

Review

Applications of Artificial Intelligence to Obesity Research: Scoping Review of Methodologies

Ruopeng An^{1*}, MPP, PhD; Jing Shen^{2*}, PhD; Yunyu Xiao³, PhD

¹Brown School, Washington University in St. Louis, St. Louis, MO, United States

²Department of Physical Education, China University of Geosciences, Beijing, China

³Weill Cornell Medical College, Cornell University, Ithaca, NY, United States

*these authors contributed equally

Corresponding Author:

Jing Shen, PhD

Department of Physical Education, China University of Geosciences

No. 29, Xueyuan Road, Haidian District

Beijing, 100083

China

Phone: 86 010 82322397

Email: shenjing@cugb.edu.cn

Abstract

Background: Obesity is a leading cause of preventable death worldwide. Artificial intelligence (AI), characterized by machine learning (ML) and deep learning (DL), has become an indispensable tool in obesity research.

Objective: This scoping review aimed to provide researchers and practitioners with an overview of the AI applications to obesity research, familiarize them with popular ML and DL models, and facilitate the adoption of AI applications.

Methods: We conducted a scoping review in PubMed and Web of Science on the applications of AI to measure, predict, and treat obesity. We summarized and categorized the AI methodologies used in the hope of identifying synergies, patterns, and trends to inform future investigations. We also provided a high-level, beginner-friendly introduction to the core methodologies to facilitate the dissemination and adoption of various AI techniques.

Results: We identified 46 studies that used diverse ML and DL models to assess obesity-related outcomes. The studies found AI models helpful in detecting clinically meaningful patterns of obesity or relationships between specific covariates and weight outcomes. The majority (18/22, 82%) of the studies comparing AI models with conventional statistical approaches found that the AI models achieved higher prediction accuracy on test data. Some (5/46, 11%) of the studies comparing the performances of different AI models revealed mixed results, indicating the high contingency of model performance on the data set and task it was applied to. An accelerating trend of adopting state-of-the-art DL models over standard ML models was observed to address challenging computer vision and natural language processing tasks. We concisely introduced the popular ML and DL models and summarized their specific applications in the studies included in the review.

Conclusions: This study reviewed AI-related methodologies adopted in the obesity literature, particularly ML and DL models applied to tabular, image, and text data. The review also discussed emerging trends such as multimodal or multitask AI models, synthetic data generation, and human-in-the-loop that may witness increasing applications in obesity research.

(*J Med Internet Res* 2022;24(12):e40589) doi: [10.2196/40589](https://doi.org/10.2196/40589)

KEYWORDS

artificial intelligence; deep learning; machine learning; obesity; scoping review

Introduction

Background

The double burden of malnutrition, characterized by the coexistence of overnutrition (eg, overweight and obesity) and

undernutrition (eg, stunting and wasting), is present at all levels of the population: country, city, community, household, and individual [1]. Obesity is a leading cause of preventable death and consumes substantial social resources in many high-income and some low- and middle-income economies [2]. Worldwide, the obesity rate has nearly tripled since 1975 [3]. In 2016, 13%

of the global population, or 650 million adults, were obese [4]. More than 340 million children and adolescents aged 5 to 19 years and 39 million children aged <5 years were overweight or obese [4]. By 2025, the global obesity prevalence is projected to reach 18% among men and 21% among women [5].

Health data are now available to researchers and practitioners in ways and quantities that have never existed before, presenting unprecedented opportunities for advancing health sciences through state-of-the-art data analytics [6]. By contrast, dealing with large-scale, complex, unconventional data (eg, text, image, video, and audio) requires innovative analytic tools and computing power only available in recent years [7,8]. Artificial intelligence (AI), characterized by machine learning (ML) and deep learning (DL), has become increasingly recognized as an indispensable tool in health sciences, with relevant applications expanding from disease outbreak prediction to medical imaging and patient communication to behavioral modification [9-14]. Over the past decade, an upsurge of the scientific literature adopting AI in health research has been witnessed [15,16]. These investigations applied a wide range of AI models: from *shallow* ML algorithms (eg, decision trees (DTs) and k-means clustering) and *deep* neural networks [17] to various data sources (eg, clinical and observational) and types (eg, tabular, text, and image) [18]. This boom in AI applications raises many questions [19-21]: How do AI-based approaches differ from conventional statistical analyses? Do AI techniques provide additional benefits or advantages over traditional methods? What are the typical AI applications and algorithms applied in obesity research? Is AI a buzzword that will eventually fall out of fashion, or will the upward trend of AI adoption to study obesity continue in the future?

Synthesizing and Disseminating AI Methodologies Adopted in Obesity Research

Three previous studies reviewed the applications of AI in weight loss interventions through diet and exercise [22-24]. They found preliminary but promising evidence regarding the effectiveness of AI-powered tools in decision support and digital health interventions [22-24]. However, to our knowledge, no study has been conducted to summarize AI algorithms, models, and methods applied to obesity research. This study remains the first methodological review on the applications of AI to measure, predict, and treat childhood and adult obesity. It serves 2 purposes: synthesizing and disseminating AI methodologies adopted in obesity research. First, we focused on summarizing and categorizing AI methodologies used in the obesity literature in the hope of identifying synergies, patterns, and trends to inform future scientific investigations. Second, we provided a high-level, beginner-friendly introduction to the core methodologies for interested readers, aiming to facilitate the dissemination and adoption of various AI techniques.

Methods

The scoping review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [25].

Study Selection Criteria

Studies that met all of the following criteria were included in the review: (1) study design: experimental or observational studies; (2) analytic approach: use of AI, including ML and DL (ie, deep neural networks), in measuring, predicting, or intervening obesity-related outcomes; (3) study participants: humans of all ages; (4) outcomes: obesity or body weight status (eg, BMI, body fat percentage [BFP], waist circumference [WC], and waist-to-hip ratio [WHR]); (5) article type: original, empirical, and peer-reviewed journal publications; (6) time window of search: from the inception of an electronic bibliographic database to January 1, 2022; and (7) language: articles written in English.

Studies that met any of the following criteria were excluded from the review: (1) studies focusing on outcomes other than obesity (eg, diet, physical activity, energy expenditure, and diabetes); (2) studies that used a rule-based (*hard-coded*) approach rather than example-based ML or DL; (3) articles not written in English; and (4) letters, editorials, study or review protocols, case reports, and review articles.

Search Strategy

A keyword search was performed in 2 electronic bibliographic databases: PubMed and Web of Science. The search algorithm included all possible combinations of keywords from the following two groups: (1) “artificial intelligence,” “computational intelligence,” “machine intelligence,” “computer reasoning,” “machine learning,” “deep learning,” “neural network,” “neural networks,” or “reinforcement learning” and (2) “obesity,” “obese,” “overweight,” “body mass index,” “BMI,” “adiposity,” “body fat,” “waist circumference,” “waist to hip,” or “waist - to - hip.” The Medical Subject Headings terms “artificial intelligence” and “obesity” were included in the PubMed search. [Multimedia Appendix 1](#) documents the search algorithm used in PubMed. Two coauthors of this review independently conducted title and abstract screening on the articles identified from the keyword search, retrieved potentially eligible articles, and evaluated their full texts. The interrater agreement between the 2 coauthors was assessed with Cohen kappa ($\kappa=0.80$). Discrepancies were resolved through discussion.

Data Extraction and Synthesis

A standardized data extraction form was used to collect the following methodological and outcome variables from each included study: authors; year of publication; country; data collection period; study design; sample size; training, validation, and test set size; sample characteristics; the proportion of female participants; age range; AI models used; input data source; input data format; input features; outcome data type; outcome measures; unit of analysis; main study findings; and implications for the effectiveness and usefulness of AI in measuring, predicting, or intervening obesity-related outcomes.

Methodological Review

We classified AI methodologies adopted by the included studies into 2 primary categories: ML and DL models. Among the ML models, methods were organized into 2 subcategories: unsupervised and supervised learning. Among the DL models, methods were classified into 3 subcategories: tabular data

modeling, computer vision (CV), and natural language processing (NLP). Rather than enumerating every single model performed by the included studies, which is unnecessary and unilluminating, we focused on the popular models used by multiple studies.

