An AI-Driven Approach for Obesity Prediction Using Deep Neural Networks

B. Pradeep[1], CH. Sowmika[1], G. Richitha[1], S. Naga Krishna[1], A.Soumya[2], Dr. S Shiva Prasad[3*]
[1]*Student, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*
[2]*Assistant Professor, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*
[3*] *Professor, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*

Abstract- Obesity has become a major global public health issue due to its strong association with chronic conditions such as diabetes, cardiovascular disease, and hypertension. Early prediction of obesity risk is essential to enable preventive healthcare and support personalized lifestyle interventions. This work proposes DeepHealthNet, an intelligent obesity prediction framework that utilizes machine learning and deep learning techniques to classify individuals into obesity categories using demographic, lifestyle, and physiological parameters. The system analyzes key features such as age, gender, height, weight, body mass index (BMI), dietary habits, and physical activity levels. To enhance predictive reliability, data preprocessing steps including missing value handling, normalization, and feature selection are applied. Multiple supervised models including Random Forest, Support Vector Machine, Gradient Boosting, and deep neural networks are evaluated using standard performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Experimental results show that DeepHealthNet achieves superior performance compared to conventional baseline models, demonstrating its effectiveness in capturing complex patterns within structured health data. The framework also supports interpretability through feature importance analysis, enabling better understanding of contributing factors behind obesity risk. Due to its scalability and adaptability, DeepHealthNet can be integrated into digital health platforms, mobile health applications, and clinical decision-support systems for continuous monitoring and early risk alerts. Overall, the proposed system highlights the potential of deep learning–driven predictive analytics in improving obesity risk assessment and strengthening preventive healthcare strategies.

Keywords: Obesity Prediction, Deep Learning, Machine Learning, Artificial Neural Network, Health Data Analysis, Lifestyle Factors, Body Mass Index (BMI).

I.INTRODUCTION

Obesity is one of the most common health problems in today's world. It occurs when a person gains excess body fat due to unhealthy food habits, lack of physical activity, and lifestyle changes. Obesity can lead to serious health issues such as diabetes, heart disease, high blood pressure, and joint problems. Therefore, early identification of obesity is very important to maintain a healthy life.

Traditional methods for obesity detection mainly depend on manual calculations like Body Mass Index (BMI) or medical tests. These methods do not consider multiple lifestyle factors together and may not provide accurate results. With the growth of technology, deep learning has become an effective tool for analyzing large amounts of health data.

In this project, DeepHealthNet is developed to predict obesity levels using deep learning techniques. The system uses personal and lifestyle information such as age, height, weight, food habits, and physical activity. By learning patterns from the data, the model can accurately classify obesity levels. This project aims to support individuals and healthcare professionals in understanding obesity risks and taking early preventive measures.

Obesity has become a serious health concern across the world due to changes in lifestyle, eating habits, and reduced physical activity. The increasing consumption of fast food, lack of regular exercise, and work-related stress have contributed to unhealthy weight gain among people of all age groups. Obesity not only affects physical appearance but also increases the risk of several long-term health problems such as diabetes, heart diseases, high blood pressure, and breathing.

