International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI: 10.55041/ISJEM05660
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

AI Driven Body Composition Risk Assessment Model

Dr.SHANMUGAPRIYA VELMURUGAN[1] ~MOHAMMED SHAMEER.K[2] ~ TAMIL SELVAN.V [3]

[1] Assistant Professor Department of Computer Science,
[2]UG Student (III B.Sc. Computer Science), Department of Computer Science,
[3] UG Student (III B.Sc. Computer Science), Department of Computer Science,
Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

tamilselvanv23bcs059@skasc.ac.in, mohammedshameerk23bcs033@skasc.ac.in, shanmugapriyav@skasc.ac.in

Abstract—This project is entitled as "AI DRIVEN BODY COMPOSITION RISK ASSESSMENT MODEL". Overweight is a major health concern influenced by dietary habits, physical activity, and lifestyle behaviours. Predicting an individual's weight condition at an early stage can help in promoting healthier living and reducing future health risks. This project focuses on overweight prediction by identifying whether a person falls under low weight, normal weight, or overweight categories using machine learning techniques. The dataset includes attributes such as age, gender, height, weight, family history, food consumption patterns, water intake, physical activity, smoking habits, alcohol consumption, and daily lifestyle activities. A Random Forest algorithm is employed to analyze the relationship between these factors and body weight outcomes, as it provides reliable performance with diverse input features. The developed model predicts the weight condition effectively based on personal and lifestyle information, supporting health awareness and informed decision-making related to overweight prevention.

1. INTRODUCTION

This project is entitled as "AI DRIVEN BODY COMPOSITION RISK ASSESSMENT MODEL". Overweight is a major health concern influenced by dietary habits, physical activity, and lifestyle behaviors. Predicting an individual's weight condition at an early stage can help in promoting healthier living and reducing future health risks. This project focuses on overweight prediction by identifying whether a person falls under low weight, normal weight, or overweight categories using machine learning techniques. The dataset includes attributes such as age, gender, height, weight, family history, food consumption patterns, water intake, physical activity, smoking habits, alcohol consumption, and daily lifestyle activities. A Random Forest algorithm is employed to analyze the relationship between these factors and body weight outcomes, as it provides reliable performance with diverse input features. The developed model predicts the weight condition effectively based on personal and lifestyle information, supporting health awareness and informed decision-making related to overweight prevention.

1.1 OBJECTIVE

Overweight has emerged as a significant global health challenge, contributing to numerous chronic diseases such as diabetes, cardiovascular diseases, and certain types of cancer. The complex nature of overweight involves a combination of factors including genetic predisposition, diet, physical activity, and lifestyle choices. Early identification of individuals at risk of overweight or those already in a particular weight category can play a crucial role in promoting preventive measures, early interventions, and healthier lifestyle adoption. Traditionally, determining an individual's weight condition involved simple metrics such as Body Mass Index (BMI), which is calculated

Using height and weight alone. However, BMI does not account for variations in body composition, such as muscle mass versus fat mass, and does not consider the full range of lifestyle factors that contribute to overweight. With advancements in machine learning, there is a growing opportunity to improve overweight risk prediction by considering a broader set of factors, such as diet, physical activity, water intake, smoking, alcohol consumption, and family history, alongside the individual's basic physical 1 measurements.

This project aims to develop an AI-driven model that predicts an individual's weight

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI: 10.55041/ISJEM05660
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

condition—low weight, normal weight, or overweight—based on various personal, physical, and lifestyle attributes.

1.2 Statements and control flow

The assignment statement (token '=', the equals sign). This operates differently than in traditional imperative programming languages, and this fundamental mechanism (including the nature of Python's version of variables) illuminates many other features of the language.

Assignment in C, e.g., x = 2, translates to "typed variable name x receives a copy of numeric value 2". The (right-hand) value is copied into an allocated storage location for which the (left hand) variable name is the symbolic address. The memory allocated to the variable is large enough (potentially quite large) for the declared type.

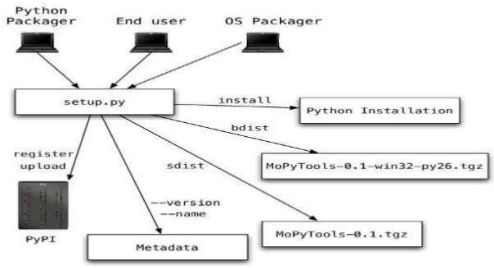
This is termed binding the name to the object. Since the name's storage location doesn't contain the indicated value, it is improper to call it a variable. Names may be subsequently rebound at any time to objects of greatly varying types, including strings, procedures, complex objects with data and methods, etc.

