Supported by tetfund

OBESITY PREDICTION MODEL USING MACHINE LEARNING TECHNIQUES.

Kabir Kasum Yusuf, Musa Fati Anisat, Taiwo Abiodun
Federal University Lokoja
Corresponding author: kabir.yusuf-msc@fulokoja.edu.ng

Received: September 14, 2023 **Accepted:** November 28, 2023

Abstract: Currently, safeguarding the community is vital in terms of finding solution to health related problems which can be achieved through medical research using the advent of technology. Obesity has become worldwide health concern as it is becoming a threat to the future. It is the most common health problems all over the world and it is associated with thousands of diseases and risks as well death. An early prediction of a disease can help both doctors and patients to act and minimize if not total eradication of the root cause or work on preventing the disease symptom from further deterioration. Going through patient's medical history is one of the methods of identifying a disease which most time consuming as processing manually and it comes with an error-prone analyses and expense. Therefore, there is need to predict an occurrence of the disease or its existence using a semi-automated or automated technique as its becoming a need of the day. In this research work, we used machine learning techniques on a public clinical available dataset to predict obesity status using different machine learning algorithms. Five machine learning algorithms were applied. Gboost Classifier, Random Forest Classifier, Decision Tree Classifier, K-Nearest Neighbor and Support Vector Machine and the model has shown promising results with as Gboost classifier achieves the highest accuracy of 99.05% as compared to other classifiers. Meanwhile, the K-Nearest Neighbor gave the poorest accuracy of 95.74%.

Key words: Obesity, prediction, machine learning, algorithms, model.

Introduction

Obesity is one of the most common health problems all over the world and it is associated with thousands of diseases and risks as well death. It is a worldwide health concern as it is becoming a threat to the future. According to (Eduado, Fabio, & Mendoza, 2019), obesity is a popular and dynamic global disease that has rapidly increased since the 80's, with serious health concerns in adults, youths and children respectively. (Eduado, Fabio, & Mendoza, 2019), also observed the problems in obesity has been rising fast, for this reason new studies involves both young and old obesity level detention. Based on this perception, it is stated that the (WHO, 2021), describes the word obesity as an abnormal or excess fat accumulation that may damage health. Lots of people above 16 years of age are suffering from alteration of their weight. This overweight can be caused by lots of intake of starchy foods that are high in fat. More also it is caused by lack of physical exercise. According to (Guterrez, 2010), obesity is a general health problem around the world and it can occur in teens, children and adults. (Hernandez, 2011) Shows that we can consider obesity as a disease that has different factors, having symptoms like uncontrolled weight increase, due to too much intake of energy giving food and fatty food. The research shows that, obesity can be caused by biological factors like hereditary background, so we have different kinds of obesity as: Polygenic monogenic, syndromic and leptin. Although, there are other risk factors such as social, eating habits and psychological as stated by (Del Cisne & Zhingre, 2015). Obesity is a major health problem all over the world. Obesity can lead to chronic diseases like cardiovascular diseases, diabetes and cancer. Detecting obesity early enough has been actively pursued in health programs. However, years of continuous research have shown that to manage and prevent obesity is not an easy task. This is as a result of our limited knowledge about obesity and the complex interactions among its various factors, including environmental and biological factors that can contribute to obesity.

This work aimed at developing a machine learning model that would be an improved version of the existing model on obesity prediction using a local dataset with obesity principle components. And the particular objectives are:

i. To determine the best set of features for obesity prediction
ii. To develop a model that will predict obesity of individual using machine learning techniques with locally available dataset.
iii. To use supervised machine-learning algorithm in order to enhance the obesity level estimation for future prediction.
iv. Test and evaluate the performance of the model using the evaluation metrics

Literature review

(Xiaolu, *et al*., 2021)Tries to examine the relationship between the weight status and the physical activities in human and also compare some machine learning and classical statistical models used in predicting obesity level. National Health and Nutrition Examination Survey Dataset was used in their model, and they made use eleven different algorithms which include the random sub space, logistic regression, decision table, Naïve Bayes, the Radial Basis Function, K-nearest neighbor, classification via regression, J48 and Multilayer perception are set of algorithms used for their implementation and evaluation. The evaluation metric used was the ROC and AUC and the algorithm that achieved the overall highest accuracy was the random subspace classifier algorithm.