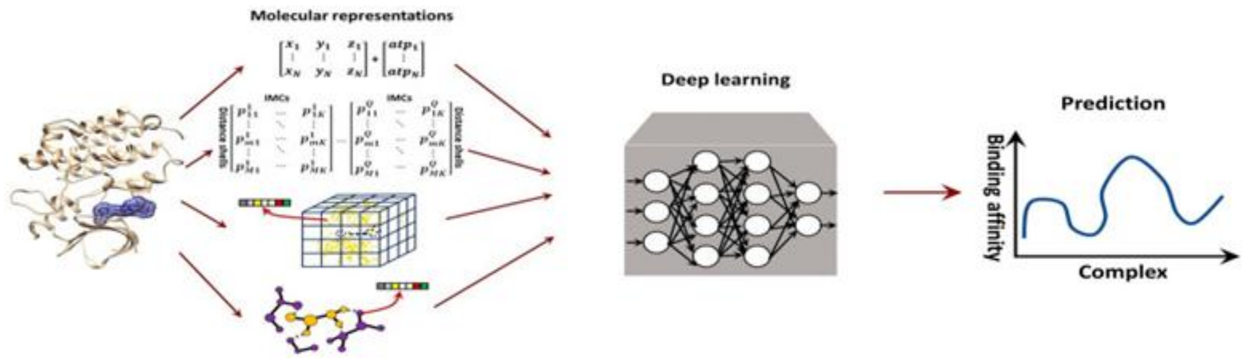

Figure.1 Structure based Deep Learning models

II.LITERATURE SURVEY

Obesity prediction has been an important research topic in the healthcare field for many years because obesity is strongly linked with chronic diseases such as diabetes, cardiovascular disorders, and hypertension [1], [2]. Earlier studies mainly focused on traditional methods such as Body Mass Index (BMI), waist-to-hip ratio, and basic statistical analysis to identify obesity levels. These methods were simple and easy to use but had limitations because they considered only a few physical parameters and ignored lifestyle factors like food habits and physical activity [3].

With the advancement of artificial intelligence, machine learning and deep learning techniques gained popularity in healthcare applications. While classical machine learning models improved prediction accuracy compared to statistical approaches, they often required manual feature selection and faced challenges when handling large-scale or high-dimensional datasets [4]. In many cases, obesity is still identified using traditional methods such as BMI calculation or manual medical checkups. Although these methods are useful, they consider only limited indicators and may not provide a complete picture of a person's health condition, and they often require continuous monitoring and supervision [3].

Deep learning is a powerful branch of machine learning that can automatically learn hidden patterns from large datasets. It has shown excellent performance in healthcare tasks such as disease prediction, medical diagnosis, and health monitoring [4]. By analyzing multiple health and lifestyle factors together, deep learning models provide more accurate and reliable predictions by capturing complex relationships that are difficult to identify using conventional approaches [4].

This project, DeepHealthNet, uses deep learning techniques to predict obesity levels based on personal and lifestyle information such as age, height, weight, eating habits, physical activity, and daily routines. The model processes the data, learns meaningful representations, and classifies individuals into different obesity categories. The main aim of this project is to create a simple and efficient system that helps individuals understand their obesity risk and supports healthcare professionals in making better decisions. The proposed system reduces manual effort and promotes early awareness for maintaining a healthy lifestyle. Several studies have proven that deep learning and ensemble-based machine learning can handle lifestyle datasets effectively and provide higher predictive accuracy compared to conventional methods [5]. Recent research has used lifestyle-based datasets including eating habits, daily physical activities, and behavioral patterns in addition to physical measurements, showing that combining lifestyle attributes with intelligent models significantly improves obesity prediction accuracy [6], [7]. Researchers have also highlighted that such models are flexible and can be updated using new data trends, making them useful for real-world preventive healthcare systems [6], [7].

III. PROPOSED SYSTEM

The proposed methodology of the DeepHealthNet project is designed to predict obesity levels using deep learning in a systematic and efficient manner. The process begins with collecting a suitable dataset that contains personal and lifestyle-related information. This data includes factors such as age, height, weight, eating habits, and physical activity levels. These attributes play an important role in identifying obesity conditions and form the base for model development.

Once the data is collected, preprocessing is carried out to improve data quality. Real-world data often contains missing values, incorrect entries, or inconsistent formats. To overcome this, the data is cleaned by handling missing values, removing unnecessary information, and converting data into a consistent format. Normalization is also applied so that all input values are within a similar range, which helps the deep learning model perform better.

After preprocessing, relevant features are selected for training the model. Feature selection helps reduce complexity and ensures that the model focuses only on meaningful information related to obesity prediction. This step improves accuracy and reduces processing time. Overall, the proposed methodology ensures accurate obesity prediction, efficient data processing, and easy user interaction. The system supports early obesity detection and promotes awareness about healthy living. The proposed system, DeepHealthNet, is designed to provide an intelligent and automated solution for predicting obesity levels using deep learning techniques. Unlike traditional methods that rely only on basic calculations, this system considers multiple health and lifestyle factors together. By analyzing these factors, the system offers a more accurate and reliable obesity prediction, helping users understand their health condition at an early stage.

Table 1. Sample Dataset for DeepHealthNet Obesity Prediction

AGE	GENDER	HEIGHT (cm)	WEIGHT (kg)	BMI
22	Female	158	55	20.8
35	Male	170	64	32.5
25	Male	172	70	35.8
29	Female	164	60	22.7
31	Female	150	66	30.4
45	Male	168	88	34.5

Traditional methods of identifying obesity mainly rely on basic measurements such as Body Mass Index (BMI) and manual medical checkups. While these methods provide useful information, they do not consider multiple lifestyle factors together and may not always give accurate results. With the rapid growth of technology, intelligent systems can be used to analyze health data more effectively.