1.3 Distutils Basics and Design Flaws

Distutils contains commands, each of which is a class with a run method that can be called with some options. Distutils also provides a Distribution class that contains global values every command can look at.

setup.py is therefore how everyone interacts with the project, whether to build, package, publish, or install it. The developer describes the content of his project through options passed to a function, and uses that file for all his packaging tasks. The file is also used by installers to install the project on a target system.

Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it.This is similar to PERL and PHP.
Python is Interactive − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.



Python is a Beginner's Language − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to www browsers to games.

2. METHODOLOGY

The methodology of this project involves collecting and analyzing a dataset containing various health and lifestyle attributes such as age, gender, height, weight, family history of overweight, food consumption habits, water intake, physical activity level, smoking habits, alcohol consumption, and daily lifestyle behaviors. The collected data is first preprocessed by handling missing values, converting categorical variables into numerical format, and normalizing the data to improve model performance. After preprocessing, the dataset is divided into training and testing sets to evaluate the model effectively.

A **Random Forest machine learning algorithm** is used to build the prediction model because of its ability to handle multiple input features and provide accurate classification results. The model is trained using the training dataset and then tested to predict the body weight category of individuals as **underweight, normal weight, or overweight**. Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. It was created by Guido van Rossum during 1985- 1990. Like Perl, Python source code is also available under the GNU General Public License (GPL). This tutorial gives enough understanding on Python programming language.

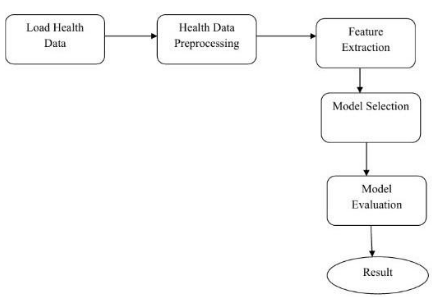
ISSN: 2583-6129
DOI: 10.55041/ISJEM05660

Table - : Prediction

Field Name	Data Type	Size	Description
PID	int	5	Prediction ID
UEmail	varchar	20	User Email ID
Uage	int	3	User Age
UGen	varchar	6	User Gender
UHeight	float	3	User Height
UWeight	float	3	User Weight
Calc	varchar	10	Alcohol Consumption Frequency
Favc	varchar	10	Frequent Consumption of High Caloric food
Ncp	int	3	Number of Main Meals
Scc	varchar	3	Calories Consumption Monitoring
Smoke	varchar	3	Smoke Habit
Ch20	int	3	Daily Water Consumption
Fhistory	varchar	3	Family History
PResult	varchar	50	Prediction Result

2.1 ARCHITECTURE DIAGRAM

The provided diagram illustrates a common workflow for developing a machine learning model using health data. This general scheme is often used in research and applications like predictive maintenance or identifying health- related entities.



2.2 IMPORTANCE OF THE STUDY

The importance of this study lies in its ability to support early detection of unhealthy body weight conditions using artificial intelligence and machine learning techniques. With the increasing prevalence of overweight and obesity worldwide, traditional methods of monitoring body composition are often limited to simple measurements like Body Mass Index (BMI).

However, BMI alone may not always provide a complete picture of a person's health status. By incorporating additional factors such as lifestyle habits, dietary patterns, physical activity levels, and family history, this AI-driven model can provide a more comprehensive and accurate assessment of an individual's body composition risk.

The model helps healthcare professionals and individuals identify potential health risks at an early stage and encourages preventive actions such as improved diet, increased physical activity, and healthier lifestyle choices. Ultimately, this study contributes to better health awareness, supports data-driven decision making, and promotes proactive management of body weight and related health risks.