(DeGregory, Patrick , DeSilvio, & Pleuss, 2018) Use machine learning methods such as linear and logistic repressor, Artificial Neural Network, Deep Learning and Decision Tree analysis to predict and/or classify obesity level from large dataset gotten from sensors, smartphones, and electronic medical health records. They concluded that machine learning will provide a sophisticated tools to predict, classify and describe obesity related risks and its outcome.

(Zachary, *et al*., 2019) Tries to predict the percentage of obesity level in adult in the United States in the nearest future by using fitting multinomial regression to estimate the prevalence of 4 BMI categories (less than 25 is normal weight, 25 to less than 30 is overweight, 30 to less than 35 is moderate obesity and above 35 is severe obesity) on a self-reporting bias dataset from the behavioral risk factor surveillance system survey of 1993-1994 and 1999- 2016. They concluded from their analysis that obesity level in adult will always be on the increase in 2030 with 48.9% obese,24.2% severely obese and the occurrence will be higher than 25% in 25 states.

(Hagai, Smadar, Shiri, Becca, Aron, & Eran, 2021) Evaluated the BMI increase pattern in kids and also develop a system that predicts those children who have high risk of being obese before it gets to critical stage. From their evaluation, they reviled that the greatest increase in BMI happens between the age of 2 & 4 and their accurate prediction happens at 5 to 6 years of age.

(Xuegin, Christopher, Felice, & Aaron, 2021) design a model that predict early childhood obesity using XGBoost, ID3 decision tree model and the Recurrent Neural Network(RNN) machine learning algorithm to model an electronic health record data. From their area under curve the XGBoost has the best AUC-value of 0.81 (0.001) and it outperformed all other models.

(Davila, DeGuzman, Johnson, & Serban, 2015), created a logistic regression model to evaluate the probability of BMI on children between the age of 2 and 17 in rural areas using publicly available dataset. The outcome shows that in small geographic areas, estimates are important to create effective involvements and to help planning of potential solutions to the problem as prevalence among census varies from 27 to 40%.

A model was presented by (Manna & Jewkes, 2014) , which uses fuzzy signature to understand and manage intricacies on children obesity dataset and a solution that could handle the related risk with children motor development and early obesity.

(Adnan M. H., 2011) Had an initial method to predicting obesity which was developed with the information collected from primary sources: Parents, caretakers and the children themselves. The paper authors tries to identify the risk factors like the obesity, education level of the parents, habits, lifestyle of the children and the environment influence on children. The proposed framework uses a hybrid approach of the decision trees and the Naïve Bayes known as the NBTree.

(Adnan, Damanhoori, & Husain, A survey on the usefulness of data mining for childhood obesity prediction., 2010) Predict children obesity using data mining. The aim of the suggested survey was to provide the required information on obesity problem. The NN, the Naïve Bayes and the Decision Trees models were used for their implementation.

(Abdullah, Manan, Ahmad, Wafa, & Shahril, 2016) These authors paper showed the classification of obesity in grade 6 children, from two separate districts in Malaysia. A classification technique was used to model the information collected. The machine learning classification models used are Decision Trees, Support Vector Machines, Neural Networks and Bayesian Networks.

An article published by (Tamara, Mukhopadhyay, Aaron , & Stephen, 2015) aims to predict the obesity in children over the age of 2 by using data collected before the second birthday of the children, they analyzed six different machine learning methods, the machine learning methods they used are the ID3, random tree, random forest, J48, Naïve Bayes and Bayes train on CHIKA data. The overall accuracy of 85% was gotten from the random forest

(Kapil, Niyati, & Prashant, 2018), used ensemble machine learning approach to predict obesity in human using attribute like the age and BMI (weight and height) of individuals. They use machine learning algorithms such as the linear model, the random forest and partial least square. They obtained an accuracy of 89.68% from random forest.

Zachary, *et al*., 2019 carried out a research on the prediction of the percentage of obesity level in adult in the nearest future using Fitting Multinomial Regression to estimate the prevalence of four (4) BMI categories (less than 25 is normal weight, 25 to less than 30 is overweight, 30 to less than 35 is moderate obesity and above 35 is severe obesity) on a self-reported bias dataset. The drawback of this research is that only one technique was used to generalize the multiclass problems and the dataset is self-report biased with Body measure index (Height and weight).

Only Body Mass Index (weight and height) is not a perfect measure to determine an accurate obese person as it can sometimes give " gives a false positive" to an athletic people whose high BMIs are not due to excess body fat but to excess muscle.