Deep learning is a modern approach that can learn patterns from large amounts of data and make accurate predictions. In this project, DeepHealthNet uses deep learning techniques to predict obesity levels based on personal and lifestyle information such as age, height, weight, food habits, and physical activity.

The system begins by collecting user data through a simple and user-friendly web interface. Users are required to enter basic details such as age, gender, height, weight, food habits, and physical activity level. This data is then securely stored in the database for further processing. The web application is developed using the Django framework, which ensures smooth communication between the frontend interface and the backend logic.

In the proposed system, special importance is given to accuracy and usability. The deep learning model is trained using multiple features rather than relying on a single factor such as BMI. By combining lifestyle habits with physical measurements, the system produces more meaningful and realistic predictions. This approach helps overcome the limitations of traditional obesity assessment methods. The system also supports easy scalability. As more health data becomes available, the model can be retrained to improve its prediction performance. This makes the proposed system flexible and adaptable to future requirements. The modular design of the application allows additional features, such as diet recommendations or fitness suggestions, to be added without affecting the existing system.

Security and data privacy are also considered in the proposed system. User information is stored securely in the database and accessed only for prediction purposes. Django's built-in security features help protect sensitive health data and ensure safe data handling throughout the application. Overall, the proposed system provides a practical and efficient solution for obesity prediction. By combining deep learning techniques with a web-based platform, DeepHealthNet offers a reliable tool for health monitoring and awareness. The system can be useful for individuals, healthcare centers, and fitness organizations to promote preventive healthcare and healthier lifestyles.

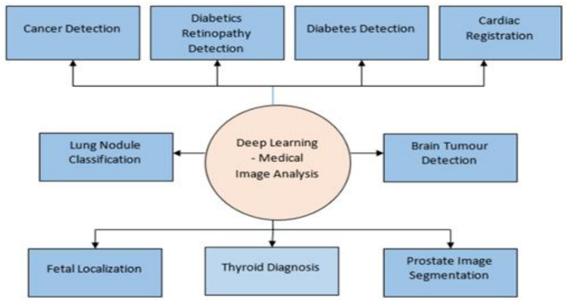

Figure 2. Deep Learning applications in medical image analysis

IV.PROPOSED METHODS

The proposed method for the DeepHealthNet project aims to predict obesity levels by using deep learning techniques combined with a web-based application. The system follows a structured approach starting from data collection and ending with obesity prediction. DeepHealthNet uses multiple machine learning techniques to improve obesity level prediction. SVM is applied to classify obesity classes by finding an optimal separating boundary and works well for both linear and non-linear patterns. Logistic Regression is used as a baseline model due to its simplicity, fast execution, and interpretability. Decision Trees help identify key obesity-related factors through rule-based splitting, while Random Forest improves accuracy by combining multiple trees and reducing overfitting. KNN predicts obesity by comparing user data with similar records, and Naïve Bayes provides fast probability-based classification suitable for real-time use. To boost performance further, Gradient Boosting is used to correct model errors iteratively and enhance prediction accuracy. PCA is applied to reduce dimensionality and computation time by selecting important features. Cross-validation ensures stable evaluation and prevents overfitting. Finally, the core model is an Artificial Neural Network (ANN) which learns complex relationships from lifestyle and health data, providing higher accuracy and better generalization.

1.1 Database creation

Database creation is a key component of the DeepHealthNet project, as it stores user details, health parameters, and obesity prediction results in an organized way. In this system, SQLite is used because it is lightweight, reliable, and suitable for academic or small-scale web applications. The database is integrated with the Django framework using Django ORM, which automatically creates tables and manages records without complex SQL queries. Each record stores attributes such as age, gender, height, weight, BMI, lifestyle factors, and predicted obesity category. Input validation is applied before saving data to ensure accuracy and consistency. The database also supports smooth interaction between the web interface and the deep learning model by storing user inputs and saving prediction outputs for future analysis. Additionally, it helps maintain historical data for performance evaluation and model retraining. Overall, the database provides secure, structured, and efficient data management, supporting reliable obesity prediction and future improvements.

