and testing upon the completion of pre-processing.

B. Experiments and Results

Table II presents the accuracy attained by all five classification predictive models: Naive Bayes, logistic regression, random forest, decision tree, and KNN. The best-performing model among the five predictive counterparts was logistic regression. Specifically, logistic regression denoted the only model with the highest number of accuracies predicted (see Table II). Logistic regression denoted the most appropriate model to predict obesity and frailty among elderly individuals.

The analytical design approach was utilized to evaluate raw data with parameters, such as age, BMIs, physical activity, protein and meat intake, peanuts, and the composition of body, body fat, and fat mass. The OPF model encompassing naive bayes, logistic regression, random forest, decision tree, and KNN were applied to four different target attribute types: "OF", "ONF", "NOF", and "PF". The most accurate models for OF were logistic regression and KNN with 70.83% accuracy. The model would attain high accuracy by predicting high genuine and low false data with 89.81% accuracy.

Notably, the decision tree outperformed all other five models by predicting the highest true values for the NOF counterpart. Meanwhile, ONF was predicted by logistic regression and KNN (see Table III).

TABLE II.   COMPARISON OF FIVE CLASSIFICATION PREDICTIVE MODELS

Classifiers (%)	Obese and Frailty	Non-Obese and Frailty	Obese and Non-Frail	Pre-Frailty
Navie Bayes	**51.36**	**41.2**	**62.5**	**97.69**
Logistic Regression	**70.83**	**78.24**	**89.81**	**97.69**
Random Forest	**68.06**	**79.63**	**87.04**	**97.69**
Decision Tree	**68.98**	**82.87**	**86.57**	**93.37**
K-Nearest Neighbour	**70.83**	**76.39**	**88.89**	**90.74**

TABLE III.   THE OPF MODEL

Current Models	Methods	Parameters used	AUC
OF	Logistic Regression KNN	**Age, BMI, physical activity, protein intake, meat intake, peanut, body composition, body fat, fat mass, diabetes, hypertension, Hyperlipidemia**	70.83
NOF	Decision Tree		89.81
ONF	Logistic Regression		89.81
PF	Naïve Bayes Logistic Regression Random Forest		97.69

The modeling implied that logistic regression and Naive Bayes offered the lowest prediction for false positives and negatives. Logistic regression earned the highest accuracy of 89.81% by predicting the truest data and minimal erroneous data. The confusion matrix for PF prediction demonstrated Naive Bayes and logistic regression to have the largest true positives, whereas the decision tree highlighted the highest true negatives. The least false positives were predicted by the decision tree while the least false negatives were forecasted by Naive Bayes and logistic regression. Naive Bayes, logistic regression, and random forest were the most optimal models that predicted PF patients with 97.69% accuracy.

Fig. 2 depicts the receiver operating characteristic (ROC) curves produced for obese and frail subjects. It was deemed possible to judge the ROC curve performance based on the closeness to both the left and top borders of the curve. Regarding obesity and frailty prediction, five algorithms were tested against one another. Logistic regression was found to be the most accurate model compared to other predictive counterparts.

Fig. 3 depicts the logistic regression model. Compared to random forest, the ROC curve was more closely associated with the left and top plot bounds. The test was performed when the curve was closer to the 45-degree diagonal of ROC space compared to when the curve was further away.