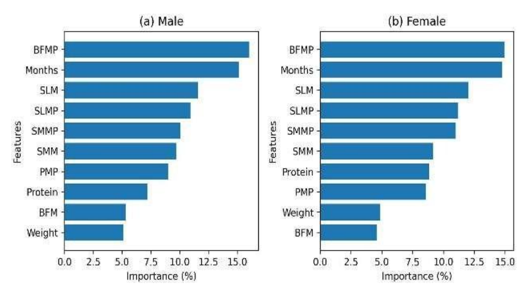
2.3 CHALLENGES

Data Quality and Availability:One of the major challenges in developing the AI-driven body composition risk assessment model is obtaining accurate and reliable data. Health- related datasets often contain missing values, inconsistent entries, or imbalanced categories. Cleaning and preprocessing such data is necessary to ensure that the machine learning model produces reliable predictions..

FeatureSelection:The dataset includes many attributes such as age, gender, height, weight, dietary habits, and lifestyle factors. Identifying which features significantly influence body composition is challenging. Irrelevant or redundant features can reduce the model's accuracy and performance.

Model Accuracy and Overfitting:Machine learning models may perform well on training data but poorly on new, unseen data. Avoiding overfitting while maintaining high prediction accuracy is a key challenge. Proper validation techniques and parameter tuning are required to build a robust model.

Interpretability of Results:Although algorithms like Random Forest provide strong predictive performance, interpreting the results in a way that healthcare professionals and users can easily understand is sometimes difficult. Presenting the outcomes clearly and meaningfully is important for practical use.



3. FUTURE ENHANCEMENT

While the developed AI-driven overweight prediction model demonstrates significant potential, there are several avenues for future enhancement and expansion. One key area of improvement is the refinement of the model's accuracy. Although the Random Forest algorithm provides strong performance, exploring additional machine learning techniques, such as deep learning models or ensemble methods, could further enhance predictive accuracy, especially when dealing with more complex and heterogeneous datasets. Incorporating more diverse and granular data, including factors such as sleep patterns, stress levels, and genetic predispositions, could lead to a more comprehensive assessment of overweight risk.

4. CONCLUSION

AI-driven overweight prediction model developed in this project provides a promising approach to assessing an individual's weight condition based on a comprehensive set of physical, personal, and lifestyle-related factors. By leveraging machine learning techniques such as the Random Forest algorithm, the system offers a reliable, data driven prediction of whether an individual falls under low weight, normal weight, or overweight categories. This model not only improves prediction accuracy but also simplifies the process of weight assessment by analyzing various interconnected factors beyond traditional methods like BMI.

The proposed system addresses several limitations of existing overweight prediction models by reducing computational complexity, focusing on straightforward weight categories, and enhancing usability for everyday health awareness. It offers a more holistic approach to overweight prevention by considering a wider range of factors, such as dietary habits, physical activity, and behavioral patterns, alongside standard physical metrics. The model's ability to handle large and diverse datasets ensures its applicability to a broad population, making it a valuable tool for public health initiatives and individual health management.

5. REFERENCES

[1] Lutz, M. (2013). Learning Python (5th Edition). O'Reilly Media. Available at: https://www.oreilly.com/library/view/learning-python/9781449355722/

[2] Tibbits, S., van der Harten, A., & Baer, S. (2011). Rhino Python Primer (3rd Edition). Robert McNeel & Associates. Available at: https://developer.rhino3d.com/guides/rhinopytho n/

[3] Downey, A. B. (2015). Think Python: How to Think Like a Computer Scientist (2nd Edition). O'Reilly Media. Available at: https://greenteapress.com/wp/think-python/

[4] Wilson, G. (2015). Data Crunching: Solve Everyday Problems Using Java, Python, and More. Pragmatic Bookshelf. Available at: https://pragprog.com/titles/gwd/data-crunching/

[5] van Rossum, G., & Drake, F. L. (2009). The Python Tutorial: An Introduction to Python. Network TheoryLtd. Available at: https://docs.python.org/3/tutorial/

[6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. Available at: http://www.deeplearningbook.org/

[7] Chollet, F. (2017). Deep Learning with Python. Manning Publications. Available at: https://www.manning.com/books/deep-learning- with-python

[8] Ketkar, N. (2017). Deep Learning with Python: A Hands-on Introduction. Springer. Available at: https://link.springer.com/book/10.1007/978-1- 4842-2766-4

[9] Chun, W. J. (2006). Core Python Programming (2nd Edition). Prentice Hall. Available at: https://www.pearson.com/en-us/subject-catalog/p/core-python-programming/P200000003480

[10] VanderPlas, J. (2016). Python Data Science Handbook. O'Reilly Media. Available at: https://jakevdp.github.io/PythonDataScienceHan dbook/