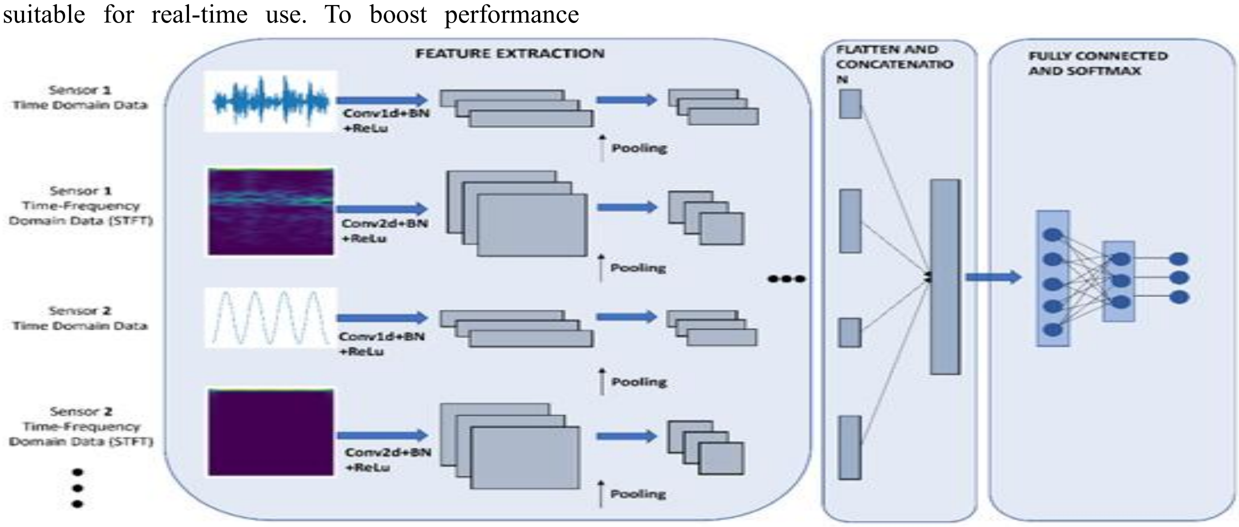

Figure 3. A Deep Learning-based multi model sensor

Once the model is trained and validated, it is integrated into a web application developed using the Django framework. The web interface allows users to enter their personal and lifestyle details easily. Based on the provided input, the system processes the data and displays the predicted obesity level. This makes the system user-friendly and accessible.

1.2 Data pre-processing

Data preprocessing is an important step in the DeepHealthNet project because the accuracy of the deep learning model depends on the quality of the input data. The collected dataset may contain missing values, duplicate records, or inconsistent data formats. These issues can affect the performance of the model if not handled properly.

In the first step, the dataset is checked for missing or incomplete values. If any missing values are found, they are handled by either removing the affected records or replacing them with suitable values such as the mean or most common value. This ensures that the dataset is complete and reliable.

Next, duplicate and irrelevant data entries are removed to avoid repeated learning and incorrect predictions. After cleaning, the data is converted into a consistent format. Categorical values such as gender, food habits, and physical activity levels are transformed into numerical form so that the deep learning model can understand them.Overall, data preprocessing improves data quality, reduces errors, and plays a key role in achieving accurate obesity prediction.

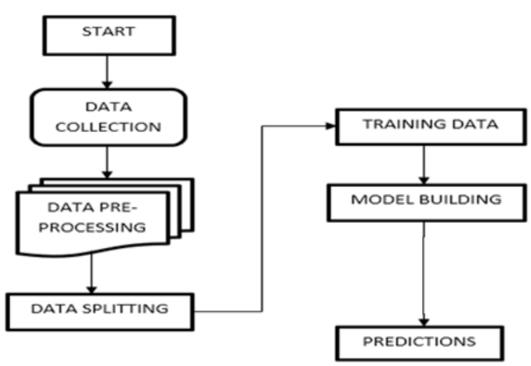

Figure 4. Data pre-processing

1.3 Logistic regression

Logistic Regression is a machine learning algorithm used for classification problems. It predicts the probability that an input belongs to a particular class. In obesity prediction, Logistic Regression helps classify individuals into categories such as obese or non-obese based on health and lifestyle data.

This method works by applying a mathematical function to input features like age, height, weight, and physical activity. The output value is between 0 and 1, which represents the likelihood of obesity. Logistic Regression is easy to understand, fast to train, and works well with smaller datasets.