Regarding logistic regression, the ROC curve in Fig. 4 presents the trade-off between sensitivity (TPR) and specificity (1-FPR). Classifiers with curves that neared the top-left corner indicated higher performance compared to the decision tree. A curve closer to the 45-degree diagonal of ROC space indicated a less accurate test.
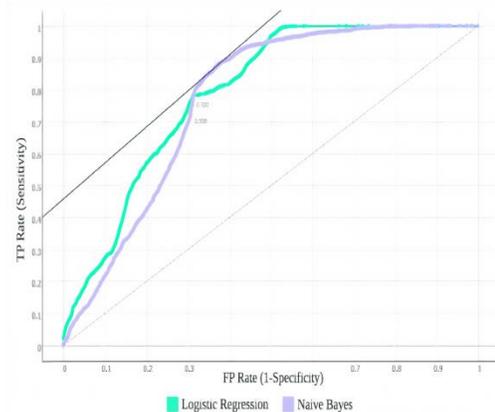
Fig. 2.   The ONF ROC Curve (Logistic Regression and Naive Bayes).
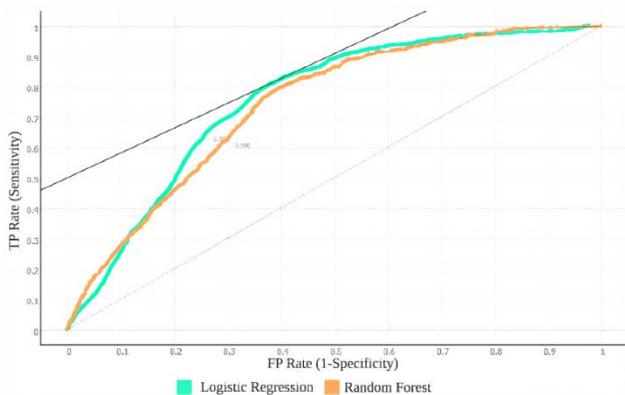
Fig. 3.   The of ROC CURVE (Logistic Regression and Random Forest).
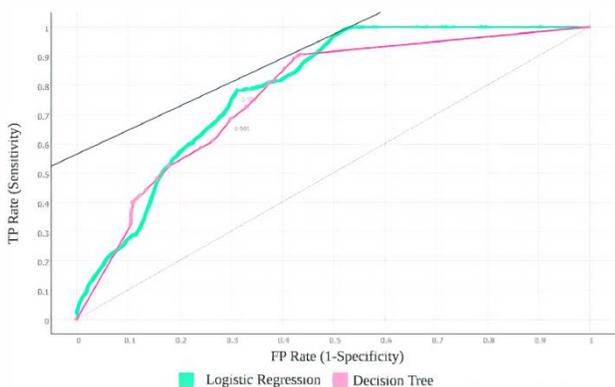
Fig. 4.   The of ROC Curve (Logistic Regression and Decision Tree).

In Fig. 5, all four figures depict the logistic regression ROC curve to be closer to the left and top plot boundaries. The ROC curve for KNN was the poorest as the KNN-logistic regression gap was the widest. The ROC analysis above demonstrated logistic regression to have the best true positive rate among the remaining algorithms while the false positive rate was low. The logistic regression accuracy was higher as high true positive and low false positive rates highly contributed to the degree of model accuracy.

Overall, the obese elderly (OE) group demonstrated a higher average lean mass (LE) compared to the remaining two groups. An increase in fat mass rather than lean mass occurred during middle age. Height loss (primarily in the spine), loss of muscle mass, increase in fat mass, central, abdominal, and muscle fat, and bone loss due to aging required due consideration. Both fat and muscle loss occurred in later stages as a sign of vulnerability. The life expectancy process occurred in catabolic life with sarcopenia. The condition also reflected composite variables that were interconnected with one another. The body composition altered over 30 years with changes in weight and fat-free muscle.

Regarding obesity in Fig. 6, the clinical definition of frailty implied individuals with low and extremely high BMIs to have higher frailty levels compared to people with normal or high BMIs. Regarding higher frailty levels among people with large waist circumferences, adiposity in the abdomen appeared to be associated with high risks. The BMI-frailty association held constant across various term definitions with specific parallels

between both elements that could be considered while examining the relationship. Low stress resistance from a single variable change could be interpreted as aging or high mortality risk following changes in hundreds of components [21]. For example, the mortality risks associated with variables were typically U or J-shaped [21]. A U-shaped association existed between age and the BMI risks of unpleasant outcomes and factors in most physiological systems [22]. At the cellular level, numerous systems that deactivated adaptive responses to protect age and BMI exhibited U or inverted U-shaped responses [23]. Although the data did not demonstrate scale invariance, the frailty "risk state" was associated with age and BMI.