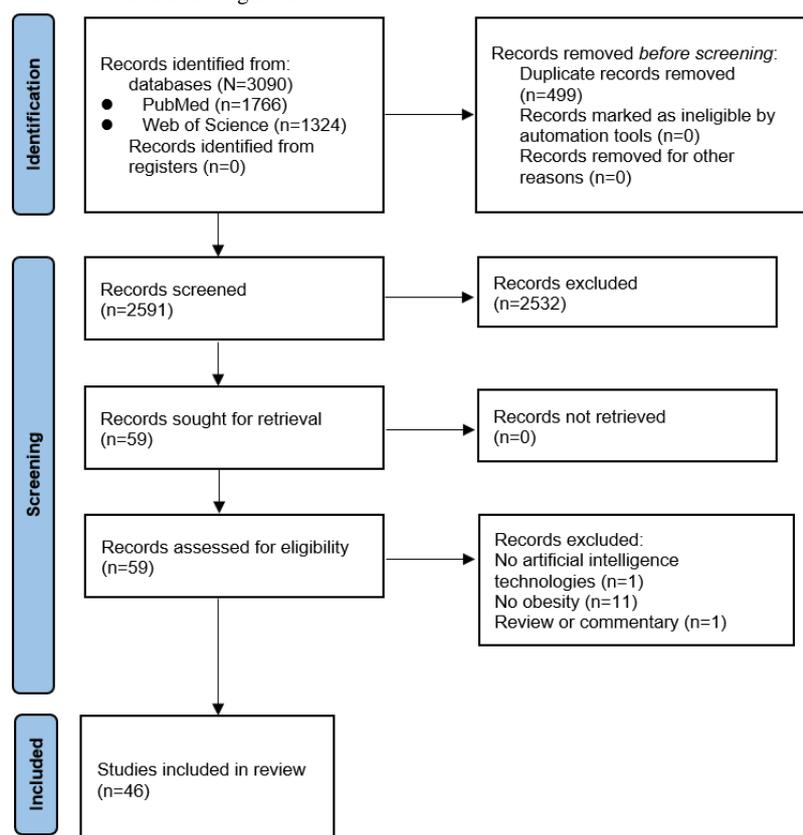
Results

Identification of Studies

Figure 1 shows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. We

identified a total of 3090 articles through the keyword search, and after removing 499 (16.15%) duplicates, 2591 (83.85%) unique articles underwent title and abstract screening. Of these 2591 articles, 2532 (97.72%) were excluded, and the full texts of the remaining 59 (2.28%) were reviewed against the study selection criteria. Of these 59 articles, 13 (22%) were excluded. The reasons for exclusion were as follows: no adoption of AI technologies (1/13, 8%), no obesity-related outcomes (11/13, 85%), and commentary rather than original empirical research (1/13, 8%). Therefore, of the 3090 articles identified initially through the keyword search, 46 (1.49%) were included in the review [26-71].

Figure 1. Identification of studies via databases and registers.



Study Characteristics

Table 1 summarizes the key characteristics of the 46 included studies. An increasing trend in relevant publications was observed. The earliest study included in the review was published in 1997; others were published in, or after, 2008; for example, 2% (1/46) each in 2008, 2012, and 2017; 4% (2/46) each in 2014 and 2016; 7% (3/46) each in 2009 and 2015; 9% (4/46) in 2018; 15% (7/46) in 2019; 20% (9/46) in 2020; and 26% (12/46) in 2021. Of the 46 studies, 16 (35%) were conducted in the United States [28,32,33,37,42,46,48,50-53,57,58,60,62,63]; 6 (13%) in China [39,40,45,56,64,65]; 3 (7%) each in the United Kingdom [27,68,69] and Korea [35,43,49]; 2 (4%) each in Italy [36,71], Turkey [41,70], Finland [44,59], Germany [54,55], and India [36,71]; and 1 (2%) each

in Saudi Arabia [26], Iran [67], Serbia [66], Portugal [61], Spain [47], Singapore [38], Australia [34], and Indonesia [29]. Of the 46 studies, 32 (70%) adopted a cross-sectional study design [26,27,29-32,37,39-42,46-50,52,55-58,60-63,65-71], 7 (15%) a prospective study design [28,33,38,43,45,54,59], 6 (13%) a retrospective study design [34-36,51,53,64], and 1 (2%) a cotwin control design [44]. Sample sizes varied substantially across the included studies, ranging from 20 to 5,265,265. Of the 46 studies, 7 (15%) had a sample size of between 20 and 82; 11 (24%) between 130 and 600; 19 (41%) between 1061 and 9524; 6 (13%) between 16,553 and 49,805; 2 (4%) between 244,053 and 618,898; and 1 (2%) study had a sample size of 5,265,265. Of the 46 studies, 23 (50%) focused on adults, 14 (30%) on children and adolescents, 1 (2%) on people of all ages, and the remaining 8 (17%) did not report the age range of participants.

Table 1. Characteristics of the studies included in the review.

Authors, year	Country	Data collection period	Study design	Sample size	Training set size	Validation set size; test set size	Sample characteristics	Female participants (%)	Age (years)	AI ^a model
Abdel-Aal and Mangoud [26], 1997	Saudi Arabia	1995	Cross-sectional	1100	800	N/A; 300	Patients	N/A ^b	≥20	NN ^c (AIM ^d abductive)
Positano et al [71], 2008	Italy	N/A	Cross-sectional	20	N/A	N/A	Participants with varying levels of obesity	N/A	Mean 52 (SD 16)	Fuzzy c-means
Ergün [70], 2009	Turkey	N/A	Cross-sectional	82	41	N/A; 41	Participants with different ranges of obesity	N/A	N/A	LR ^e , MLP ^f
Yang et al [69], 2009	United Kingdom	N/A	Cross-sectional	507	N/A	N/A	Patients	N/A	N/A	SVM ^g
Zhang et al [68], 2009	United Kingdom	1988 to 2003	Cross-sectional	16,553	11,091	N/A; 5462	Children	N/A	Birth to 3	NB ^h , SVM, DT ⁱ , NN
Heydari et al [67], 2012	Iran	2010	Cross-sectional	414	248	N/A; 104	Healthy military personnel	N/A	Mean 34.4 (SD 7.5)	NN, LR
Kupusinac et al [66], 2014	Serbia	N/A	Cross-sectional	2755	1929	413; 413	Adults	48.3	18 to 88	NN
Shao [65], 2014	China	N/A	Cross-sectional	248	174	N/A; 74	N/A	N/A	N/A	MR ^j , MARS ^k , SVM, NN
Chen et al [64], 2015	China	N/A	Retrospective	476	N/A	N/A	Participants with different ranges of obesity	62.4	22 to 82	NN (ELM ^l)
Dugan et al [63], 2015	United States	N/A	Cross-sectional	7519	6767	N/A; 752	Children	49	2 to 10	DT, RF ^m , NB, NN (BN ⁿ)
Nau et al [62], 2015	United States	2010	Cross-sectional	22,497	15,073	N/A; 7424	Children	N/A	10 to 18	RF
Almeida et al [61], 2016	Portugal	2009 to 2013	Cross-sectional	3084	1537	N/A; 664	School-age children	49.7	9	LR, NN
Lingren et al [60], 2016	United States	N/A	Cross-sectional	428	257	N/A; 86	Children	N/A	1 to 6	SVM, NB
Syednasrollah et al et al [59], 2017	Finland	1980 to 2012	Prospective	2262	1625	N/A; 637	Adults	N/A	≥18	GB ^o
Hinojosa et al [58], 2018	United States	2003 to 2007	Cross-sectional	5,265,265	N/A	N/A	School-age children: grades 5, 7, and 9	N/A	N/A	RF
Maharana and Nsoesie [57], 2018	United States	2017	Cross-sectional	1695	508	N/A; 339	Adults	N/A	≥18	NN (CNN ^p)
Wang et al [56], 2018	China	2014 to 2015	Cross-sectional	139	111	N/A; 28	Participants with different ranges of obesity	36.7	27 to 53	SVM, KNN ^q , DT, LR
Duran et al [55], 2018	Germany	1999 to 2004	Cross-sectional	1999	1333	N/A; 666	Children	42.8	8 to 19	NN

Authors, year	Country	Data collection period	Study design	Sample size	Training set size	Validation set size; test set size	Sample characteristics	Female participants (%)	Age (years)	AI ^a model
Gerl et al [54], 2019	Germany	2012; 1991 to 1994	Prospective	1061	796	206; 250	N/A	53.8	N/A	Cubist, LASSO ^f , PLS ^s , GB, RF, LM ^t
Hammond et al [53], 2019	United States	2008 to 2016	Retrospective	3449	482	N/A; 207	Children	49.2	4.5 to 5.5	LASSO, RF, GB
Hong et al [52], 2019	United States	2008	Cross-sectional	1237	1400	N/A; 600	Patients	N/A	≥18	LR, SVM, DT, RF
Ramyaa et al [51], 2019	United States	1993 to 1994	Retrospective	48,508	33,956	N/A; 14,552	Postmenopausal women	100	50 to 79	SVM, KNN, DT, PCA ^u , RF, NN
Scheinker et al [50], 2019	United States	2018	Cross-sectional	3138	N/A	N/A	Census population	49.9	All ages	LM, GB
Shin et al [49], 2019	Korea	N/A	Cross-sectional	163	143	N/A; 20	Amateur athletes	37.4	17 to 25	NN
Stephens et al [48], 2019	United States	N/A	Cross-sectional	23	N/A	N/A	Youth with obesity symptoms	57	Range 9.78-18.54	NN
Blanes-Selva et al [47], 2020	Spain	N/A	Cross-sectional	49,805	39,844	N/A; 9961	Patients	N/A	N/A	PU ^v learning
Dunstan et al [46], 2020	United States	2008	Cross-sectional	79	N/A	N/A	Adults	N/A	≥20	SVM, RF, GB
Fu et al [45], 2020	China	1999 to 2003	Prospective	2125	1143	381; 382	Children	40.6	4 to 7	GB
Kibble et al [44], 2020	Finland	N/A	Cotwin control	43	N/A	N/A	Young adult monozygotic twin pairs	53	22 to 36	GFA ^w
Park et al [43], 2020	Korea	N/A	Prospective	76	75	N/A; 1	Adolescents	6.8; N/A	Mean 11.94 (SD 3.13); mean 13.42 (SD 3.25)	LASSO
Phan et al [42], 2020	United States	2017 to 2018	Cross-sectional	18,700 images	14,960	N/A; 3740	Adolescents and adults	N/A	N/A	LM, NN (CNN)
Taghiyev et al [41], 2020	Turkey	2019	Cross-sectional	500	325	N/A; 175	Female patients	100	≥18	DT, LR
Xiao et al [40], 2020	China	2007 to 2010	Cross-sectional	9524	N/A	N/A	Residents	54	≥18	LR, NN (CNN)
Yao et al [39], 2020	China	N/A	Cross-sectional	67; 24	N/A	N/A	Smartphone users	N/A; 41.7	Mean 25.19; range 18-46	NN
Alkutbe et al [27], 2021	United Kingdom	2014; 2015 to 2016	Cross-sectional	1223	977	N/A; 246	Children	61.8	8 to 12	GB
Bhanu et al [38], 2021	Singapore	2003 to 2006	Prospective	130	104	N/A; 26	Older adults	69.5	Mean 67.85 (SD 7.90)	NN (U-Net)