Hence, it is essential to use more features to build advanced prediction model using machine learning approaches rather than using simpler methods like the manual or statistics approaches. This research will focus on improving an already existing model to obtain a more efficient and less human interference in predicting obesity status using a native dataset with standard components.

Base on this point, this paper considered using a machine learning techniques to build a prediction model of obesity in individuals using a locally available dataset from a health care diagnostic center with principle features. These form the research gap.

Materials and method

Description of existing system

In obesity researches, data mining techniques have been widely engaged to investigate the risk factors. Different advance machine learning algorithms have been used to that effect. Algorithms of K-nearest neighbor, Logistic Regression, Random Forest, Support Vector Machine, Multilayer Perceptron, Adaptive Boosting, Naïve Bayes, Gradient Boosting Classifier, and Decision Tree were used in designing models for prediction. (Ferdowsy et al., 2021)

Review of existing system

In the early years, traditional and clinical procedures are being use in predicting obesity and this can be time-consuming, as it demands the employment of trained physicians in the processes needed to obtain diagnosed result. Most times, complications are due to late intervention as manual methods are mostly carried out when symptoms has manifested in patients. In recently, various researches have been carried out in the area of obesity prediction using machine learning techniques. These techniques being employ by different researchers produced varying results of different accuracy depending on its challenges as its related to the individual.

To diagnosed obesity, physical exams are performed on the patient, then tests are recommended by experts, these test and examinations generally include; taking the patient health history, generally examining the patient physically (checking height, blood pressure, heart rate and temperature), calculating the body mass index (BMI), measuring the waist circumference, lifestyle, hereditary background and checking for other problems. These processes require a high level expertise and lots of time.

Problem with the existing system

The following are the problems facing the existing system

i. Inaccurate obesity diagnosis as patients are not diagnoses by an expert.
ii. Only one feature is used which is the body measure thus prediction might not be correct.
iii. Prone to low accuracy

Description of the proposed system

The proposed model will be build using the Python programming language, applying machine learning algorithm on the dataset collected from a healthcare center to model a system that will be capable of predicting obesity in patients with dataset that has been technically automated to suit the purpose of this research.

Justification of the proposed system

The advancement of machine learning which a branch of artificial intelligence is has led to different researches in different domains to help with lots of process that has been carried out manually. This model when built will help medical personnel in the prediction of obesity in patients with less time involvement. It will also help to handle the prediction of the class or level of obesity which will aid decision for early intervention.

The system will have the following advantages.

i. Increase the reliability of result diagnosed
ii. Reduce the trouble of diagnosis of obesity to the barest minimal
iii. Provide real time system for classification of obesity
iv. Saves time and cut cost required for various medical test.

Feasibility Studies of the Proposed System

The feasibility study of the proposed model assesses its working status, cost restrictions, user requirements, contribution to its related organization, and technical restrictions during development. Based on various physiognomies, the analysis will determine the feasibility study if the system can be implemented. These include:

I. Schedule Feasibility: This is the time taking to finish modeling the project and its functionality to meet expected needs. This will be determined by the availability of resources and its relevance to the field of study while evaluating the timetable feasibility. A Gantt chart can aid the suggested date schedule for the completion of the work.
II. Economic Feasibility: This deals with the system cost effectiveness. The analyses here are to show that the proposed model will be realistic in terms of the cost involved.
III. Operational Feasibility: This will deals with the metric of how successfully the model will interact with the users. The operation of the model will provides a simple user interface that will be practical to improve ease of operation, flexibility, and accuracy of the intended output.
IV. Technical feasibility: This is the assessment of the model to its actual execution as well as its specific technical measurement in terms of technological resources and expertise.

High Level Model of the Proposed System

The high level model of the proposed system has two modules which are patient information management modules and prediction analysis module. The patient information management module is responsible for the documentation of information about the patient while the prediction analysis module is responsible for the prediction of obesity level.