Based on the findings in Fig. 7, comparisons between OF and hypertension OF reflected a higher level of hyperlipidemia than hypertension. Lastly, hyperlipidemia was the most chronic issue among ONF, whereas hypertension was the most prevalent condition among PF (189 respondents). The added impact of lower-extremity obesity on frailty was discovered among older adults, thus highlighting the essentiality of regular physical activities as an alternative. Predictably, abdominal obesity among older people with low BMIs would become a novel therapy target in the future parallel to the study findings. Increased physical activities could reduce belly fat [42] while endurance through exercise training could develop mitochondria [41]. Notably, physical activity benefits could be due to reduced belly adiposity and enhanced oxidative activity (independent of weight loss benefits) [41].
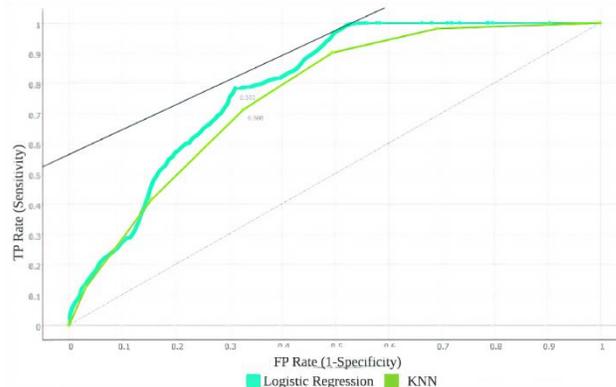
Fig. 5.   The of ROC Curve (Logistic Regression and KNN).
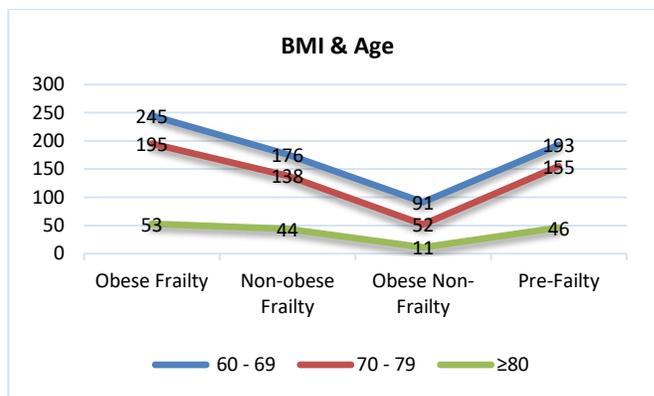
Fig. 6.   The of with Aging Individuals Aged 60 and Older with Bmi and Degrees of of.
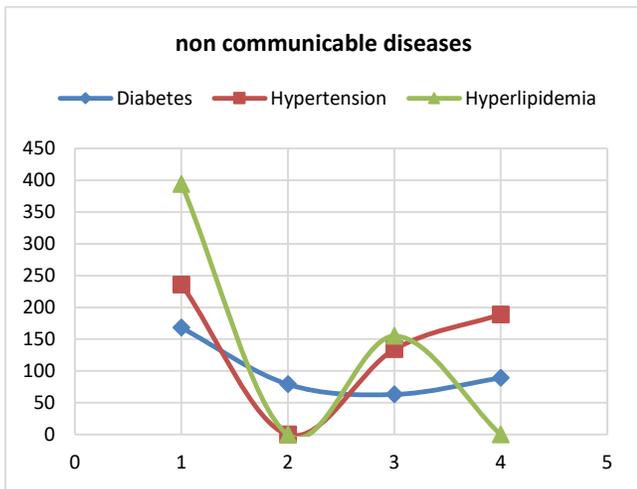
Fig. 7.   The of Proportion through Non-Communicable Diseases with of Levels.