Authors, year	Country	Data collection period	Study design	Sample size	Training set size	Validation set size; test set size	Sample characteristics	Female participants (%)	Age (years)	AI ^a model
Cheng et al [37], 2021	United States	2003 to 2004; 2005 to 2006	Cross-sectional	7162	N/A	N/A	Adults	48.6	20 to 85	NB, KNN, MEFC ^x , DT, NN (MLP)
Delnevo et al [36], 2021	Italy	N/A	Retrospective	221	176	N/A; 45	Participants with different ranges of obesity	N/A	N/A	GB, RF
Lee et al [35], 2021	Korea	2015 to 2020	Retrospective	3159	2370	N/A; 789	Obstetric patients and their newborns	100	20 to 44	LM, RF, NN
Lin et al [34], 2021	Australia	2010 to 2019	Retrospective	2495	882	N/A; 1613	Participants with different ranges of obesity	67.4	21 to 36	Two-step cluster analysis, k-means
Pang et al [33], 2021	United States	2009 to 2017	Prospective	27,203	21,762	N/A; 5441	Children	49.2	<2	DT, NB, LR, SVM, GB, NN
Park et al [32], 2021	United States	2014 to 2016	Cross-sectional	5000 tweets	4500	N/A; 500	Twitter users	60.7	Mean 51.91 (SD 17.20)	NB, SVM, NN (CNN, LSTM ^y)
Rashmi et al [31], 2021	India	2020	Cross-sectional	600 images	420	120; 60	Children	50	8 to 11	SVM, NB, RF
Snehalatha and Sangamithirai [30], 2021	India	N/A	Cross-sectional	2700 images	2000	500; 200	Adults	N/A	Mean 45 (SD 2.5)	NN (VGG, ResNet, DenseNet)
Thamrin et al [29], 2021	Indonesia	2018	Cross-sectional	618,898	557,008	N/A; 61,890	Adults	N/A	≥18	DT, NB, LR

Authors, year	Country	Data collection period	Study design	Sample size	Training set size	Validation set size; test set size	Sample characteristics	Female participants (%)	Age (years)	AI ^a model
Zare et al [28], 2021	United States	2003 to 2019	Prospective	244,053	162,702	N/A; 81,351	Children	49	5 to 6	DT, LR, RF, NN

^aAI: artificial intelligence.

^bN/A: not applicable.

^cNN: neural network.

^dAIM: abductory induction mechanism.

^eLR: logistic regression.

^fMLP: multilayer perceptron.

^gSVM: support vector machine.

^hNB: naïve Bayes.

ⁱDT: decision tree.

^jMR: multiple regression.

^kMARS: multivariate adaptive regression splines.

^lELM: extreme learning machine.

^mRF: random forest.

ⁿBN: BayesNet.

^oGB: gradient boosting.

^pCNN: convolutional neural network.

^qKNN: k-nearest neighbor.

^rLASSO: least absolute shrinkage and selection operator.

^sPLS: partial least squares.

^tLM: linear model.

^uPCA: principal component analysis.

^vPU: positive and unlabeled.

^wGFA: group factor analysis.

^xMEFC: multiobjective evolutionary fuzzy classifier.

^yLSTM: long short-term memory.

Data Sources and Outcome Measures

Table 2 summarizes the data sources and outcome measures of the studies included in the review. Input data were obtained from a variety of sources, including health surveys (eg, National Health and Nutrition Examination Survey), electronic health records, magnetic resonance imaging (MRI) scans, social media data (eg, tweets), and geographically aggregated data sets (eg,

InfoUSA and Dun & Bradstreet). Of the 46 studies, 34 (74%) analyzed tabular data (eg, spreadsheet data) [26-29,33-37,39,41,44-47,49-51,53-56,58-68,70], 8 (17%) analyzed digital image data [30,31,38,40,42,43,57,71], and 4 (9%) analyzed text data [32,48,52,69]. Obesity-related measures used across the studies included anthropometrics (eg, body weight, BMI, BFP, WC, and WHR) and biomarkers.

Table 2. Data sources and measures of outcomes in the studies included in the review.

Authors, year	Input data source	Input data format	Input features (independent variables)	Outcome data type	Outcome measures	Unit of analysis
Abdel-Aal and Mangoud [26], 1997	Medical survey data	Tabular	13 health parameters	Continuous	WHR ^a	Individual
Positano et al [71], 2008	MRI ^b	Image	Subcutaneous adipose tissue and visceral adipose tissue	Binary	Abdominal adipose tissue distribution	Individual
Ergün [70], 2009	Obtained from participants	Tabular	24 obesity parameters	Binary	Classification of obesity	Individual
Yang et al [69], 2009	Clinical data	Text	Clinical discharge summaries	Binary	Obesity status	Individual
Zhang et al [68], 2009	Objective measure	Tabular	Data recorded regarding the weight of the child during the first 2 years of the child's life	Binary	Obesity	Individual
Heydari et al [67], 2012	Questionnaire and objective measure	Tabular	Age, systole, diastole, weight, height, BMI, WC ^c , HC ^d , and triceps skinfold and abdominal thicknesses	Binary	Obesity	Individual
Kupusinac et al [66], 2014	Objective measure	Tabular	Gender, age, and BMI	Continuous	BFP ^e	Individual
Shao [65], 2014	Objective measure	Tabular	13 body circumference measurements	Continuous	BFP	Individual
Chen et al [64], 2015	Objective measure	Tabular	18 blood indexes and 16 biochemical indexes	Continuous	Overweight	Individual
Dugan et al [63], 2015	Questionnaire and objective measure	Tabular	167 clinical data attributes	Continuous	Obesity	Individual
Nau et al [62], 2015	Two secondary data sources (InfoUSA and Dun & Bradstreet)	Tabular	44 community characteristics	Binary	Obesogenic and obesoprotective environments	Community
Almeida et al [61], 2016	Objective measure	Tabular	Age, sex, BMI z score, and calf circumference	Continuous	BFP	Individual
Lingren et al [60], 2016	EHR ^f	Tabular	EHR data	Binary	Obesity	Individual
Seyednasrollah et al [59], 2017	Objective measure	Tabular	Clinical factors and genetic risk factors	Binary	Obesity	Individual
Hinojosa et al [58], 2018	Objective measure	Tabular	School environment	Binary	Obesity	School
Maharana and Nsoesie [57], 2018	Objective measure	Image	Built environment	Continuous	Prevalence of obesity	Census tract
Wang et al [56], 2018	Objective measure	Tabular	Single-nucleotide polymorphisms	Binary	Obesity risk	Individual
Duran et al [55], 2018	NHANES ^g	Tabular	Age, height, weight, and WC	Binary	Excess body fat	Individual
Gerl et al [54], 2019	Objective measure	Tabular	Human plasma lipidomes	Binary and continuous	Obesity: BMI, WC, WHR, and BFP	Individual
Hammond et al [53], 2019	EHR and publicly available data	Tabular	EHR data	Binary and continuous	Obesity status	Individual
Hong et al [52], 2019	EHR	Text	Discharge summaries	Binary	Identification of obesity	Individual
Ramyaa et al [51], 2019	Questionnaire	Tabular	Energy balance components	Binary and continuous	Energy stores: body weight	Individual