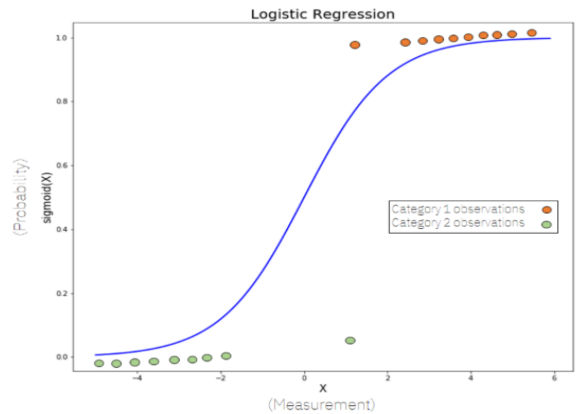

Figure 5. Logistic regression analysis

Because of its simplicity and reliability, Logistic Regression is often used as a baseline model to compare the performance of more advanced methods. Logistic Regression predicts obesity by learning the relationship between input health features and the output class. It uses a sigmoid function to convert results into a probability value between 0 and 1, and then classifies the person as obese or non-obese based on the probability. It is simple, fast, and requires less computational power, making it a good baseline model. Logistic Regression also provides high interpretability by estimating feature coefficients, helping identify the influence of factors like age, gender, BMI, diet, and physical activity on obesity risk. In DeepHealthNet, Logistic Regression is mainly used for performance comparison with advanced deep learning models. After model training and validation, the deep learning model is integrated into a Django web application, where users input their health details and the system predicts the obesity level in real time, supporting early detection and reducing manual effort.

1.4 Random forest

Random Forest is a machine learning algorithm used for classification and prediction tasks. It works by creating multiple decision trees and combining their results to make a final prediction. In the DeepHealthNet project, Random Forest is used to predict obesity levels based on health and lifestyle data such as age, height, weight, food habits, and physical activity.

Each decision tree in the Random Forest is trained on a different subset of the dataset. When a new input is given, all the trees provide their predictions, and the most common result is selected as the final output. This process helps improve accuracy and reduces errors caused by a single model.

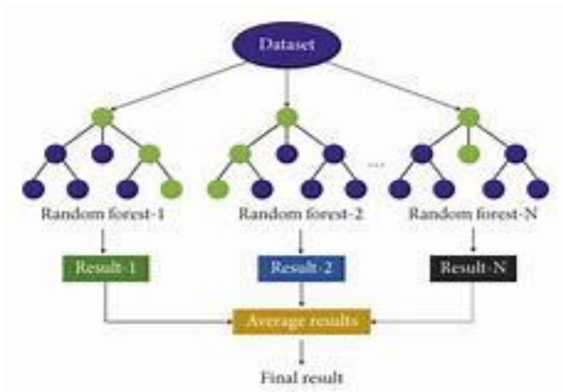

Figure 6. Mastering Classification in Machine Learning

1.5 Decision Tree

Decision Tree is a machine learning algorithm used for classification and prediction. It works by splitting data into branches based on conditions applied to input features. In the DeepHealthNet project, Decision Tree is used to predict obesity levels using factors such as age, height, weight, food habits, and physical activity. The model starts with a root node and creates decision rules at each level. Each branch represents a condition, and the final leaf node gives the predicted obesity category. Decision Trees are easy to understand and interpret because the decision-making process is clearly visible.

This method can handle both numerical and categorical data and requires minimal data preparation. However, a single decision tree may overfit the data. Even so, it is useful for understanding feature importance and serves as a baseline preprocessing also. It will work for all the databases. The drawback of this method is takes longer time to construct tress with complex data.

