A Survey Paper on Obesity Detection

[1] Prof. Akash Mhetre, [2] Preet Pochat, [3]Aniket Dhane, [4]Niranjan Desai, [5]Mitali Bora

Sinhgad Institute of Technology and Science, Pune 411041, Maharashtra, India.

Abstract

Obesity is a global health epidemic with significant implications for public health. Timely and accurate detection of obesity is crucial for effective prevention and intervention. This abstract provides an overview of the current state of obesity detection methods, their challenges, and the potential role of
emerging technologies in addressing these issues.
Traditional methods of obesity detection primarily rely on body mass index (BMI) calculations, which have limitations in accurately assessing body composition and health risks associated with obesity. More advanced approaches, such as dual-energy X-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and magnetic resonance imaging (MRI), offer improved accuracy in assessing body fat percentage, but they are often costly and less accessible in routine clinical settings. Support Vector Machine, Decision Trees, Random
Forest and Logistic Regression can be used for classification. Recent developments in artificial intelligence and machine learning have shown
promise in automating and enhancing obesity detection. These technologies can process diverse data sources, including medical imaging, electronic health records, and wearable devices, to provide a more comprehensive and personalized assessment of an individual's obesity status. Machine learning models can also integrate multiple factors, such as genetics, diet, and physical activity, to offer a holistic view of obesity risk.

Keywords: SVM, DXA, BIA, BMI.

1. Introduction:-

Obesity has become a global health epidemic, with significant implications for public
health, healthcare costs, and overall well-being.
The World Health Organization (WHO) defines obesity as abnormal or excessive fat accumulation that presents a risk to health. It is a complex and multifactorial condition influenced by genetics, environment, lifestyle, and socio-economic factors. Identifying and addressing obesity is critical to reducing the burden of chronic diseases such as diabetes, heart disease, and certain cancers. World Health Organization (WHO) defines overweight and obesity as "abnormal or excessive fat accumulation that presents a risk to health".[1] One of the most challenging healthcare problems that the world is facing today is that of childhood and adolescent obesity. It is not limited to a single country but has become a global public health crisis. For the year 2017-2018 about 14.4 million children and adolescents (aged 2-19 years) were affected by obesity in the United States. Among 2-5 years old, its prevalence was 13.4%, for 6-11 years old, it was 20.3%, and 21.2% for 12-19 years old.

Data also shows that the distribution is not uniform, as certain populations (Hispanic and non-Hispanic Black) are more vulnerable to obesity. Among Hispanic children, the prevalence of obesity was 25.6%, 24.2% among nonHispanic Black children, for non-Hispanic White children it was 16.1%, and 8.7% among non-Hispanic Asian children.[2] In not just the United States, the prevalence of obesity among children and adolescents is increasing all over the world. Globally about 41 million children under the age of 5 were overweight. Children under the age of 5 living in Asia contributed half to this number and about one quarter belonged to Africa.[3] Diseases such as Type 2 diabetes, heart disease, blood vessel disease, and obesity-related depression, and social isolation, usually associated with adults are now being detected in children as well. Once obesity sets in, it is very difficult to treat it as the causes of childhood and adolescent obesity are complex and multifaceted which makes it a very challenging task. Therefore, immediate steps must be taken to prevent obesity. If obesity is not curbed at childhood itself then there is a greater chance of its persistence into adulthood. Adult obese people have a higher risk of developing diabetes, high blood pressure, and heart disease.
Adults are classified as overweight or obese based on Body Mass Index (BMI) which is defined as the weight in kilograms divided by the square of the height in meters (kg/m^2). For children and adolescents, this metric of classification is not that simple because as they grow, their bodies undergo several physiological changes. Factors such as age, puberty, and growth rate influence the rate of fat deposition and removal. As a result, defining a standard for overweight and obesity that incorporates all age groups is very difficult [1]. To assess obesity, three classification systems are used at the international level: International

Obesity Task Force (IOTF), the United States Centers for Disease Control and Prevention (CDC) growth charts issued in 2000, and the World Health Organization (WHO) criteria. International Obesity Task Force (IOTF) [2], constructed in 2000 and updated in 2002 with the help of datasets from 6 countries (Singapore, Netherlands, Brazil, Hong Kong, the UK, the USA), uses sex-specific BMI curves that match adult BMI values of 25kg/m^2 (Overweight) and 30kg/m^2 (Obesity) at 18 years. The World Health Organization (WHO) criteria [3] supplemented with data from the WHO Child Growth Standards for children aged 5 years and younger, was developed in 2007 using the 1977 National Center for Health Statistics (NCHS) growth reference from 5 to 19 years. It defines overweight as a BMI > 1 standard deviation (SD) and obesity as a BMI > 2 SD from the mean of the WHO reference population. CDC growth charts [4] were a revision of the NCHS 1977 growth reference that incorporated data from five national surveys conducted between 1963 and 1994 in the United States of America. This system defines overweight as a BMI > 85th percentile of the reference population and obesity as a BMI > 95th percentile.