Authors, year	Input data source	Input data format	Input features (independent variables)	Outcome data type	Outcome measures	Unit of analysis
Scheinker et al [50], 2019	2018 Robert Wood Johnson Foundation County Health Rankings	Tabular	Demographic factors, socioeconomic factors, health care factors, and environmental factors	Continuous	Obesity prevalence	County
Shin et al [49], 2019	Objective measure	Tabular	Upper body impedance and lower body anthropometric data	Continuous	BFP	Individual
Stephens et al [48], 2019	From recorded dialogue	Text	Dialogue	Binary	Weight management program	Individual
Blanes-Selva et al [47], 2020	EHR of HULAFE ^h	Tabular	32 variables	Binary	Identification of obesity	Individual
Dunstan et al [46], 2020	Euromonitor data set	Tabular	National sales of a small subset of food and beverage categories	Continuous	Nationwide obesity prevalence	Country
Fu et al [45], 2020	Clinical data	Tabular	Demographic characteristics, maternal anthropometrics, perinatal clinical history, laboratory tests, and postnatal feeding practices	Binary	Obesity	Individual
Kibble et al [44], 2020	Clinical data	Tabular	42 clinical variables	Binary	Mechanisms of obesity	Individual
Park et al [43], 2020	Openly accessible database	Image	Neuroimaging biomarkers	Continuous	BMI	Individual
Phan et al [42], 2020	Objective measure	Image	Neighborhood built environment characteristics	Binary, continuous	Obesity	State
Taghiyev et al [41], 2020	EHR	Tabular	Results of blood tests	Binary	Obesity	Individual
Xiao et al [40], 2020	Objective measure	Image	Vertical greenness level	Binary	Obesity	Individual
Yao et al [39], 2020	Objective measure	Tabular	Characteristics of body movement captured by smartphone's built-in motion sensors	Continuous	BMI	Individual
Alkutbe et al [27], 2021	Self-reported and objective measures	Tabular	Weight, height, age, and gender	Binary and continuous	BFP	Individual
Bhanu et al [38], 2021	MRI	Image	SAT ⁱ and VAT ^j	Binary	Abdominal fat	Individual
Cheng et al [37], 2021	Objective measure	Tabular	Physical activity	Binary	Obesity	Individual
Delnevo et al [36], 2021	Questionnaire	Tabular	Positive and negative psychological variables	Binary and continuous	BMI values and BMI status	Individual
Lee et al [35], 2021	Objective measure	Tabular	64 independent variables: nationwide multicenter ultrasound data and maternal and delivery information	Continuous	BMI	Individual
Lin et al [34], 2021	Objective measure	Tabular	Key clinical variables	Binary	Obesity classification criterion	Individual
Pang et al [33], 2021	EHR data from pediatric big data repository	Tabular	Demographic variables and 54 clinical variables	Binary	Obesity	Individual
Park et al [32], 2021	Corpus of geotagged tweets	Text	Tweets	Binary and continuous	BMI and obesity	Individual
Rashmi et al [31], 2021	Objective measure	Image	600 thermograms	Binary	Obesity	Individual
Snehalatha and Sangamithirai [30], 2021	Objective measure	Image	Thermal imaging	Binary	Diagnosis of obesity	Individual

Authors, year	Input data source	Input data format	Input features (independent variables)	Outcome data type	Outcome measures	Unit of analysis
Thamrin et al [29], 2021	Publicly available health data	Tabular	Risk factors for obesity	Binary	Obesity	Individual
Zare et al [28], 2021	BMI panel data set	Tabular	Kindergarten BMI z score	Binary	Obesity by grade 4	Individual

^aWHR: waist-hip ratio.

^bMRI: magnetic resonance imaging.

^cWC: waist circumference.

^dHC: hip circumference.

^eBFP: body fat percentage.

^fEHR: electronic health record.

^gNHANES: National Health and Nutrition Examination Survey.

^hHULAFE: Hospital Universitari i Politècnic La Fe.

ⁱSAT: subcutaneous adipose tissue.

^jVAT: visceral adipose tissue.

Main Findings

Table 3 summarizes the estimated effects and main findings of the studies included in the review. Four key findings have emerged.

First, the studies found that ML or DL models were generally effective in detecting clinically meaningful patterns of obesity or relationships between covariates and weight outcomes; for example, ML and DL models were found useful in classifying obesity severity [30,47,52], identifying anthropometric [34] and genetic characteristics of obesity [56], and predicting obesity onset in children [28,53,63]. ML algorithms (eg, random forest [RF] and conditional RF) revealed meaningful relationships between school and neighborhood environments and overweight and obesity [45,58,62]. DL algorithms (eg, convolutional neural network [CNN]) effectively extracted built environment features from satellite images to assess their associations with the local obesity rate [57].

Second, most (18/22, 82%) of the studies comparing AI models with conventional statistical methods reported that the AI models achieved higher prediction accuracy on test data, whereas others (4/22, 18%) found similar model performances; for example, ML and DL models were found to explain a larger proportion of variations in county-level obesity prevalence than

conventional statistical approaches [50]. ML models showed flexibility in handling various variable types [36,41] and large-scale data sets [32] and producing robust, generalizable inferences [41,54,64,65] with higher prediction accuracy [61,66]. By contrast, Cheng et al [37] reported that ML algorithms and conventional statistical approaches had similar performance.

Third, some (5/46, 11%) of the studies comparing the performances of different AI models yielded mixed results, reflecting the interdependence between model and data or task; for example, logistic regressions were reported to achieve higher prediction accuracy than DTs, naïve Bayes (NB) [29], and DL [35]. By contrast, Heydari et al [67] found that logistic regressions and DL models performed equally well in solving classification problems. Zhang et al [68] and Ergün [70] reported that data mining and DL techniques outperformed logistic regressions in classification accuracy.

Fourth, newer studies increasingly adopted state-of-the-art DL models to address CV and NLP tasks; for example, chatbots built on NLP models were used to support pediatric obesity treatment [48]. CNN-based CV models were used to construct indicators for the built environment using images from Google Street View [42]. DL-based tools were used to efficiently visualize and analyze abdominal visceral adipose tissue and subcutaneous adipose tissue [38].

Table 3. Estimated effects and main findings of the studies included in the review.

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
Abdel-Aal and Mangoud [26], 1997	<ul style="list-style-type: none"> Models for WHR^b as a continuous variable predict the actual values within an error rate of 7.5% at the 90% confidence limits. Categorical models predict the correct logical value of WHR with an error in only 2 of the 300 evaluation cases. Analytical relationships derived from simple categorical models explain global observations on the total survey population to an accuracy rate as high as 99%. Simple continuous models represented as analytical functions highlight global relationships and trends. There is a strong correlation between WHR and diastolic blood pressure, cholesterol level, and family history of obesity. 	<ul style="list-style-type: none"> Compared with other statistical and neural network approaches, AIM^c abductive networks provide a faster and more automated model synthesis.
Positano et al [71], 2008	<ul style="list-style-type: none"> CV^d values in VAT^e, SAT^f, and VAT/SAT ratio assessment by the standard algorithm without image inhomogeneities correction were 10.7%, 11.9%, and 17.3%, respectively. Correlation coefficients were $r=0.97$, $r=0.93$, and $r=0.95$, respectively (all $P<.001$). When correction for field inhomogeneities was applied, VAT, SAT, and VAT/SAT ratio CVs became 9.8%, 6.7%, and 13.1%, respectively. Correlation coefficients became $r=0.97$, $P<.001$ for VAT; $r=0.99$, $P<.001$ for SAT; and $r=0.97$, $P<.001$ for VAT/SAT ratio. 	<ul style="list-style-type: none"> The CV between manual and unsupervised analyses was significantly improved by inhomogeneities correction in SAT evaluation. Systematic underestimation of SAT was also corrected. A less critical performance improvement was found in VAT measurement. The compensation of signal inhomogeneities improves the effectiveness of the unsupervised assessment of abdominal fat. Correction of intensity distortions is necessary for SAT evaluation but less significant in VAT measurement.
Ergün [70], 2009	<ul style="list-style-type: none"> The classification rate of neural networks in obesity is 90.2%, and the classification rate of logistic regression in obesity is 87.8%. After these classifications, in obesity, the BMI is more affected than the divergent arteries. 	<ul style="list-style-type: none"> The classifying performance of a neural network is better than that of logistic regression.
Yang et al [69], 2009	<ul style="list-style-type: none"> The implemented method achieved the macroaveraged F-measure of 81% for the textual task and 63% for the intuitive task. The microaveraged F-measure showed an average accuracy of 97% for textual annotations and 96% for intuitive annotations. 	<ul style="list-style-type: none"> Text mining may provide an accurate and efficient prediction of disease statuses from clinical discharge summaries.
Zhang et al [68], 2009	<ul style="list-style-type: none"> Prediction at 8 months' accuracy is improved very slightly, in this case by using neural networks, whereas for prediction at 2 years, the obtained accuracy is enhanced by >10%, in this case by using Bayesian methods. 	<ul style="list-style-type: none"> SVM^g and Bayesian algorithms seem to be the best algorithms for predicting overweight and obesity from the Wirral database. The incorporation of nonlinear interactions could be important in childhood obesity prediction. Data mining techniques are becoming sufficiently well established to offer the medical research community a valid alternative to logistic regression.
Heydari et al [67], 2012	<ul style="list-style-type: none"> Regarding logistic regression and neural networks, the respective values were 80.2% and 81.2% for correct classification 80.2% and 79.7% for sensitivity, and 81.9% and 83.7% for specificity; the values for the area under the receiver operating characteristic curve were 0.888 and 0.884, respectively, and the values for the kappa statistic were 0.600 and 0.629, respectively. Abdominal thickness, weight, BMI, and HC^h were significantly associated with obesity. 	<ul style="list-style-type: none"> Neural networks and logistic regression were good classifiers for obesity detection but were not significantly different with regard to classification.
Kupusinac et al [66], 2014	<ul style="list-style-type: none"> The predictive accuracy of an ANNⁱ solution is 80.43%. ANN showed higher predictive accuracy ranging from +1.23% to +3.12%. 	<ul style="list-style-type: none"> An ANN is a new approach to predicting BFP^j with the same complexity and costs but with higher predictive accuracy.