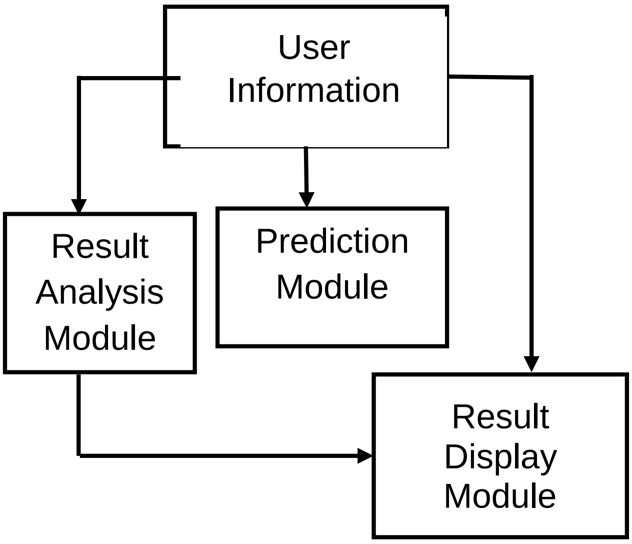


Figure 3.1 High-Level Model of Obesity Prediction Model

- User Information Module: This holds the information of the patient or user which consists of details of the features being exhibited. The features details includes full name, gender, age, family history, physiological, cholesterol level, blood pressure level, body mass index and life style of the patient.
- Prediction Module: This is the model's core module; it takes the patient's details features from the user information module and analyzes it using the created model, producing a prediction result that would be sending to the result analysis module.
- Result Analysis Module: At this module, the outcome of the prediction from the prediction module is interpreted into an understandable format. The patient's overall features will be analyzed and clarifications are given to identify the level at its risk. The interpreted details are sent to the result display module.
- Result Display Module: The patient's personal information (full name, age and gender) from the user information module, the diagnosis outcome from the prediction module along with the remarks on the individual features details

from the analysis module are received at the result display module to give output on the screen.

Data flow Diagram of the Proposed System

The data flow diagram, is the representation of the flow of data from a process to a system, that is, it demonstrated how information flows from a process or system. The data flow diagram provides detailed information on the inputs and the corresponding outputs of each entity and the process itself. With data flow diagram, there are no decision rules, control flow or loops.

Methodology

This aspect discusses the methodology employed to achieve the aim and objectives of the model. A robust, well proven and one of the most used methodologies in machine learning projects is adopted to achieve this goal. It helps to plan, organize and implement machine learning project from one phase to another. The Python programming language, machine learning algorithms and other libraries designed specifically for machine learning researches was used for this work.