2. Literature Survey:-

The paper [1], The postoperative health status of an obesity patient indicates the outcome of the surgical treatment. By each postoperative revisit, physicians need to go through the previous patient records to recall the patient status and to evaluate the postoperative risk of readmission. In order to support in this process, we develop a method to extract indicators and to analyse weight changes, so that potential complications and risks of clinical readmission can be recognized timely. In this paper, we will compare two approaches that are based on traditional machine learning and neural networks. The performance of traditional machine learning on the task of obesity related entity extraction is compared with one variation of attentive recurrent neural networks. We conclude that for processing a small data set using neural networks, a data balancing method should firstly be applied to achieve an extended corpus and a general representation, which can apparently increase the differentiability of the input data. A fine-tuning in the networks can provide further enhancement of the performance.

The paper [2], In this work, we present FREGEX a method for automatically extracting features from biomedical texts based on regular expressions. Using Smith-Waterman and Needleman-Wunsch sequence alignment algorithms, tokens were extracted from biomedical texts and represented by common patterns. Three manually annotated datasets with information on obesity, obesity types, and smoking habits were used to evaluate the effectiveness of the proposed method. Features extracted using consecutive sequences of tokens (ngrams) were used for comparison, and both types of features were mathematically represented using the TF-IDF vector model. Support Vector Machine and Na¨ıve Bayes classifiers were trained, and their performances were ultimately used to assess the ability of the feature extraction methods. Results indicate that features based on regular expressions not only improved the performance of both classifiers in all datasets but also use fewer features than n-grams, especially in those datasets containing information related to anthropometric measures (obesity and obesity types).

The paper [3], In addition to the rapid growth of Machine Learning in biomedical and healthcare communities, the accurate analysis of medical data benefits early disease detection, patient care and community services. However, the analyses accuracy of a disease is reduced when the intake and the quality of medical data is unexplored and incomplete. Obesity happens to be a highly ignored yet a grave health issue that gives rise to a lot of potentially harmful cardiovascular diseases which if neglected could prove fatal. Many people overlook this condition not realizing that by taking small steps to improve their health they can keep many other diseases at bay. Moreover, the underlying problem that persists is not only identifying obesity as a physical medical problem but also speculating the different factors that give rise to it. The key to combatting this real-life problem and staying in the pink of health would be to thoroughly investigate and survey the many reasons as to why people get obese in the first place and eventually nip all the other probable deadly diseases in the bud because prevention is better than cure! This work aims to overcome the above-mentioned limitations by developing a state-of-the-art system that streamlines machine learning algorithms for the effective prediction of Obesity and its related diseases considering the population of India. The government indulgent regulation on food restriction provides easy accessibility to unhealthy, processed food. There felt a need to develop a system that consider parameters affecting an individual physically, internally, mentally, psychologically and emotionally that contribute to the occurrence of Obesity and also suggest healthier alternatives to curb this problem. Some of the common diseases due to Obesity included in the system are Diabetes, Heart Attack, Hypertension, Osteoarthritis and Varicose Veins. The developed system will undeniably be beneficial for predicting obesity, its related diseases and for the future betterment of an individual.

The paper [4], Metabolic disorders such as type 2 diabetes mellitus, obesity and metabolic syndrome have a high incidence in the population of developed countries and require continuous clinical and pharmacological treatments throughout their progression. Obesity is associated with insulin resistance in over 90obese subjects are “protected” from this condition. Infrared spectroscopy has been investigated as a non–invasive tool on biofluids, together with signal processing techniques, in the research of novel and predictive biomarkers. In the following, a study based on saliva profiling using infrared spectroscopy on a population of metabolic abnormal and normal obese compared with control subjects is presented. Analysis has been carried out to design a consistent and standardized protocol for saliva profiling in different molecular regions of interest. Results obtained through an unsupervised classification technique allowed the grouping of patients belonging to a specific population on the basis of the characteristic molecular signatures in the regions of Amide I, glucose and thiocyanate.