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
Shao [65], 2014	<ul style="list-style-type: none"> Although the 13 body circumference measurements are involved in the real data set, the proposed models can provide better predictions with fewer body circumference measurements. It is much more convenient to predict BFP with fewer body circumference measurements for most people. 	<ul style="list-style-type: none"> Compared with traditional single-stage approaches, the proposed hybrid models—multiple regression, ANN, multivariate adaptive regression splines, and support vector regression techniques—can effectively predict BFP.
Chen et al [64], 2015	<ul style="list-style-type: none"> The most important correlated indexes are creatinine, hemoglobin, hematocrit, uric acid, red blood cells, high-density lipoprotein, alanine transaminase, triglyceride, and γ-glutamyl transpeptidase. 	<ul style="list-style-type: none"> The ELM^k performs much more efficiently than the SVM and BPNN^l and with higher recognition rates. The proposed ELM-based approach for overweight detection in biomedical applications holds promise as a new, accurate method for identifying participants' overweight status. It provides a viable alternative to traditional overweight modeling tools by offering excellent predictive ability.
Dugan et al [63], 2015	<ul style="list-style-type: none"> The ID3^m model trained on the CHICAⁿ data set demonstrated the best overall performance with an accuracy of 85% and sensitivity of 89%. In addition, the ID3 model had a positive predictive value of 84% and a negative predictive value of 88%. Being overweight between the ages of 12 and 24 months is a key risk factor for obesity after the second birthday. Furthermore, it is more of a risk factor if the child was not overweight before 12 months. 	<ul style="list-style-type: none"> Data from a production clinical decision support system can be used to build an accurate ML^o model to predict obesity in children after the age of 2 years.
Nau et al [62], 2015	<ul style="list-style-type: none"> After examining 44 community characteristics, the researchers identified 13 features of the social, food, and physical activity environment that, in combination, correctly classified 67% of communities as obesoprotective or obesogenic using the mean BMI z score as a surrogate. Social environment characteristics emerged as the most critical classifiers and might leverage intervention. 	<ul style="list-style-type: none"> CRF^p allows consideration of the neighborhood as a system of risk factors.
Almeida et al [61], 2016	<ul style="list-style-type: none"> All BFP-grade predictive models presented a good global accuracy ($\geq 91.3\%$) for obesity discrimination. Both overfat and obese as well as obese prediction models showed, respectively, good sensitivity (78.6% and 71%), specificity (98% and 99.2%), and reliability for positive or negative test results ($\geq 82\%$ and $\geq 96\%$). For boys, the order of parameters, by relative weight in the predictive model, was BMI z score, height, WHtR^q squared variable (_Q), age, weight, CC^r_Q, and HC^s_Q (adjusted R²=0.847 and RMSE^t=2.852); for girls, it was BMI z score, WHtR_Q, height, age, HC_Q, and CC_Q (adjusted R²=0.872 and RMSE=2.171). 	<ul style="list-style-type: none"> BFP can be graded and predicted with relative accuracy from anthropometric measurements (excluding skinfold thickness). Fitness and cross-validation results showed that the multivariable regression model performed better in this population than in some previously published models.
Lingren et al [60], 2016	<ul style="list-style-type: none"> Overall, the rule-based algorithm performed the best: 0.895 (CCHMC^u) and 0.770 (BCH^v). 	<ul style="list-style-type: none"> The rule-based exclusion algorithm performed better than the ML algorithm. The best feature set for ML used Unified Medical Language System concept unique identifiers; International Classification of Diseases, Ninth Revision, codes; and R^xNorm codes.
Syednasrollah et al [59], 2017		<ul style="list-style-type: none"> WGRS19 improves the prediction of adulthood obesity. The model helps screen children with a high risk of developing obesity. Predictive accuracy is highest among young children (aged 3-6 years), whereas among older children (aged 9-18 years), the risk can be identified using childhood clinical factors.

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
	<ul style="list-style-type: none"> Replication in the BHSw confirmed the researchers' findings that WGRSx19 and WGRS97 are associated with BMI. WGRS19 improved the accuracy of predicting adulthood obesity in the training data (area under the curve=0.787 vs area under the curve=0.744; $P<.001$) and validation data (area under the curve=0.769 vs area under the curve=0.747; $P=.03$). WGRS97 improved the accuracy in the training data (area under the curve=0.782 vs area under the curve=0.744; $P<.001$) but not in the validation data (area under the curve=0.749 vs area under the curve=0.747; $P=.79$). Higher WGRS19 is associated with a higher BMI at 9 years and WGRS97 at 6 years. 	
Hinojosa et al [58], 2018	<ul style="list-style-type: none"> Violent crime, English learners, socioeconomic disadvantage, fewer physical education and fully credentialed teachers, and diversity index were positively associated with obesity. By contrast, the academic performance index, physical education participation, mean educational attainment, and per capita income were negatively associated with obesity. The most highly ranked built or physical environment variables were distance to the nearest highway and green spaces, 10th and 11th most important, respectively. 	<ul style="list-style-type: none"> An RFy algorithm effectively identifies the relative importance of school environment attributes.
Maharana and Nsoesie [57], 2018	<ul style="list-style-type: none"> Features of the built environment explained 64.8% (RMSE=4.3) of the variation in obesity prevalence across all US census tracts. Individually, the variation explained was 55.8% (RMSE=3.2) for Seattle, Washington (213 census tracts); 56.1% (RMSE=4.2) for Los Angeles, California (993 census tracts); 73.3% (RMSE=4.5) for Memphis, Tennessee (178 census tracts); and 61.5% (RMSE=3.5) for San Antonio, Texas (311 census tracts). 	<ul style="list-style-type: none"> CNN^z can be used to automate the extraction of features of the built environment from satellite images for studying health indicators. Understanding the association between specific features of the built environment and obesity prevalence can lead to structural changes that could encourage physical activity and decrease obesity prevalence.
Wang et al [56], 2018	<ul style="list-style-type: none"> The SVM model significantly outperformed other classifiers based on the same training features. The SVM model exhibits 70.77% accuracy, 80.09% sensitivity, and 63.02% specificity. The selected SNPs^{aa} were effective in the detection of obesity risk. 	<ul style="list-style-type: none"> The ML-based method provides a feasible means for conducting preliminary analyses of genetic characteristics of obesity.
Duran et al [55], 2018	<ul style="list-style-type: none"> In female participants, the sensitivity of the BMI, WC^{bb}, and ANN approaches to predict excess body fat was 0.751 (95% CI 0.730 - 0.771), 0.523 (95% CI 0.487 - 0.559), and 0.782 (95% CI 0.754 - 0.810), respectively. In male participants, the sensitivity of the BMI, WC, and ANN approaches to predict excess body fat was 0.721 (95% CI 0.699 - 0.743), 0.572 (95% CI 0.549 - 0.594), and 0.795 (95% CI 0.768 - 0.821). 	<ul style="list-style-type: none"> The diagnostic performance in identifying excess body fat was better in male participants when an ANN approach was used than when BMI and WC z scores were applied. The ANN and BMI z scores performed comparably and significantly better, respectively, than WC z scores in female participants.
Gerl et al [54], 2019	<ul style="list-style-type: none"> The lipidome, based on a LASSO^{cc} model, predicted BFP the best ($R^2=0.73$). In this model, the strongest positive predictor and strongest negative predictor were sphingomyelin molecules, which differ by only 1 double bond, implying the involvement of an unknown desaturase in obesity-related aberrations of lipid metabolism. The regression was used to probe the clinically relevant information in the plasma lipidome and found that the plasma lipidome also includes information on body fat distribution because WHR ($R^2=0.65$) was predicted more accurately than BMI ($R^2=0.47$). 	<ul style="list-style-type: none"> ML can model and validate obesity estimates better than classical clinical parameters such as total triglycerides and cholesterol.
Hammond et al [53], 2019	<ul style="list-style-type: none"> LASSO regression predicted obesity with an area under the receiver operating characteristic curve of 81.7% for girls and 76.1% for boys. In each of the separate models for boys and girls, the researchers found that the weight-for-length z score, BMI between 19 and 24 months, and the last BMI measure recorded before the age of 2 years were the most important features for prediction. 	<ul style="list-style-type: none"> Comparable to cohort-based studies, EHR^{dd} data with area under the receiver operating characteristic curve values could be used to predict obesity at the age of 5 years, reducing the need for investment in additional data collection.
Hong et al [52], 2019		

