

Anticipating Kid Weight Using Machine Learning

Dr.S.Manju¹, Harish Kumar S²

¹Associate Professor, Dept of Computer Application (MCA)

²Dept of Computer Application (MCA)

^{1,2}PSG College of Arts & Science, Coimbatore, Tamil Nadu, India.

Abstract- *In ultramodern times, rotundity has come a significant trouble each over the world. Rotundity means an unnatural or inordinate quantum of fat that's present in our bodies. People are constantly moving towards an unhealthy life, eating inordinate junk food, late- night sleep, spend a long time sitting down. Especially, adolescents are being affected because of their unconscious stations. It's a medical problem known as a veritably complex complaint. It promotes the spread of complex ails, stroke, heart complaint, liver cancer. In this paper we have referred various research paper to understand the different types of algorithms used for predicting the obesity using machine learning and checks the performance of the model.*

I. INTRODUCTION

Youth weight has arisen as an unavoidable and squeezing worldwide wellbeing challenge in ongoing many years, with significant ramifications for the prosperity of youngsters and the manageability of medical care frameworks around the world. The heightening predominance of life as a youngster heftiness isn't simply a question of actual appearance yet addresses an intricate medical problem with broad outcomes. Weight in youth not just expands the impending gamble of medical conditions yet additionally inclines people toward a heap of ongoing illnesses in adulthood, including diabetes, cardiovascular sicknesses, and specific kinds of malignant growth. In this period of uncommon mechanical headway, where information driven arrangements have altered different areas, bridling the force of Machine Learning (ML) to anticipate kid heftiness has turned into an objective and promising road for general wellbeing research. Exact expectation of stoutness risk during adolescence holds the possibility to change preventive techniques, engage medical care experts with early intercession instruments, and eventually lighten the weight of corpulence related medical problems. According to [10], Food is a wellspring of energy to empower the body's inward capabilities (blood course, body heat age, breath, and so forth) and activities (strolling, running, moving appendages, and so on.). A sound weight can be kept up with reasonable food consumption and the joining of standard activity. By and large, a male with moderate actual work requires 2500 calories each day, though a female with moderate actual work requires

2,000 calories each day. These calorie prerequisites could appear to be copious; in any case, an enormous feast with fries, a burger, and a virus drink could add up to 1,400 calories. Another issue is the absence of actual work in people to adjust the overconsumption of calories. Consequently, the extra calories consumed will add to muscle to fat ratio. Weight doesn't show up precipitously.

It creates with time and, generally speaking, because of unfortunate dietary decisions, which might incorporate eating handled food, cheap food, and focus points; the utilization of food with a high fat and sugar content; drinking liquor and different refreshments; eating out routinely; eating enormous parcels; and eating more than required. ML procedures give an alluring displaying strategy to dissecting early clinical information to foresee later weight in kids. Such procedures can envelop the intricacy of this issue more totally than easier displaying strategies like direct relapse and other measurable methods [7]. Clinical information are innately uproarious, however, and ML procedures give more strong techniques to taking care of missing and inaccurately recorded information. We conjectured that ML could be a powerful way to deal with foreseeing future heftiness among kids more youthful than two years [2].

II. MACHINE LEARNING MODELS

Machine learning models are computational algorithms or mathematical representations that learn patterns and relationships from data. These models can be trained on labeled datasets to make predictions, classify data, or provide insights without being explicitly programmed. Machine learning models are a fundamental component of the broader field of machine learning and play a central role in various applications. Here are some common types of machine learning models:

A. K-NEAREST NEIGHBORS:

The k-Closest Neighbors (k-NN) calculation is a straightforward and natural Machine Learning calculation utilized for both grouping and relapse undertakings. It is a sort of occasion based or languid learning calculation, meaning it doesn't fabricate a model during preparing however rather

retains the whole preparation dataset. While making forecasts on new data of interest, it takes a gander at the "k" closest pieces of information from the preparation set and uses them to decide the result.

B. SUPPORT VECTOR MACHINE:

Support Vector Machine (SVM) is a strong managed machine learning calculation principally utilized for order errands, in spite of the fact that it tends to be reached out for relapse too. SVM means to find the ideal hyperplane that best isolates information guides having a place toward various classes in a manner that expands the edge between the classes.