The paper [5], In this recent research found that genetics plays an important role in obesity risk analysis besides life styles. Many literatures are focusing on analyzing the effect of Single Nucleotide Polymorphism (SNPs) towards obesity to facilitate personalized medication. However, SNPs data are normally large and noisy, which affects the accuracy and computational complexity on data processing and analysis. Therefore, efficient data reduction is essential to yield better analysis results and reduce computational complexity in the experimentations. In this paper, we investigated feature selection process in obesity related SPNs analysis using Forward attribute reduction based on neighbourhood rough set model (FARNeM).But, FARNeM did not achieve good reduction rate when applied to the experimental data set. However, the overall analysis showed that, it is encouraging to include feature selection process before the learning algorithms.

The paper [6], In this a fuzzy medical diagnostic decision system for helping support to evaluate patients with anginal chest pain and obesity clinical condition is proposed in this paper. Such an approach is based on the Braunwald symptomatic classification, the fuzzy set theory and fuzzy logic, and a risk obesity factor determined by a simplified Fuzzy Body Mass Index (FBMI). The fuzzy Braunwald symptomatic classification intertwined with the fuzzy obesity risk factor overwhelm the current rapid access chest pain clinic approaches that do not discriminate the obesity comorbidity or takes into account the subjectiveness, uncertainty, imprecision, and vagueness concerning such a clinical health condition. The resulting fuzzy obesity–based Braunwald symptomatic chest pain assessment is an alternative to support healthcare professionals in primary health care for patients with anginal chest pain worsened by the obesity clinical condition.

The paper [7], The key to combatting this real-life problem and staying in the pink of health would be to thoroughly investigate and survey the many reasons as to why people get obese in the first place and eventually nip all the other probable deadly diseases in the bud because prevention is better than cure! This work aims to overcome the abovementioned limitations by developing a state-of-the-art system that streamlines machine learning algorithms for the effective prediction of Obesity and its related diseases considering the population of India. The government indulgent regulation on food restriction provides easy accessibility to unhealthy, processed food. Also, India’s growing career-oriented lifestyle gives rise to irregular biological patterns especially in younger generation who prefer indoor games rather than playing outside. The behavioural and socio-psychological factors such as sleep, stress, ethnicity in addition to hormonal disorders are influential in giving rise to obesity in today’s era.

3. Problem Definition: -

Obesity is a complex medical condition characterized by the excessive accumulation of body fat, which can have serious implications for an individual's health and well-being. To address the growing concern of obesity and its associated health risks, there is a need to develop an effective classification system that can accurately identify and categorize individuals into different obesity classes based on their body mass index (BMI) and other relevant features.

4. Objectives: -

1) To accurately categorize individuals based on their body weight and composition to identify and understand the prevalence and severity of obesity within a population.

2) Develop a clear and consistent definition of obesity based on measurable criteria, such as body mass index (BMI), waist circumference, or body fat percentage.

3) Evaluate the accuracy and reliability of different measurement methods and diagnostic tools used in obesity classification.

4) Assess the effectiveness of interventions and treatment strategies for obesity by utilizing the classification system to measure outcomes and compare results across different patient populations.

5. Motivation: -

Motivation for obesity detection is of paramount importance given the increasing
prevalence of obesity worldwide and its significant impact on individual health and society as a whole. Obesity is a major risk factor for numerous chronic diseases, including heart disease, type 2 diabetes, certain cancers, and more. Detecting obesity early allows for proactive measures to reduce the risk of these debilitating and potentially life-threatening conditions.

6. Methodology: -

SVM is a supervised learning algorithm used for both classification and regression tasks. SVM aims to find an optimal hyper-plane that separates the data points of different classes or predicts continuous target values. Decision Trees are versatile supervised learning algorithms used for classification and regression tasks. Decision Trees create a tree-like model of decisions and their possible consequences based on the features in the input data. Random Forest is a popular machine learning algorithm used for classification and regression tasks due to its high accuracy.

7. Conclusion: -

In conclusion, our obesity detection project utilized methods, such as machine learning algorithms, data analysis techniques, etc., to predict and classify obesity based on such as BMI, waist circumference, and other body coordinates. Our findings indicate accuracy of prediction, factors contributing to obesity.