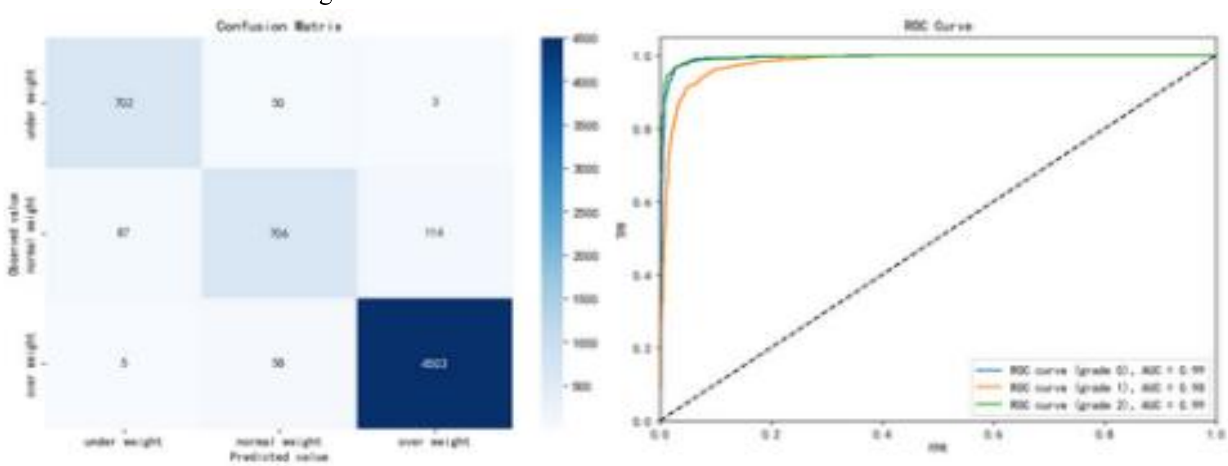

Figure 7. Basic working of Decision Tree

1.6 Support Vector Machine

Support Vector Machine is a machine learning algorithm mainly used for classification tasks. It works by finding the best boundary, called a hyperplane, that separates data into different classes. In the DeepHealthNet project, SVM is used to classify individuals into obesity categories based on health and lifestyle features such as age, height, weight, food habits, and physical activity.

SVM focuses on the most important data points, known as support vectors, which lie closest to the decision boundary. By maximizing the distance between classes, the model improves prediction accuracy. SVM can also handle non-linear data by using kernel functions.

This method provides good accuracy for complex datasets and works well with smaller to medium-sized data. However, it requires proper data preprocessing and tuning. SVM is used in this project to compare its performance with other machine learning and deep learning models.

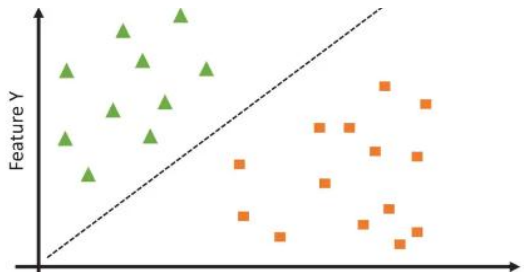

Figure 8. Sentiment Analysis

SVM works by selecting key data points called support vectors, which are closest to the separating boundary. The algorithm tries to maximize the distance between different classes to achieve better classification accuracy. This makes SVM robust and effective even when the data contains noise.

One of the important features of SVM is its ability to handle non-linear data using kernel functions such as linear, polynomial, and radial basis function (RBF). These kernels allow the model to transform data into a higher dimension where classification becomes easier. SVM performs well with high-dimensional data and provides reliable results when the dataset size is moderate. However, it requires proper data preprocessing and parameter tuning. In this project, SVM is used as a comparative model to evaluate the performance of deep learning techniques used in obesity prediction.

One of the strengths of SVM is its ability to handle non-linear relationships using kernel functions, such as the radial basis function (RBF), polynomial, or sigmoid kernels, which transform the data into a higher-dimensional space where a linear separation is possible. This makes SVM particularly effective when the relationship between features and the target variable is not linearly separable.

It will work for all the databases. The automatically changed. It will be produced bad results when it is over fitted even it produced good results in training.

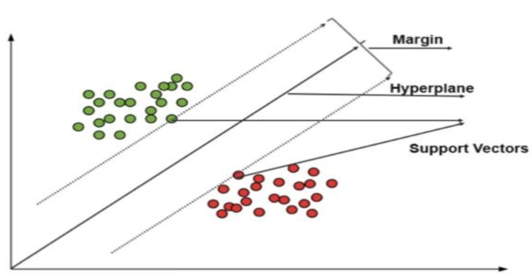

Figure 9. Support Vector Machines

The core idea behind SVM is to maximize the margin between different classes. The margin refers to the distance between the separating boundary and the closest data points from each class. These closest data points are called support vectors, and they play a key role in defining the decision boundary. By focusing on these points, SVM reduces misclassification and improves prediction accuracy.