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
	<ul style="list-style-type: none"> As the results of the 4 ML classifiers showed, the RF algorithm performed the best with micro F1-score 0.9466 and macro F1-score 0.7887 and micro F1-score 0.9536 and macro F1-score 0.6524 for intuitive classification (reflecting medical professionals' judgments) and textual classification (reflecting the decisions based on explicitly reported information of diseases), respectively. The MIMIC^{ee}-III obesity data set was successfully integrated for prediction with minimal configuration of the NLP^{ff}2FHIR^{gg} pipeline and ML models. 	<ul style="list-style-type: none"> The FHIR-based EHR phenotyping approach could effectively identify the obesity status and multiple comorbidities using semistructured discharge summaries.
Ramyaa et al [51], 2019	<ul style="list-style-type: none"> SVM, neural network, and KNN^{hh} algorithms performed modestly for the numerical predictions, with mean approximate errors of 6.70 kg, 6.98 kg, and 6.90 kg, respectively. K-means cluster analysis improved prediction using numerical data and identified 10 clusters suggestive of phenotypes, with a minimum mean approximate error of approximately 1.1 kg. A classifier was used to phenotype participants into the identified clusters, with mean approximate errors of <5 kg for 15% of the test set (approximately, n=2000). SVM performed the best (54.5% accuracy), followed closely by the bagged tree ensemble and KNN algorithms. 	<ul style="list-style-type: none"> SVM regression was the best-suited predictive and inferential tool for this task, closely followed by neural network and KNN algorithms. Although the overall data model showed a good fit and predictive ability, clustering produced relatively superior fit statistics.
Scheinker et al [50], 2019	<ul style="list-style-type: none"> Multivariate linear regression and gradient boosting machine regression (the best-performing ML model) of obesity prevalence using all county-level demographic, socioeconomic, health care, and environmental factors had R² values of 0.58 and 0.66, respectively (P<.001). 	<ul style="list-style-type: none"> ML may be used to explain more variation in county-level obesity prevalence than traditional epidemiologic models. The top-performing ML model explained two-thirds of the variation in county-level obesity prevalence, significantly more than conventional multivariate linear models.
Shin et al [49], 2019	<ul style="list-style-type: none"> The performance of the proposed system was compared with those of 2 commercial systems that were designed to measure body composition using either a whole body or upper body impedance value. The results showed that the correlation coefficient (R²) value was improved by approximately 9%, and the SE of the estimate was reduced by 28%. 	<ul style="list-style-type: none"> The test results validated that the inclusion of anthropometric data helped to improve accuracy, primarily when a DLⁱⁱ approach was used to predict the regression values.
Stephens et al [48], 2019	<ul style="list-style-type: none"> Adolescent patients reported experiencing positive progress toward their goals 81% of the time. The 4123 messages exchanged and patients' reported usefulness ratings (96% of the time) illustrate that adolescents engaged with the chatbot and viewed it as helpful. 	<ul style="list-style-type: none"> An AI chatbot is feasible as an adjunct to treatment. The feasibility and benefit of support through AI, specifically in a pediatric setting, could be scaled to serve larger groups of patients.
Blanes-Selva et al [47], 2020	<ul style="list-style-type: none"> The PU^{jj} learning algorithm presented a high sensitivity (98%) and predicted that approximately 18% of the patients without a diagnosis were obese. 	<ul style="list-style-type: none"> The implementation of the PU learning methodology in identifying obesity produced results that were satisfactory, providing high sensitivity, and consistent with the World Health Organization's obesity report.
Dunstan et al [46], 2020	<ul style="list-style-type: none"> Using only 5 categories, RF could predict obesity prevalence with absolute error <10% for approximately 60% of the countries considered and absolute error <20% for 87%. The most relevant food category with regard to predicting obesity consists of baked goods and flours, followed by cheese and carbonated drinks. 	<ul style="list-style-type: none"> RF shows the best performance for predicting obesity from food, followed closely by XGB^{kk}.
Fu et al [45], 2020		

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
	<ul style="list-style-type: none"> The 2 most important features—trajectory of infant BMI z score change and maternal BMI at enrollment—were identified from the ML algorithm. The aforementioned features showed similar predictive capacity compared with all features (area under the curve=0.68 vs 0.68; $P=.83$; DeLong test). The sensitivity analyses identified the same 2 features (ie, trajectory of infant BMI z score change and maternal BMI at enrollment), and the ranking of these features' Shapley additive explanations value was unchanged. In the independent test cohort, the area under the curve for childhood overweight and obesity classification using the aforementioned 2 features was 0.71 (95% CI 0.66 to 0.76), which was comparable to that based on all features (0.72, 95% CI 0.67 to 0.76). 	<ul style="list-style-type: none"> An ML algorithm is applied to identify risk factors contributing to childhood overweight or obesity based on a large longitudinal study and addresses the relationships between all collected features and outcomes without any assumption. A novel unified framework, Shapley additive explanations, is used to interpret predictions, and the identified predictive factors are robust.
Kibble et al [44], 2020	<ul style="list-style-type: none"> New potential links between cytokines and weight gain are identified, as well as associations among dietary, inflammatory, and epigenetic factors. 	<ul style="list-style-type: none"> An integrative ML method called group factor analysis was used to identify the links between multimolecular-level interactions and the development of obesity.
Park et al [43], 2020	<ul style="list-style-type: none"> The actual and predicted ΔBMI showed a significant intraclass correlation value with a low RMSE, and classification between people with increased BMI and those with nonincreased BMI resulted in a high area under the receiver operating characteristic curve value using only the degree centrality values obtained at the baseline visit. 	<ul style="list-style-type: none"> The constructed model using functional connectivity of the selected regions provides robust neuroimaging biomarkers for predicting BMI progression.
Phan et al [42], 2020	<ul style="list-style-type: none"> A DNN^{ll} was used for neighborhood indicator recognition and achieved high accuracies (85%-93%) for the separate recognition tasks. 	<ul style="list-style-type: none"> DL techniques were used to create indicators for neighborhood-built environment characteristics.
Taghiyev [41], 2020	<ul style="list-style-type: none"> The proposed hybrid system demonstrated 91.4% accuracy, which is higher than that of other classifiers (ie, 4.6% higher than the performance of logistic regression and 2.3% higher than the performance of DT^{mm}). 	<ul style="list-style-type: none"> The proposed hybrid system provides a more accurate classification of patients with obesity and a practical approach to estimating the factors affecting obesity.
Xiao et al [40], 2020	<ul style="list-style-type: none"> All aspects of horizontal greenery, vertical greenery, and proximity of green levels affected body weight; however, only the VGIⁿⁿ consistently had an adverse effect on weight and obesity. 	<ul style="list-style-type: none"> The VGI of the DL approach using Baidu Street View images could effectively capture the eye-level greenness in high-density-population areas. Thus, VGI can be used to effectively promote walking and other physical activities to prevent obesity.
Yao et al [39], 2020	<ul style="list-style-type: none"> Jogging may be a more suitable activity of daily living for BMI prediction than walking and walking up stairs. 	<ul style="list-style-type: none"> The proposed DL model with the motion entropy-based filtering strategy outperforms the baseline approaches significantly.
Alkutbe et al [27], 2021	<ul style="list-style-type: none"> For the gradient boosting models, the predicted fat percentage values were more aligned with the actual value than those in regression models. Gradient boosting achieved better performance than the regression equation because it combined multiple simple models into a single composite model to take advantage of this weak classifier. The developed predictive model archived RMSE values of 3.12 for girls and 2.48 for boys. 	<ul style="list-style-type: none"> ML models and newly developed centile charts could be valuable tools for estimating and classifying BFP.
Bhanu et al [38], 2021	<ul style="list-style-type: none"> The accuracy of segmentation was superficial SAT: 0.92, deep SAT: 0.88, and VAT: 0.9. The average Hausdorff distance was <5 mm. Automated segmentation significantly correlated $R^2>0.99$ ($P<.001$) with ground truth for all 3-fat compartments. Predicted volumes were within 1.96 SD from Bland-Altman analysis. 	<ul style="list-style-type: none"> DL-based, comprehensive superficial SAT, deep SAT, and VAT analysis tools showed high accuracy and reproducibility and provided a comprehensive fat compartment composition analysis and visualization in <10 seconds.
Cheng et al [37], 2021		