It was deemed essential to understand the frailty-obesity relationship and the criteria validity utilized to categorize the factors. Specific interventions, such as weight loss and exercise among older adults could alleviate frailty with a direct impact on the downstream implications of central obesity [29, 30]. As distinct entities, PF, frailty, and obesity were related to impairment, high healthcare usage, institutionalization, and premature total mortality [29]. Resultantly, physical activity and contributors of age, BMI, physical activity, protein and meat intake, peanuts, body composition, body fat, fat mass, diabetes, hypertension, and hyperlipidemia demonstrated the strongest connection to frailty with a potentially superior predictor of frailty risk in the OE population. The simple anthropometric measure might be easily and affordably incorporated into clinical procedures and provide prognostication among older people.

Regarding the overall test scores in Table IV, the logistic regression model yielded higher scores than neural network, random forest, and SVM. Concerning the highest accuracy, logistic regression generated more area of curve (AUC), CA, F1, precision, and recall scores. Conclusively, logistic regression predicted fewer false positives and negatives compared to other algorithms.

Fig. 8 depicts a conceptual overview of the obesity-frailty relationship, diseases, and lifestyles. Failure of heart muscles and blood vessels, diabetes, hypertension, and heart diseases were all associated with obesity and frailty. Subsequently, obesity and frailty were connected to related diseases and lifestyle factors. Active people consumed highly-sugared beverages, peanuts, and red meat and watched television for longer periods, thus contributing to obesity. Weight gain that affected various organs implied obesity. Additionally, aging bodies eventually degraded and posed high-risk and adverse health outcomes. Both combination levels led to various diseases and lifestyle-related health complexities and caused exhaustion, perspiration, anxiety, irritability, and joint and back pains that prevented sleep.

TABLE IV.   OVERVIEW TEST SCORES OF THE PREDICTIVE MODELLING

Model	AUC	CA	F1	Precision	Recall
Logistic Regression	0.982422	0.950463	0.948411	0.948032	0.950463
Neural Network	0.963104	0.941204	0.938094	0.937523	0.941204
Random Forest	0.950443	0.933565	0.922497	0.933513	0.933565
SVM	0.940531	0.936111	0.925756	0.936804	0.936111
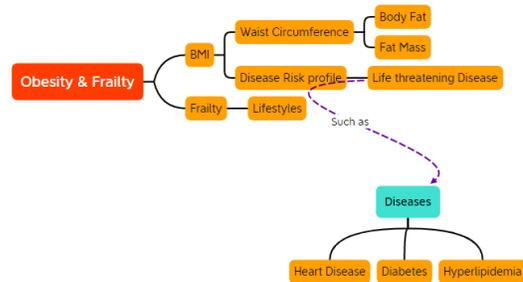
Fig. 8.   Conceptual Model of Relationship with Diseases and Lifestyles (OPF).

Regarding the study limitations, it was deemed impossible to establish a causal relationship between the clinical variables and study outcome due to the restrictive cross-sectional study design. The study results were not generalizable as the sample represented a specific community. In this vein, logistic regression studies should be performed to further investigate the relationships and confirm the transition between frailty levels parallel to the case severity and reversibility in medium and long terms.

VII. CONCLUSION AND FUTURE WORK

This study aimed to create and apply a model that differed from current models by including obesity-related diseases and physical frailty. The present models only facilitated single classification with accuracy as low as 83%. Resultantly, the proposed model reflected 97.69% of accuracy and was associated with physical frailty and obesity. The OPF conceptual model was also introduced in this study. With a corresponding improvement in the logistic regression outcomes, optimal results were obtained in bridging the research gap. Current and novel approaches to machine learning fulfilled the demand for advanced and high-level prediction and description by utilizing both established and emerging machine-learning methods. Future studies intended to include logistic regression and neural network that involved participants to achieve higher accuracy. The same data were also recommended to predict and create novel modeling against OPF.