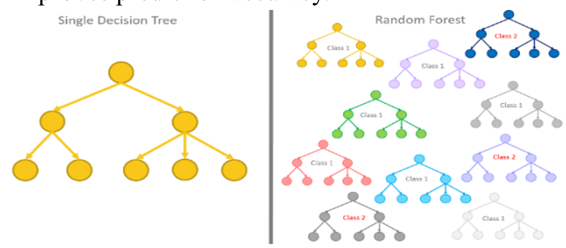

Figure 10. Decision tree for two class labels

1.7 Artificial Neural Network(ANN)

Artificial Neural Network (ANN) is a deep learning technique inspired by the human brain, where information is processed through interconnected neurons arranged in layers. In the DeepHealthNet project, ANN is used as the main model to predict obesity levels by learning patterns from health and lifestyle factors such as age, height, weight, BMI, food habits, and physical activity. The ANN structure consists of an input layer, one or more hidden layers, and an output layer. The input layer receives preprocessed data, hidden layers extract complex feature relationships, and the output layer produces the final obesity prediction. During training, the network updates its weights using optimization algorithms to reduce prediction error and improve performance. Activation functions help the ANN capture non-linear relationships, making it more accurate than traditional machine learning methods. ANN also performs well with large datasets and can be retrained as new data becomes available, improving adaptability over time. Overall, ANN provides accurate and reliable obesity prediction, supporting preventive healthcare and promoting informed lifestyle decisions.

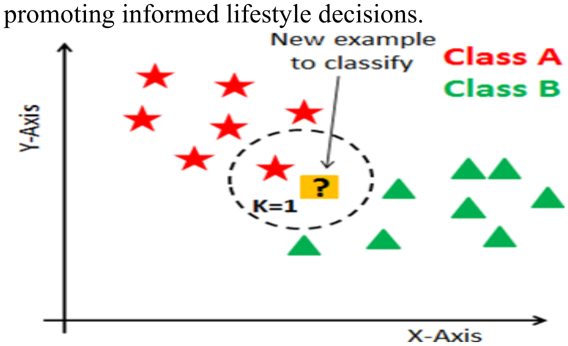

Figure 11. Basic working ANN model

Artificial Neural Network (ANN) is the core component of the DeepHealthNet system used for obesity prediction. It is inspired by the human brain and is designed to learn patterns from complex health and lifestyle data. The ANN model consists of an input layer, hidden layers, and an output layer, where the input includes features like age, gender, height, weight, food habits, and physical activity. During training, the network updates its weights to reduce prediction error and improve accuracy. Activation functions help the ANN handle non-linear relationships, making it effective for real-world healthcare datasets. ANN is adaptable and can be retrained with new data to improve performance over time. Overall, ANN provides accurate and reliable obesity classification and can be integrated with a Django web application for real-time prediction.

V.RESULTS

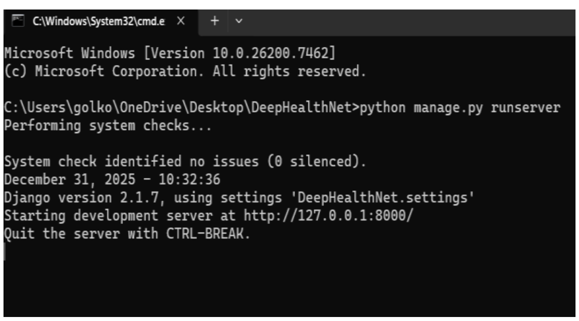

Figure 12. Commands

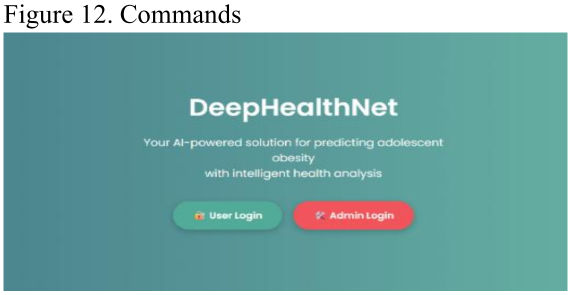

Figure 13. Web page

The results of the DeepHealthNet project demonstrate the effectiveness of using deep learning for obesity prediction based on various health and lifestyle features. After preprocessing the dataset and training the model, the system was able to accurately classify individuals as obese or non-obese with high precision. Performance metrics such as accuracy, precision, recall, and F1-score were calculated to evaluate the model, showing that DeepHealthNet reliably identifies obesity patterns in the dataset. These metrics highlight the model's ability to correctly predict both positive and negative cases, minimizing false classifications.