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
	<ul style="list-style-type: none"> Physical activity was an important factor in predicting weight status, with gender, age, and race or ethnicity being less important factors associated with weight outcomes. The durations of vigorous-intensity activity in 1 week and moderate-intensity activity in 1 week were essential attributes. 	<ul style="list-style-type: none"> With physical activity and basic demographic information of all methods analyzed, the random subspace classifier algorithm achieved the highest overall accuracy and area under the receiver operating characteristic curve value. In general, most algorithms showed similar performance. Logistic regression was middle ranking in terms of overall accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve value among all methods.
Delnevo et al [36], 2021	<ul style="list-style-type: none"> The psychological variables in use allow one to predict both BMI values (with a mean absolute error of 5.27-5.50) and BMI status with an accuracy of >80% (metric: F1-score). 	<ul style="list-style-type: none"> Certain psychological variables such as depression are highly predictive of BMI. ML has several advantages over traditional statistics and can be used to compare the impact of many variables on predicting a chosen outcome and can handle various types of variables.
Lee et al [35], 2021	<ul style="list-style-type: none"> For predicting a newborn's BMI, linear regression (2.0744) and RF (2.1610) were better than ANN with 1, 2, and 3 hidden layers (150.7100, 154.7198, and 152.5843, respectively) in the mean squared error. On the basis of variable importance from the RF, the major predictors of a newborn's BMI were the first abdominal circumference value and estimated fetal weight in week 36 or later, gestational age at delivery, the first abdominal circumference value during week 21 to week 35, maternal BMI at delivery, maternal weight at delivery, and the first biparietal diameter value in week 36 or later. 	<ul style="list-style-type: none"> ML approaches based on ultrasound measures would be a useful noninvasive tool for predicting a newborn's BMI. Linear regression and RF were better models than ANNs for predicting a newborn's BMI.
Lin et al [34], 2021	<ul style="list-style-type: none"> ML revealed the following 4 stable metabolically distinct obesity clusters in each cohort: <ul style="list-style-type: none"> Metabolic healthy obesity (44% of the patients) was characterized by a relatively healthy metabolic status with the lowest incidents of comorbidity. Hypermetabolic obesity-hyperuricemia (33% of the patients) was characterized by extremely high uric acid and an increased incidence of hyperuricemia (adjusted odds ratio 73.67 to metabolic healthy obesity, 95% CI 35.46-153.06). Hypermetabolic obesity-hyperinsulinemia (8% of the patients) was distinguished by overcompensated insulin secretion and an increased incidence of polycystic ovary syndrome (adjusted odds ratio 14.44 to metabolic healthy obesity, 95% CI 1.75-118.99). Hypometabolic obesity (15% of the patients) was characterized by extremely high glucose levels, decompensated insulin secretion, and the worst glucolipid metabolism (diabetes: adjusted odds ratio 105.85 to metabolic healthy obesity, 95% CI 42.00-266.74; metabolic syndrome: adjusted odds ratio 13.50 to metabolic healthy obesity, 95% CI 7.34-24.83). The assignment of patients in the verification cohorts to the main model showed a mean accuracy of 0.941 in all clusters. 	<ul style="list-style-type: none"> ML automatically identified 4 subtypes of obesity in clinical characteristics in 4 independent patient cohorts. This proof-of-concept study provided evidence that a precise diagnosis of obesity can potentially guide therapeutic planning and decisions for different subtypes of obesity.
Pang et al [33], 2021	<ul style="list-style-type: none"> XGB yielded a mean area under the curve value of 0.81 (SD 0.001), which outperformed all other models. It also achieved a statistically significant better performance than all other models on standard classifier metrics (sensitivity fixed at 80%): precision, mean 30.9% (SD 0.22%); F1-score, mean 44.6% (SD 0.26%); accuracy, mean 66.14% (SD 0.41%); and specificity, mean 63.27% (SD 0.41%). 	<ul style="list-style-type: none"> The presented ML model development workflow can be adapted to various EHR-based studies and is valuable for developing other clinical prediction models.
Park et al [32], 2021		

Authors, year	Estimated effects of AI ^a technologies on obesity prevention or treatment	Main findings
	<ul style="list-style-type: none"> ML algorithms were used to determine the stances of tweets on Black Lives Matter. ML models showed better performance than lexicon-based sentiment analysis (accuracy: 61%). The NB^{oo} model had an overall accuracy of 85%, slightly higher than that of the CNN model (83.8%); both had higher accuracy than the other models. However, NB had the highest recall and F1-score for predicting the against stance, whereas CNN performed poorly on identifying the against stance. 	<ul style="list-style-type: none"> The study demonstrated the strengths of ML techniques in handling large data sets. Social scientists can use ML techniques to scale up traditional content analysis.
Rashmi et al [31], 2021	<ul style="list-style-type: none"> The PCA^{PP} method provides the best classification accuracy for SVM (98%), followed by NB and RF (97%). 	<ul style="list-style-type: none"> The regional thermography and computer-aided diagnostic tool with ML classifier could be used as a primary noninvasive prognostic tool for evaluating obesity in children.
Snehalatha and Sangamithirai [30], 2021	<ul style="list-style-type: none"> Among the region of interest studied, the abdomen region exhibited a high temperature difference of 4.703% between normal participants and participants who were obese compared with other regions. The proposed custom network-2 provided an overall accuracy of 92%, with an area under the curve value of 0.948. By contrast, the pretrained model VGG16 produced an accuracy of 79% and an area under the curve value of 0.90 for discrimination into obese and normal thermograms. 	<ul style="list-style-type: none"> The DL system based on custom CNN provided a reliable classification performance to identify the occurrence of obesity in test participants. Custom CNN network-2 provided a commendable accuracy in classifying normal participants and participants who were obese from the thermal images. The trained custom-2 CNN model can be used for computer-aided screening of test participants for obesity detection.
Thamrin et al [29], 2021	<ul style="list-style-type: none"> Location, marital status, age group, education, sweet drinks, fatty or oily foods, grilled foods, preserved foods, seasoning powders, soft drinks or carbonated beverages, alcoholic beverages, mental or emotional disorders, diagnosed hypertension, physical activity, smoking, and fruit and vegetable consumption are significant in predicting obesity status in adults. The classification prediction using the logistic regression method achieves the best performance based on the accuracy metric (72%), specificity (71%), precision (69%), kappa (44%), and Fβ-score (70%). Classification prediction by the classification and regression tree method achieves the highest sensitivity (82%) and the highest F1-score (72%). With regard to the area under the receiver operating characteristic curve performance of the respective classification methods with 10-fold cross-validation, the logistic regression classifier has the highest average area under the receiver operating characteristic curve value (0.798). 	<ul style="list-style-type: none"> Logistic regression has a better performance than the classification and regression tree and NB methods. Kappa coefficients show only moderate concordance between predicted and measured obesity. The constructed obesity classification model can evaluate and predict the risk of obesity using ML methods for the population of Indonesia, which can then be applied to publicly available open data.
Zare et al [28], 2021	<ul style="list-style-type: none"> The kindergarten BMI z score is the most important predictor of obesity by grade 4. Including the kindergarten BMI z score of students in the model meaningfully increases the prediction accuracy. Logistic regression, RF, and neural network algorithms performed similarly in terms of accuracy, sensitivity, specificity, and area under the curve values. The 95% CIs around the area under the curve overlap among these 3 algorithms. The DT showed lower performance with an area under the curve value that was statistically lower than the area under the curve values from each of the other algorithms. Nevertheless, the performance of the DT algorithm was close to that of the others. 	<ul style="list-style-type: none"> Data from the Arkansas, United States, BMI screening program significantly improve the ability to identify children at a high risk of obesity to the extent that better prediction can be translated into more effective policy and better health outcomes. The ability to predict obesity by grade 4 was robust across the ML algorithms and logistic regression with these data.

^aAI: artificial intelligence.^bWHR: waist-to-hip ratio.^cAIM: abductor induction mechanism.^dCV: coefficient of variation.^eVAT: visceral adipose tissue.

^fSAT: subcutaneous adipose tissue.
^gSVM: support vector machine.
^hHC: hip circumference.
ⁱANN: artificial neural network.
^jBFP: body fat percentage.
^kELM: extreme learning machine.
^lBPNN: back propagation neural network.
^mID3: iterative dichotomizer 3.
ⁿCHICA: Child Health Improvement Through Computer Automation.
^oML: machine learning.
^pCRF: conditional random forest.
^qWHtR: waist-to-height ratio.
^rCC: calf circumference.
^sHC: hip circumference.
^tRMSE: root mean square error.
^uCCHMC: Cincinnati Children's Hospital and Medical Center.
^vBCH: Boston Children's Hospital.
^wBHS: Bogalusa Heart Study.
^xWGRS: weighted genetic risk score.
^yRF: random forest.
^zCNN: convolutional neural network.
^{aa}SNP: single-nucleotide polymorphism.
^{bb}WC: waist circumference.
^{cc}LASSO: least absolute shrinkage and selection operator.
^{dd}EHR: electronic health record.
^{ee}MIMIC: Multiparameter Intelligent Monitoring in Intensive Care.
^{ff}NLP: natural language processing.
^{gg}FHIR: Fast Healthcare Interoperability Resources.
^{hh}KNN: k-nearest neighbor.
ⁱⁱDL: deep learning.
^{jj}PU: positive and unlabeled.
^{kk}XGB: extreme gradient boosting.
^{ll}DNN: deep neural network.
^{mm}DT: decision tree.
ⁿⁿVGI: Visible Green Index.
^{oo}NB: naïve Bayes.
^{pp}PCA: principal component analysis.

Methodological Review

AI Overview

AI symbolizes the effort to automate intellectual tasks usually performed by humans [72]. In general, AI consists of 2 domains or developmental periods: symbolic AI and modern AI [73]. Symbolic AI prevailed from the 1950s to the 1980s, characterized by the endeavors to achieve human-level intelligence by having programmers handcraft a sufficiently large set of explicit rules for manipulating knowledge [74]. Although symbolic AI proved suitable for solving well-defined, logical problems, such as a rule-based question-answer system, it became intractable when creating rules to solve more complex, fuzzy issues such as image classification, speech recognition, and language translation [74]. The definition of ML is “the field of study that gives computers the ability to learn without being explicitly programmed” [75]. Instead of hard coding all the rules in the symbolic AI, researchers provide examples (eg,

images with labels that identify the objects in them) to *train* modern ML models to output rules [74]. As a subdomain of ML, DL is based on artificial neural networks in which multiple (*deep*) layers of artificial neurons are used to progressively extract higher-level features from data [76]. This layered representation enables the modeling of more complex, dynamic patterns compared with traditional ML (which sometimes is called *shallow learning* in contrast to DL), which finds its utility in analyzing *big data*: data massive in scale and *messy* to work with (eg, unstructured texts and images) [77]. The first ML and DL algorithms were developed in the 1950s, attracting initial excitement but then lying dormant for several decades [72]. Since the late 1980s, partly because of the rediscovery of backpropagation algorithms, the invention of CNNs, and the strong growth in computational capacity, ML and DL have regained their popularity vis-à-vis symbolic AI [72].