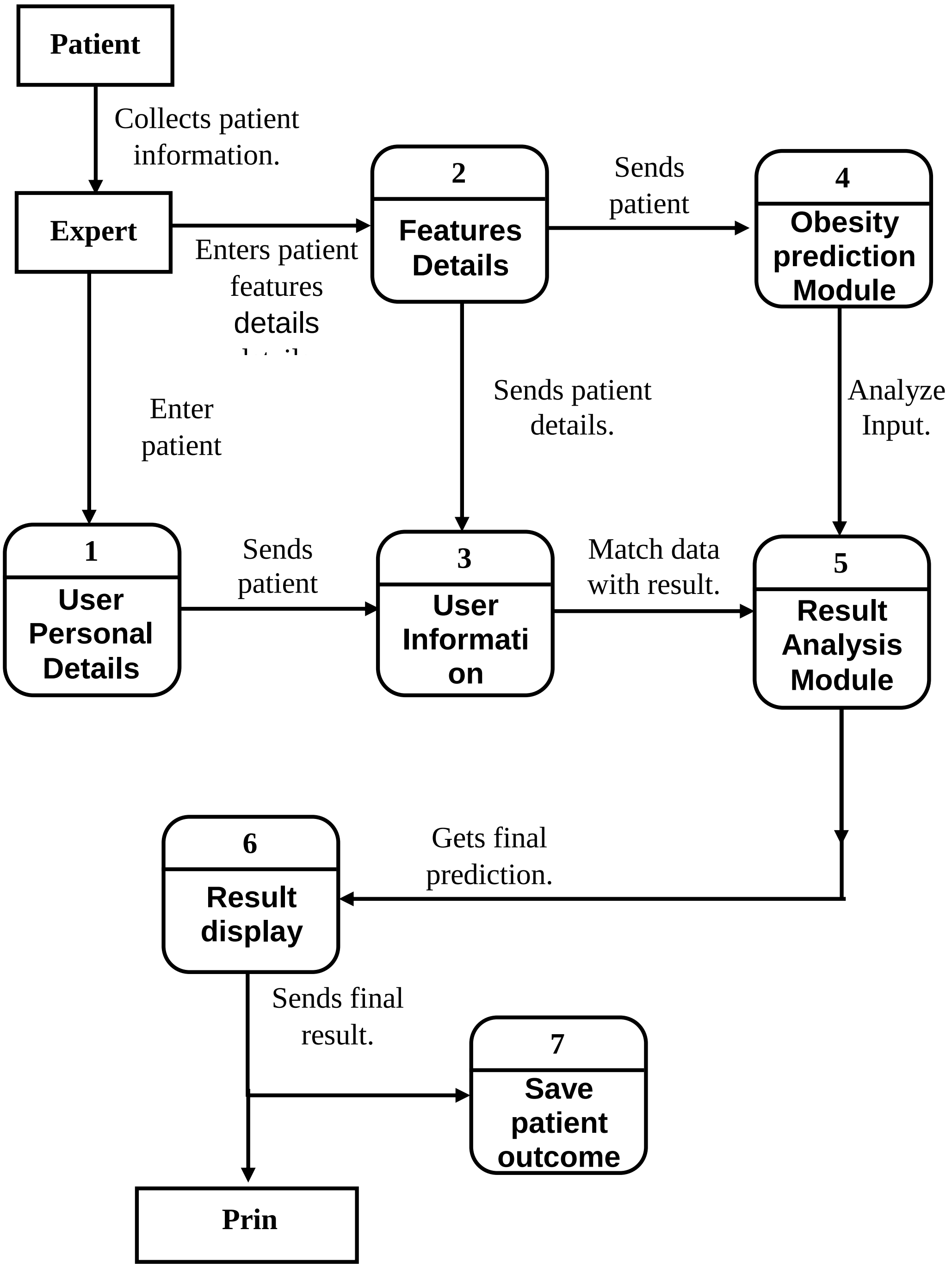


Figure 3.2 The Proposed System data flow

Specification and Justification for the Selected Methodology

The six phase method is an iterative process in which assessment of progress can be measured and evaluated.

The process start from understanding the concept of the research processes to the final stage of modeling as well as deployment. The data for this project was gotten from a healthcare center database which makes the methodology suitable for this research and considering the various processes involved that makes the analysis and building of the proposed model easily achievable. Each phase of the methodology includes:

- Business understanding
- Understanding the Data
- Preparation of the Data
- Modeling
- Evaluation
- Deployment

a. ***Business understanding:*** At this stage we met with expert in order to learn and gain more knowledge about obesity in human i.e. the causes and the symptoms and how the existing methods used in predicting it.

b. ***Understanding the data:*** getting familiarize with the data which includes identifying the quality of data and the relationship between attributes present in the data.

c. ***Preparation of the data:*** preparing the data involve the activities like data cleaning (cleaning the data to confront with the format), data selection (selecting the required features), data reduction (dropping the irrelevant data) and data transformation (transforming the data to machine readable format).

d. ***Modeling:*** this is the core of any machine learning research. This is the stage where the model is being trained and tested to ascertain if the project goal satisfied/achieved. This is where the machine learning algorithms are applied. The machine learning algorithms that was used in this research are;

i. Gboost Classifier
ii. Random Forest Classifier
iii. Decision Tree Classifier
iv. K-Nearest Neighbor
v. Support Vector Machine

Gboost Classifier: Gradient boosting is a class of ensemble algorithm for machine learning that is used for regression or classification prediction modeling problems. At a period, trees are added to the ensemble to fit correctly to the prediction errors made by prior models (boosting) and models fittings using any arbitrary differentiable loss function and gradient descent optimization algorithm brand the techniques its name as gradient boosting (Gboost). (Brownlee, 2020)

Random Forest Classifier: Random forest classifier is a flexible and easy to use machine learning algorithm that can be applied to both classification and regression tasks by building an ensemble of decision trees (multiple) trained by bagging method which is the combination of learning models that can increases the overall result of the model. (Donges, 2021)

Decision Tree Classifier: Decision tree classifier is under the class of supervised learning algorithm used to solve problems of regression and classification. Tree representation is used to solve problem in which each leaf node is corresponded to the attributes and class label are represented on the internal node of the tree. All training dataset are considered as the root before feature values are categorized and used statistical methods for gathering attributes of the internal node. (www.geeksforgeeks.org, 2022)

K-Nearest Neighbor: (Joby, 2021)**,** described KNN (K-Nearest Neighbor) algorithm as a classification method of data used in estimating the likelihood that a data point will become a member of a group or the other based on what group is the nearest pointing data belong to. The k-nearest neighbor algorithm is a supervised machine learning algorithm which can also be used to solve regression and classification problems but mostly used for classification problems because it doesn't perform training when you supply training data and it doesn't perform any calculations. The algorithm looks at the state of a data at which its point near it to determine whether a data point is in group A or B group.