C. RANDOM TREE:

"Random Tree" can allude to a few distinct ideas or calculations with regards to AI and software engineering. Without explicit setting, it might allude to:

1. Random Decision Tree: An irregular choice tree, otherwise called an arbitrary tree or irregular woodland, is an outfit learning technique. It joins various choice trees to make forecasts. Each tree in the woodland is developed utilizing an irregular subset of the preparation information and an arbitrary subset of the highlights. This haphazardness lessens overfitting and works on the model's presentation and vigor.
2. Randomized Paired Search Tree: In information structures, an irregular tree can allude to a randomized double hunt tree. It's a kind of paired search tree where the request where components are embedded is randomized to guarantee that the tree stays adjusted overall. This keeps up with proficient inquiry and addition times.
3. Irregular Tree in Diagram Hypothesis: In chart hypothesis, an arbitrary tree might allude to different kinds of irregular diagrams or trees produced by unambiguous likelihood conveyances. For instance, a "irregular traversing tree" of a diagram is a tree that traverses all vertices of the chart and is picked indiscriminately.
4. Irregular Tree in Game Hypothesis: In game hypothesis, an irregular tree can be utilized to address methodologies in specific games. The tree structure examines the potential results of various moves in a game, taking into account the component of haphazardness or vulnerability.

D. ID3:

ID3, which means "Iterative Dichotomiser 3," is an exemplary machine learning calculation utilized for choice tree development. It was created by Ross Quinlan during the 1980s and is basically utilized for characterization errands. ID3 is one of the essential calculations that prepared for further developed choice tree calculations like C4.5 and CART.

E. J48:

J48 is a machine learning calculation and a well known execution of the C4.5 calculation, which is utilized for building choice trees. C4.5 was initially evolved by Ross Quinlan during the 1990s as a development of his previous ID3 calculation. J48 is explicitly the Java execution of C4.5, and it is a piece of the Weka machine learning programming suite, which is generally utilized for information mining and machine learning undertakings.

F. NAÏVE BAYES:

Naïve Bayes is a basic yet powerful machine learning calculation utilized for grouping errands, particularly in the domain of text and record order. It depends on Bayes' hypothesis and the presumption of autonomy among highlights, which is the reason it's classified "guileless." Regardless of its straightforwardness, it frequently performs shockingly well in different true applications.

G. BAGGING:

Bagging, short for Bootstrap Amassing, is an outfit machine learning strategy used to work on the presentation and vigor of different calculations, particularly choice trees and different models that are delicate to the difference in the information. Bagging works by making different subsets of the preparation information through resampling and preparing a different model on every subset.

H. MULTILAYER PERCEPTRON:

A Multilayer Perceptron (MLP) is a type of artificial neural network used in machine learning and deep learning. It is a feedforward neural network consisting of multiple layers of interconnected nodes or neurons. MLPs are a fundamental building block of deep learning models and are used for various tasks, including classification, regression, and pattern recognition.

I. GENETIC ALGORITHMS:

Genetic algorithms (GAs) are a type of optimization and search algorithm inspired by the process of natural selection and genetics. They are used to find approximate solutions to complex optimization and search problems by mimicking the process of evolution.

III. ANALYSIS OF PREDICTION MODELS:

S.NO	Author Name	Published Year	Model Used for Prediction	Performance Obtained
1	Dugan	2015	Random Tree, Random Forest, ID3, J48, Naive Bayes, and Bayes Net trained on CHICA which are clinical decision support system[2].	They got the best performance from the model ID3, which was highly accurate at 85% and sensitive at nearly 90%[2].
2	Jindal	2018	They used Python interface also used leverages generalized linear model, random forest, and partial least squares for their prediction model[6].	Their predicted value of obesity was 89.68%[6].
3	Montañez	2017	Their applied machine learning algorithm was SVM algorithm, decision tree, decision rule, and k-NN algorithm to predict susceptibility to chronic hepatitis using SNPs data. From those algorithms, SVM gave the best result for their prediction model[8].	Their simulation result showed that SVM generated the highest area under the curve value of 90.5%[8].
4	Adnan	2012	It works on a hybrid approach for prediction and parameter optimization using Naive Bayes and genetic algorithms[9].	They got the highest accuracy from genetic algorithm optimization. From their implementation, they identified a weakness of the Naive Bayes algorithm which is known as "zero value parameters". Their initial test presented that their structure was usable and that it accurately observed 92% of the zero value parameter samples[9].
5	Singh and Tawfik	2020	Their applied algorithms were k-NN, J48 pruned tree, Random forest, and Bagging, support vector machine, multilayer perception, and voting the effectiveness of all the algorithms was tested on a sample of an unaltered, unbalanced dataset[11].	The precision value is 96% for the MLP algorithm. The result of the F1-score was 93.96%[11].