ACKNOWLEDGMENT

REFERENCES

[1] Gregg EW, Cheng YJ, Cadwell BL, Imperatore G, Williams DE, Flegal KM, et al. Secular trends in cardiovascular disease risk factors according to body mass index in US adults. Jama. 2016;293(15):1868–74. [PubMed: 15840861].

[2] Seidell JC, Halberstadt J. The global burden of obesity and the challenges of prevention. Ann Nutr Metab. 2017;66 Suppl 2:7–12.

[3] Porter Starr KN, McDonald SR, Bales CW. Obesity and physical frailty in older adults: a scoping review of lifestyle intervention trials. J Am Med Dir Assoc. 2017;15(4):240–50. [PubMed: 24445063].

[4] Hubbard RE, Lang IA, Llewellyn DJ, Rockwood K. Frailty, body mass index, and abdominal obesity in older people. J Gerontol A Biol Sci Med Sci. 2017;65(4):377–81. [PubMed: 19942592].

[5] Ying Yi Liao, Hsuan Chen, Ray Effects of Kinect based exergaming on frailty status and physical performance in prefrail and frail elderly: A randomized controlled trail. 2019;1(2):47–9. [PubMed: 25364705].

[6] Allyson K Palmer, James L Kirkland, Aging and adipose tissue: potential interventions for diabetes and regenerative medice, 2016.

[7] Laurent Maimoun, Thibault Mura, Body Composiion in Individual with obesity according to age and sex: A cross sectional study, 2020.

[8] K. W. DeGregory1, P. Kuiper1, "A review of machine learning in obesity", 2018.

[9] Jacy Aurelia Vieira de Sousa, Maria Helena Lenardt " Physical frailty prediction model for the oldest old", 2018.

[10] A.A Pereira, F.S.A Borim, "Comparison of two models of frailty for the prediction of mortality in Brazilian Community Dewelling older adults: the fibra study. 2019.

[11] Comans TA, Peel NM, Hubbard RE, Mulligan AD, Gray LC, Scuffham PA. The increase in healthcare costs associated with frailty in older people discharged to a post-acute transition care program. Age Ageing. 2016;45(2):317–20. [PubMed: 26769469].

[12] Avila-Funes JA, Amieva H, Barberger-Gateau P, Le Goff M, Raoux N, Ritchie K, et al. Cognitive impairment improves the predictive validity of the phenotype of frailty for adverse health outcomes: the three-city study. J Am Geriatr Soc. 2019;57(3):453–61. [PubMed: 19245415].

[13] Dee A, Kearns K, O'Neill C, Sharp L, Staines A, O'Dwyer V, et al. The direct and indirect costs of both overweight and obesity: a systematic review. BMC Res Notes. 2019;7:242. [PubMed: 24739239].

[14] Visscher TL, Seidell JC. The public health impact of obesity. Annu Rev Public Health. 2017;22:355–75. [PubMed: 11274526].

[15] Hirani V, Naganathan V, Blyth F, Le Couteur DG, Seibel MJ, Waite LM, et al. Longitudinal associations between body composition, sarcopenic obesity and outcomes of frailty, disability, institutionalisation and mortality in community-dwelling older men: The Concord Health and Ageing in Men Project. Age Ageing. 2016.

[16] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2017;56(3):M146–56. [PubMed: 11253156].

[17] Walston J, McBurnie MA, Newman A, Tracy RP, Kop WJ, Hirsch CH, et al. Frailty and activation of the inflammation and coagulation systems with and without clinical comorbidities: results from the Cardiovascular Health Study. Arch Intern Med. 2020;162(20):2333–41. [PubMed: 12418947].

[18] Stenholm S, Strandberg TE, Pitkala K, Sainio P, Heliovaara M, Koskinen S. Midlife obesity and risk of frailty in old age during a 22-year follow-up in men and women: the Mini-Finland Follow-up Survey. J Gerontol A Biol Sci Med Sci. 2019;69(1):73–8. [PubMed: 23640762].