Support Vector Machine*:* a support vector machine which is also known as SVM according to (Dwivedi, 2021)is an essential and versatile algorithm in machine learning that is proficient of task in linear and nonlinear classification, regression and outlier detection. Support vector machine classifier/algorithm is used for both classifications as well as regression problems but widely used for classification task. Because of its less computation and give remarkable accuracy, it is preferred over other classification algorithm.

e. ***Evaluation:*** In this stage the result of our model will be appraised and validated to check the correctness or accuracy of the model. The evaluation metrics of AUC value from our ROC-curve will be used to evaluate the accuracy of the model.

f. ***Deployment****:* This is the final and completed phase where the system is being presented to the end-user. It should be understandable and in useful mode to accomplish.

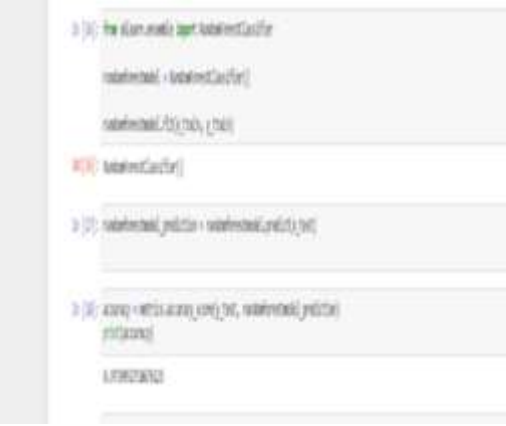
Result

The prediction model was evaluated based on the accuracy of each model to know how well it performs towards achieving the aim. Below are the performance evaluations for the five algorithms used; Gboost Classifier, Random Forest Classifier, Decision Tree Classifier, K-Nearest Neighbor, and Support Vector Machine.

Table 4.1: Model Performance Evaluation

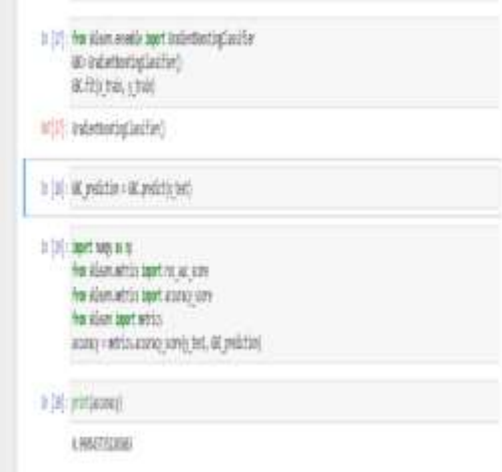
Model	**Accuracy**
Gboost Classifier	99.05%
Random Forest Classifier	96.93%
Decision Tree Classifier	96.93%
K-Nearest Neighbor	95.74%
Support Vector Machine	97.16%

Result for random forest classifier



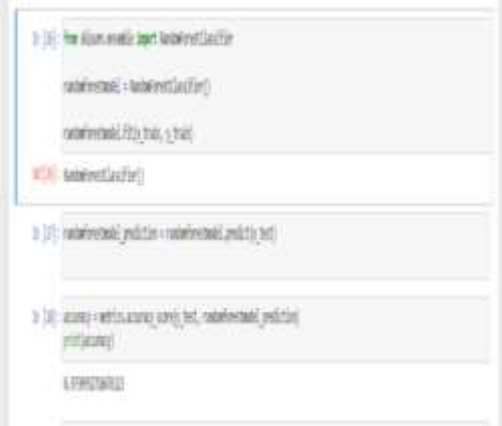
```
from sklearn.ensemble import RandomForestClassifier

randomforestmodel = RandomForestClassifier()

randomforestmodel.fit(x_train, y_train)

RandomForestClassifier()

randomforestmodel_prediction = randomforestmodel.predict(x_test)

accuracy = metrics.accuracy_score(y_test, randomforestmodel_prediction)
print(accuracy)

[illegible]
```