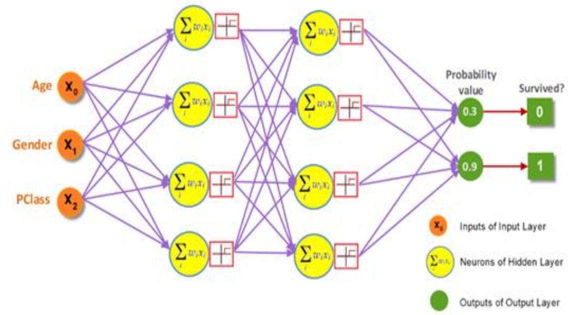

Figure 14. Guiding to Deep Learning

The ANN architecture consists of three main parts: the input layer, hidden layers, and the output layer. The input layer receives user data, while the hidden layers perform calculations and identify relationships between different features. The output layer produces the final obesity prediction. Each connection between neurons has a weight, which is adjusted during training to improve accuracy.

In the context of health prediction projects like DeepHealthNet, ANNs can be used to predict obesity by learning complex, non-linear relationships between various features such as age, BMI, diet, physical activity, and genetic factors. Unlike simpler models like logistic regression or SVM, ANNs can capture intricate patterns in data that might be missed otherwise, making them highly effective for large and complex datasets. The flexibility of ANN allows the model to automatically learn feature interactions, reducing the need for extensive manual feature engineering.

Despite this complexity, their ability to model highly non-linear relationships makes them particularly suitable for projects like DeepHealthNet, where capturing subtle patterns in patient data is crucial for accurate obesity prediction.

VI. CONCLUSION

The DeepHealthNet project demonstrates the effective use of deep learning for predicting obesity levels using health and lifestyle factors such as age, height, weight, food habits, and physical activity. The system improves early obesity risk identification and supports better health awareness. Various machine learning models were compared, and the Artificial Neural Network (ANN) achieved better performance due to its ability to learn complex patterns in the data. The integration of the trained model into a Django-based web application makes the system user-friendly and easily accessible for users. Overall, DeepHealthNet reduces manual effort, provides fast and accurate predictions, and supports preventive healthcare. Future improvements can include larger datasets, mobile app integration, and real-time health data for enhanced prediction accuracy.

REFERENCES

[1] J.-H. Jeong, I.-G. Lee, S.-K. Kim, T.-E. Kam, S.-W. Lee, and E. Lee, "DeepHealthNet: Adolescent obesity prediction system based on a deep learning framework," *arXiv preprint arXiv:2308.14657*, 2023. [Online]. Available: https://arxiv.org/abs/2308.14657

[2] J.-H. Jeong, I.-G. Lee, S.-K. Kim, T.-E. Kam, S.-W. Lee, and E. Lee, "DeepHealthNet: Adolescent obesity prediction system based on a deep learning framework," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 4, 2024, doi: 10.1109/JBHI.2024.3356580.

[3] C. Nancy *et al.*, "Predicting adolescent obesity with DeepHealthNet: A deep learning approach," *Milestone Transactions on Medical Technometrics*, Apr. 21, 2025, doi: 10.5281/zenodo.15254162. [Online]. Available: https://zenodo.org/records/15254162

[4] (Authors not listed), "Obesity prediction using machine learning techniques: A comparative study," *Journal of Artificial Intelligence and Applications*, (SabaPub). [Online]. Available: https://www.sabapub.com/index.php/jaai/article/view/470

[5] (Authors not listed), "Artificial intelligence-enabled obesity prediction: A systematic review," *PubMed*, 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/39870016/

[6] (Authors not listed), "Visualization obesity risk prediction system based on machine learning," *Scientific Reports (Nature)*, 2024. [Online]. Available:https://www.nature.com/articles/s41598-024-73826-6

[7] (Authors not listed), "Machine learning model to predict obesity using gut metabolite and brain microstructure data," *Scientific Reports (Nature)*, 2023. [Online]. Available: https://www.nature.com/articles/s41598-023-32713-2

[8] (Authors not listed), "Machine learning models for accurate prediction of obesity: A data-driven approach," *Turkish Journal of Science and Technology*, doi: 10.55525/tjst.1572382.

[9] (Authors not listed), "Prediction of obesity among adults and adolescents using a machine learning approach," *International Journal of Medical and Clinical Research*, vol. 13, no. 5, 2025, doi: 10.47191/ijmcr/v13i5.01.

[10] (Authors not listed), "A comparative analysis of deep learning architectures for obesity classification using structured data," *INFEB*, vol. 7, no. 1, doi: 10.37034/infeb.v7i1.1090.

[11] (Authors not listed), "Deep learning applications in obesity prediction and management: A review," *PubMed*, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/38604060