[19] Stevens J, Cai J, Pamuk ER, Williamson DF, Thun MJ, Wood JL. The Effect of Age on the Association between Body-mass Index and Mortality. N Engl J Med. 2016; 338(1):1–7. [PubMed: 9414324].

[20] Waters DL, Ward AL, Villareal DT. Weight loss in obese adults 65years and older: A review of the controversy. Exp Gerontol. 2013.

[21] Yashin AI, Arbeev KG, Kulminski A, Akushevich I, Akushevich L, Ukraintseva SV. Health decline, aging and mortality: how are they related? *Biogerontology*. 2017;8(3):291–302.

[22] Whelan B, Bennett K, O'Riordan D, Silke B. Serum sodium as a risk factor for in-hospital mortality in acute unselected general medical patients. *QJM*. 2009;102(3):175–182.

[23] M. Jura and L. P. Kozak, "Obesity and related consequences to ageing," *CrossMark,* 2016.

[24] Nur Aina Daud et al, "Predictive Analytics : The Application of J48 Algorithm on Grocery Data to Predict Obesity," in *ResearchGate*, 2018.

[25] Singh et al, "Machine learning approach for the early prediction of the risk of overwieght". 2018.

[26] J.A. Vaccaro, T.Gaillard, Huffman, "Motivational Strategies to prevent frailty in older adults with diabetes", Hindawi, p.8, 2019.

[27] Y. Y. D. e. al, "Multidimensional predictors of physical frailty in older people: identifying how and for whome they exert their effects," 2017.

[28] V. Salvestrini, "Obesity May Accelerate the Aging Process," *Frontiers in Endocrinology,* p. 266, 2019.

[29] Department of Statistics Malaysia, 2020, "The source of Malaysia's Official Statistics".

[30] M.-V. Em., "Obesity the the elderly," *PMID: 22772735,* pp. 533-44, 2019.

[31] "Sensitivity and specificity of waist circumference overweight and obesity malaysian adults," p. vol 66, 2011.

[32] T. T. N. H. a. M. Survey, "NHMS III," *Institure Public Helth National Insti,* 2011.

[33] T. o. i. ASEAN, "Tackling Obesity in Asean," *Economist Intelligence1Unit Report .* 2017.

[34] K. K. L. Ying Ying Chan1*, "Physical activity and overweight/obesity among Malaysian adults," 2015.

[35] H. James O and J. C.Peters, "Environmental Contributions to the Obesity Epidemic," *web of Science,* p. Vol 28. 2018.

[36] Yashin AI, Arbeev KG, Kulminski A, Akushevich I, Akushevich L, Ukraintseva SV. Health decline, aging and mortality: how are they related? *Biogerontology*. 2017;8(3):291–302.

[37] Whelan B, Bennett K, O'Riordan D, Silke B. Serum sodium as a risk factor for in-hospital mortality in acute unselected general medical patients. *QJM*. 2019;102(3):175–182.

[38] Zhang Q, Pi J, Woods CG, Jarabek AM, Clewell HJ, Andersen ME. Hormesis and adaptive cellular control systems. *Dose Response*. 2018;6(2):196–208.

[39] Lakka TA, Laaksonen DE. Physical activity in prevention and treatment of the metabolic syndrome. A*ppl Physiol N*utr *Metab*. 2017;32(1):76–88.

[40] Short KR, Vittone JL, Bigelow ML, et al. Impact of aerobic exercise training on age-related changes in insulin sensitivity and muscle oxidative capacity. *Diabetes*. 2016;52(8):1888–1896.

[41] Han, J., Kamber, M, "Data preprocessing. In: Kamber, M. (ed.) Data Mining Concepts and Techniques," Elsevier, Amsterdam, 2019.

[42] J. A. V. d. Sousa, "Physical frailty prediction model for the oldest old," 2018.