Result for GBoosting



```
from sklearn.ensemble import GradientBoostingClassifier
GBC = GradientBoostingClassifier()
GBC.fit(x_train, y_train)

GradientBoostingClassifier()

GBC_prediction = GBC.predict(x_test)

import numpy as np
from sklearn.metrics import roc_auc_score
from sklearn.metrics import accuracy_score
from sklearn import metrics
accuracy = metrics.accuracy_score(y_test, GBC_prediction)

print(accuracy)

[illegible]
```

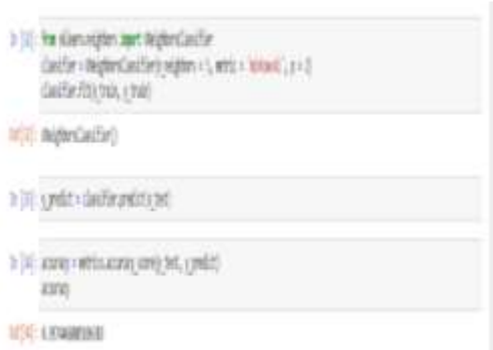
Result for random forest



```
from sklearn.ensemble import RandomForestClassifier

randomforestmodel = RandomForestClassifier()

randomforestmodel.fit(x_train, y_train)

RandomForestClassifier()

randomforestmodel_prediction = randomforestmodel.predict(x_test)

accuracy = metrics.accuracy_score(y_test, randomforestmodel_prediction)
print(accuracy)

[illegible]
```

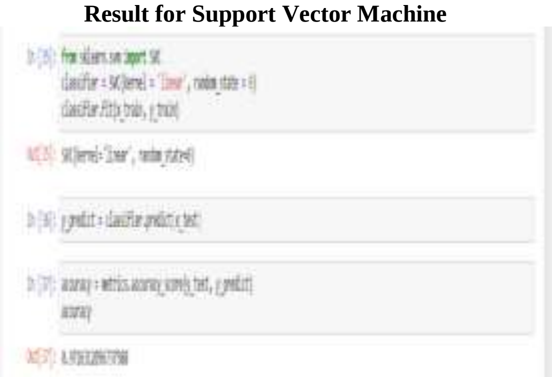
Result for decision tree

Result for KNN



```
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p = 2)
classifier.fit(x_train, y_train)

KNeighborsClassifier()

y_predict = classifier.predict(x_test)

accuracy = metrics.accuracy_score(y_test, y_predict)
accuracy

[illegible]
```

Result for Support Vector Machine



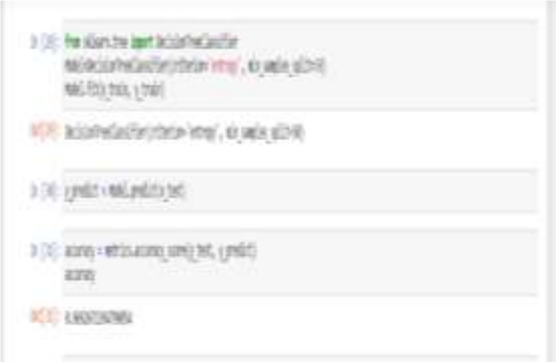
```
from sklearn.svm import SVC
classifier = SVC(kernel = 'linear', random_state = 0)
classifier.fit(x_train, y_train)

SVC(kernel='linear', random_state=0)

y_predict = classifier.predict(x_test)

accuracy = metrics.accuracy_score(y_test, y_predict)
accuracy

[illegible]
```

Conclusion

Diagnosing obesity is difficult, as it is a complex disease that varies in nature. Improvement of the diagnosis of obesity is needed in the health sector to help reduce the risk/implications to the barest minimal. To determine the obesity status of a patient, a physician must carry out some physical assessment on the patient and examine the results of a patient's test to assess prior judgments because it is subject to the doctor's interpretation. This project with the development of the Obesity prediction model would solve these issues of validity of diagnosis, time-consuming factor and also provide a reliable diagnosis system that can be used for all gender.

A reliable and time-saving obesity prediction model has been developed in this research. The machine learning model was designed using the python programming language. Algorithms that were used to achieve the aim of this research are Gradient boosting classifier, random forest classifier, K-nearest neighbor, and Support vector machine. The Gradient boosting algorithm outperformed the other algorithms compared, with an accuracy of 99.05%.