While the accuracy achieved was promising, there were certain limitations encountered during the project. These limitations include availability of data, sample size and many others. Addressing these limitations could enhance the robustness and generalizability of our obesity detection model in future iterations.

Furthermore, it is important to note that obesity is a complex issue influenced by various socio-economic, cultural, and environmental factors. Thus, while our model provides valuable insights, it should be complemented with interdisciplinary approaches for a comprehensive understanding of obesity and its prevention.

Additionally, the application of our obesity detection model could extend beyond clinical settings to public health interventions, facilitating targeted interventions and personalized healthcare strategies.

In conclusion, our obesity detection project contributes to the growing body of knowledge aimed at combating obesity. By leveraging advanced technologies and data-driven approaches, we aim to empower individuals and healthcare professionals in the prevention and management of obesity, ultimately improving public health outcomes.

8. Future Scope: -

The future scope of an obesity detection project involves expanding its applicability, improving accuracy, and addressing emerging challenges. Here are some potential future directions for such a project:

1. Integration of Additional Data Sources: Incorporating diverse data sources such as genetic information, dietary patterns, physical activity levels, and environmental factors can enhance the accuracy and comprehensiveness of obesity detection models. This holistic approach enables a deeper understanding of the multifactorial nature of obesity.

2. Development of Personalized Models: Moving towards personalized obesity detection models tailored to individual characteristics and risk factors can optimize intervention strategies. Machine learning techniques, such as deep learning and reinforcement learning, can be employed to adaptively refine predictions based on ongoing monitoring and feedback.

3. Validation and Generalization Studies: Conducting validation studies on larger and more diverse populations is crucial for assessing the generalizability and reliability of obesity detection models. Collaborations with healthcare institutions and community organizations can facilitate access to diverse datasets, ensuring the robustness of the models across different demographics and settings.

4. Real-Time Monitoring and Intervention: Leveraging wearable devices, mobile applications, and IoT (Internet of Things) technology enables real-time monitoring of physiological parameters and behaviors related to obesity. Integrating predictive analytics and decision support systems can facilitate timely interventions, empowering individuals to make informed lifestyle choices and healthcare professionals to deliver personalized care.

References: -

[1] HERA SIDDIQUI AJITA RATTANI School of Computing Department of Public Health Sciences. "A Survey on Machine and Deep Learning Models For Childhood and Adolescent Obesity" In IEEE Access 2022

[2] Yihan Deng , Peter Dolog, Medical Informatics Bern University of Applied Sciences, Computer Science Aalborg University "Obesity entity extraction from real outpatient records: When learning-based methods meet small imbalanced Medical data sets" In IEEE , 2019

[3] Christopher A. Flores, Rosa L. Figueroa, and Jorge E. Pezo "FREGEX: A Feature Extraction Method for Biomedical Text Classification using Regular Expression" In IEEE 2019

[4] Naomi Christianne Pereira ,Naomi Christianne Pereira, Department of Computer Engineering. "OBESITY RELATED DISEASE PREDICTION FROM HEALTHCARE COMMUNITIES USING MACHINE LEARNING" In IEEE 2019

[5] S.A. Pullano, M. Greco, M.G. Bianco ,Dept. Of Health Sciences University Magna Graecia of Catanzaro "FT–IR Saliva Profiling in Patients with Obesity and Obesity–Related Insulin Resistance" In IEEE 2019

[6] Yulistia Elsa Nursida, Bayu Erfianto, School of Computing Telkom University. "Foot Plantar Pressure to Detect Obesity " In IEEE 2020

[7] Carlos Martinez-Millana, Antonio Martinez-Millana, Carlos Fernandez-Llatas ITACA Universitat Politecnica de Val ' encia ' "Comparing data Base engines for building big data analytics in obesity detection" In IEEE 2019*

[8] Farshid Keivanian, Raymond Chiong, School of Electrical Engineering and Computing, The University of Newcastle. "A Fuzzy Adaptive Binary Global Learning Colonization-MLP model for Body Fat Prediction" In IEEE 2019

[9] SR Rathod, UM Chaskar, Department of Instrumentation and Control Engineering College of Engineering Pune. "Machine Learning Approach to Study the Impact of Obesity on Autonomic Nervous System using Heart Rate Variability Features" In IEEE 2021