6	Gerl	2019	Works to predict different measures of obesity based on a large population cohort[4].	For BFP, they identified a perplexing signature and also could predict 8% of the full range of BFP with error and interpret 73% of its variants based on age, gender, and lipids[4].
7	Dunstan	2019	SVM, RF, and extreme gradient boosting are applied for their model[3].	Their imitate confirm that they used the five categories, approximately 60% of the countries considered, and 10% (with respect to a complete prevalence range), and below 20% for the 87% of countries can predict an obesity prevalence with an absolute error[3].
8	Hammond	2019	They applied logistic regression, used another application of random forest classifier, and gradient boosting model for predicting their dichotomous measures of low obese/ medium obese/ high obese for prophecy their continuous BMI values they employed LASSO regression[5].	They ran the bootstrap 100 times to get the best performance of the models for the final model[5].
9	Borrell and Samuel	2014	Using Cox proportional hazards regression they estimate the rate advancement period for all cause and the rate of dying and depending on normal-weight counter parts they estimate CVD-specific mortality for adults who are suffering overweight and obese[1].	They state that The CVD mortality rate in obese adults was at least more than 20% which is compared with normal weight adults[1].

IV. CONCLUSION

In this paper we have referred various research paper to understand the different types of algorithm used for predicting the obesity level using machine learning approach. We conclude that in Table (1) the highest performance rate obtained at Singh and Tawfik paper which was published in the year 2020. Their precision value is 96% for the MLP algorithm. The result of the F1-score was 93.96%.

REFERENCES

[1] Borrell, L.N., Samuel, L., 2014. Body mass index categories and mortality risk in us adults: the effect of overweight and obesity on advancing death. Am. J. Public Health 104 (3).

[2] Dugan, T.M., Mukhopadhyay, S., Carroll, A., Downs, S., 2015. Machine learning techniques for prediction of early childhood obesity. Appl. Clin. Inform. 6 (3).

[3] Dunstan, J., Aguirre, M., Bastías, M., Glass, T.A, Tobar, F., 2019. Predicting nationwide obesity from food sales using machine learning. Health Inf. J. 26 (issue:1), 652–663 pages.

[4] Gerl, M.J., Klose, C., Surma, M.A., Fernandez, C., Melander, O., Männistö, S., Borodulin, K., Havulinna, A.S., Salomaa, V., Ikonen, E., Cannistraci, C.V., Simons, K., 2019. Machine learning of human plasma lipidomes

- for obesity estimation in a large population cohort. Journal.
- [5] Hammond, R., Athanasiadou, R., Curado, S., Aphinyanaphongs, Y., Abrams, C., Messito, M.J., Gross, R., Katzow, M., Jay, M., Razavian, N., Elbel, B., 2019. Correction: Predicting childhood obesity using electronic health records and publicly available data. J., PLS ONE 14 (4).
- [6] Jindal,K, Baliyan, N., Rana, P.S., 2018. Obesity prediction using ensemble machine learning approaches.In: Proceedings of the 5th ICACNI 2017, 2, pp.355-362.
- [7] Michie D, Spiegelhalter DJ, Taylor CC. Machine learning, neural and statistical classification. 1994.
- [8] Montañez, C.A.C., et al., 2017. Machine learning approaches for the prediction of obesity using publicly available genetic profiles. In: 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, pp. 2743–2750. doi:10.1109/IJCNN.2017.7966194.
- [9] Muhamad Adnan, M.H.B., Husain, W., Rashid, N.Abdul, 2012. A hybrid approach using Naïve Bayes and Genetic Algorithm for childhood obesity prediction. In: 2012 International Conference on Computer & Information Science (ICCIS), Kuala Lumpur, Malaysia, pp. 281–285. doi:10.1109/ICCISci.2012.6297254s.
- [10] Saeed Ali Alsareii, Muhammad Awais, Abdulrahman ManaaAlamri, Mansour Yousef AlAsmari, Muhammad Irfan, Mohsin Raza and Umer Manzoor.,2023. Machine-Learning-Enabled Obesity Level Prediction Through Electronic Health Records. DOI: 10.32604/csse.2023.035687
- [11]Singh and B., Tawfik, H., 2020. Machine learning approach for the early prediction of the risk of overweight and obesity in young people. Int. Conf. Comput. Sci. 12140, 523–535.