This prediction model is highly recommended for hospitals, clinics, diagnostic centers and the health sector in general, as it will help them to accurately predict obesity status in patients before it gets to a complex stage.



```
from sklearn.tree import DecisionTreeClassifier
model=DecisionTreeClassifier(criterion='entropy', [illegible])
model.fit(x_train, y_train)
```

```
DecisionTreeClassifier(criterion='entropy', [illegible])
```

```
y_predict = model.predict(x_test)
```

```
accuracy = metrics.accuracy_score(y_test, y_predict)
accuracy
```

```
[illegible]
```

Reference

Abdullah, F. S., Manan, N. S., Ahmad, A., Wafa, S. W., & Shahril, M. R. (2016). Data miningtechniques for classification of childhood obecity among year 6 school children. proceeding of the international conference on soft computing and data mining (pp. 465-474). Springer: IEEE Xplore press.

Adnan, M. H. (2011). A framework for childhood obesity classifications and predictions using NBtree. Proceedings of the 7th International Conference on Information Technology in Asia (pp. 1-6). Kuching, Sarawak, Malaysia: IEEE Xplore Press.

Adnan, M. H., Damanhoori, F. D., & Husain, W. (2010). A survey on the usefulness of data mining for childhood obesity prediction. Proceedings of the 8th AsiaPacific Symposium on Information and Telecommunication Technologies (pp. 1-6). Kuching, Malaysia: IEEE Xplore Press.

Brownlee, J. (2020). Ensemble Learning. Machine Learning Mastery.

Davila, P., DeGuzman, C. M., Johnson, K., & Serban. (2015). Estimating prevalence of overweight or obese children and adolescents in small geographic areas using publicly available data. IEEE.

DeGregory, K. W., Patrick , K., DeSilvio, T., & Pleuss, J. D. (2018). A review of machine learning in obesity: machine learning in obesity reaserch. ResearchGate.

Del Cisne, P., & Zhingre, O. (2015). Factors that Influence Obesity. 10-13.

Donges, N. (2021). A Compete Guide to the Random Forest Algorithm. Expert Contributor Network.

Dwivedi, R. (2021, January 29). Machine Learning. United States Artificial Intelligence Institute.

Eduado, D., Fabio, E., & Mendoza, P. (2019). Obesity Level Estimation Software based on Decision Tree. Jornal of Computer Science, 67-77.

Guterrez, H. M. (2010). Diez problems de la poblacion de jalisco. Una perspective sociodeografica, 25-30.

Hagai, R., Smadar, S., Shiri, B.-H., Becca, F., Aron, W., & Eran, S. (2021). Prediction of Childhood Obesity from Nationwide Health Records. The Journal of Pediatrics, 132-140.

Hernandez, J. (2011). Obesity and its causes. International Journal for Medical Image.

Joby, A. (2021, July 19). learn.g2.com. Retrieved from G2 Learn Hub.

Kapil, J., Niyati, B., & Prashant, S. R. (2018). Obesity prediction using ensemble machine learning approaches . Researchgate.

Manna , S., & Jewkes, A. M. (2014). Understanding early childhood obesity risks: An empirical study using fuzzy signatures. IEEE international confluence (pp. 1333-1339). Beijing China : Xplore press.

Tamara, M. D., Mukhopadhyay, S., Aaron , C., & Stephen, M. D. (2015). Machine lerning techniques for prediction on early childhood obesity. Researchgate.

WHO. (2021, June 23). Obesity Related Diseases. Retrieved from World Health Organisation: www.who.int

www.geeksforgeeks.org. (2022). www.geeksforgeeks.org. Retrieved from www.geeksforgeeks.org.

Xiaolu, C., Shuo-yu, L., Jin, L., Shiyong , L., Jun, Z., Peng, N., et al. (2021). Does Physical Activity Predict Obesity-A Machine Learning and Statistical Method-Based Analysis. International Journal of environmental research and public Health.

Xuegin, P., Christopher, B. F., Felice, L.-S., & Aaron, J. M. (2021). Prediction of Early Childhood Obesity with Machine Learning and Electronic Health Record Data. International Journal of Medical Informatics.

Zachary, J., Ward, M. P., Sara, N. B., Angie, L., Cradock, Jessical, L., et al. (2019). Projected US State Level Prevalence of Adult Obesity and Severe Obesity. The New England Journal of Medicine, 2440-2450.