AI Versus Conventional Statistical Methods

Admittedly, the concept of conventional statistical methods is dubious at best because the development of statistical theories and algorithms is continual in time and intertwines at all levels [78]. Indeed, many *conventional* models fall into the ML domain, such as linear and logistic regressions. Despite the poorly defined domain and overlapping algorithms, at least 2 distinctions could be made between modern AI (ie, ML and DL) and other statistical methods. In terms of aims, the objective of AI models and their evaluation metrics predominantly concern prediction precision (often at the cost of compromising interpretability as models become complex) [78,79]. By contrast, conventional statistical approaches usually attempt to reveal relationships among variables (*statistical inference*) and focus on model interpretability [80]. In terms of procedures, it is standard practice to split data into training, validation, and test sets so that an AI model can be trained using the training set with the aim of achieving the optimal performance on some predefined evaluation metrics (eg, accuracy and mean squared error) when testing on the validation set [81,82]. The fine-tuned AI model is subsequently tested on the test set. The utility of the validation set is to prevent model overfitting (ie, too tailored to the training set while losing generalizability to new, unseen data) and fine-tune hyperparameters (ie, parameters external to the model, whose values cannot be automatically learned from data). The test set is preserved to test the final model's performance on unseen data. By contrast, conventional statistical methods do not usually fit and evaluate models using training, validation, and test sets but use other model selection criteria (eg, adjusted R-squared and Akaike and Bayesian information criteria) to evaluate model performance [83].

ML Subcategories

Overview

ML is classified into 2 subcategories: unsupervised ML and supervised ML [84]. Unsupervised ML analyzes and clusters unlabeled data sets, discovering hidden patterns or data groupings without the need for human intervention [85]. Its capability to reveal similarities and differences in information makes it ideal for exploratory data analysis. Unsupervised ML models are used for 3 main tasks: clustering, association, and dimensionality reduction [86]. Clustering algorithms (eg, k-means clustering, hierarchical clustering, and Gaussian mixture) group unlabeled data based on similarities [86]. Association algorithms (eg, Apriori, Eclat, and FP-Growth) identify rules and relations among variables in large databases [87]. Dimensionality reduction algorithms (eg, principal component analysis [PCA], singular value decomposition, and multidimensional scaling) deal with an excessive number of features during data preprocessing, reducing them to a manageable size while preserving the integrity of the data set as much as possible [88]. Supervised ML uses a training set consisting of input-output pairs to enable the algorithm to learn a function that maps input to output over time [89]. The algorithm measures its accuracy through the loss function, adjusting until the error is minimized sufficiently. The critical difference between supervised ML and unsupervised ML is that the former requires labeled data (ie, input-output pairs), whereas

the latter only requires inputs (ie, unlabeled data) [84]. Supervised ML models are used for 2 main tasks: classification and regression [84]. Classification algorithms assign data to specific categories (eg, obese or nonobese). Regression algorithms learn the relationship between input features and continuously distributed outcomes and are commonly used for projections (eg, BMI in 5 years).

Unsupervised ML

K-means Clustering

K-means clustering is an iterative algorithm that tries to partition the data set into a total of k nonoverlapping groups (ie, clusters) [86,90]. Each data point belongs to only 1 group. The algorithm attempts to make the intracluster data points as similar as possible while keeping the clusters apart. In particular, it assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (ie, arithmetic mean of all the data points belonging to that cluster) is minimized. As the number of clusters k needs to be determined before implementing the algorithm, silhouette coefficients are commonly used to identify the optimal k value. Lin et al [34] used k-means clustering to classify patients with obesity into 4 groups based on 3 biomarkers concerning glucose, insulin, and uric acid.

Fuzzy C-means Clustering

In nonfuzzy clustering (also known as hard clustering; for example, k-means clustering), data are divided into distinct clusters, where each data point can only belong to 1 cluster [86]. In fuzzy clustering, data points can potentially belong to multiple clusters [91]. Fuzzy c-means clustering assigns each data point membership from 0% to 100% in each cluster center [92]. The fuzzy partition coefficient is often used to determine the optimal number of clusters with a value ranging from 0 (worst) to 1 (best) [93]. Positano et al [71] used the fuzzy c-means algorithm to classify MRI pixels into clusters to assess abdominal fat.

Group Factor Analysis

Factor analysis describes relationships among the individual variables of a data set [94]. Group factor analysis (GFA) extends this classical formulation into describing relationships among groups of variables, where each group represents either a set of related variables or a data set [95]. GFA is commonly formulated as a latent variable model consisting of 2 hierarchical levels: the higher level models the relationships among the groups, and the lower-level models the observed variables given the higher level [95]. Kibble et al [44] used GFA to jointly analyze 5 large multivariate data sets to understand the multimolecular-level interactions associated with obesity development.

PCA for Large Data Sets

Large data sets are increasingly common nowadays. PCA is a classic, widely adopted method to reduce the dimensionality of a large data set while preserving as much statistical information (ie, variability) as possible [86]. In particular, PCA attempts to find new variables, called principal components, that are linear functions of those in the original data set. The new variables are uncorrelated with each other (ie, orthogonal) and maximize the projected data variance. Rashmi et al [31] used PCA to

reduce the feature dimensions of a thermal imaging data set to classify children by their obesity severity level.

Supervised ML

Linear Regression

Linear regression is considered a conventional statistical model and a classical architecture to develop a predictive model [96], but it fulfills all criteria from an ML point of view and is widely used as an ML algorithm to predict continuous outcomes such as BMI or BFP [97]. Trainable weights (ie, coefficients) of linear regression are commonly estimated using ordinary least squares or gradient descent. Compared with many other ML models, linear regression has the advantages of simplicity and interpretability [98]. It is easy to understand how the model reaches its predictions. Wang et al [56] used linear regressions to identify features of single-nucleotide polymorphisms that predict obesity risk. Phan et al [42] used linear regressions to estimate the associations between built environment indicators and state-level obesity prevalence.

Regularized Linear Regression

The bias-variance tradeoff is a fundamental issue faced by all ML models [86,99]. Bias is an error from erroneous assumptions in a learning algorithm. High bias may cause the algorithm to miss the relevant relations between features and outputs (called underfitting). Variance is an error from a learning algorithm's sensitivity to small fluctuations in the training set. A high variance may result from the algorithm modeling the random noise in the training data, often leading to the algorithm's poor generalizability to new, unseen data (called overfitting). In general, decreasing variance increases bias and vice versa, and ML algorithms need to be fine-tuned to balance these 2 properties. Regularization is an essential technique to prevent model overfitting and improve generalizability (at the cost of increasing bias) by adding a penalty term of trainable weights to the loss function [86]. Optimization algorithms that minimize the loss function will learn to avoid extreme weight values and thus reduce variance. The penalty term with the sum of squared trainable weights is called L2 regularization, used in Ridge regression. The penalty term with the sum of the absolute values of trainable weights is called L1 regularization, used in the least absolute shrinkage and selection operator (LASSO) regression. Unlike Ridge regression, LASSO regression often shrinks some feature weights to absolute zero, making it useful for feature selection. Finally, ElasticNet regression uses a weighted sum of L1 and L2 regularizations. Gerl et al [54] used LASSO regression to estimate the relationship between human plasma lipidomes and body weight outcomes, including BMI, WC, WHR, and BFP.

Logistic Regression

In its simplest form, logistic regression uses a logistic function, called the sigmoid function, to model a binary outcome [100]. A sigmoid function is a continuous, smooth, differentiable S-shaped mathematical function that maps a real number to a value in the range of 0 and 1, making it ideal for modeling probabilities. The estimated probabilities are converted to predictions (0 or 1, denoting exclusive group membership) based on some predefined threshold (eg, >0.5). In ML, logistic

regression often incorporates regularizations (L1, L2, or both) to prevent overfitting. Another common extension of logistic regression in ML is to solve multiclass classification problems when classification tasks involve >2 (exclusive) classes. A typical strategy uses the one-vs-rest method (also called one-vs-all) that fits 1 classifier (eg, a logistic regression) per class against all the other classes [101]. A data point is assigned to the class with the highest confidence score among all classifiers. Thamrin et al [29] used logistic regressions to assess the predictability of various obesity risk factors. Cheng et al [37] used logistic regressions to classify obesity status based on participants' physical activity levels.

NB Classifier

NB algorithms apply the Bayes theorem with the *naïve* assumption of conditional independence among each pair of features given the value of the class [102]. Despite this oversimplified assumption, NB classifiers have been widely used and have worked well in solving many real-world problems. The decoupling of conditional feature distributions allows each distribution to be independently estimated as 1D, making the training of NB classifiers much faster than more sophisticated ML models [86]. By contrast, the predicted probabilities of NB classifiers are less trustworthy owing to the algorithm's *naïve* assumption. Rashmi et al [31] used NB to classify childhood obesity based on thermogram images. Thamrin et al [29] adopted NB to predict adult obesity using Indonesian health survey data [29].

K-nearest Neighbor

K-nearest neighbor (KNN) is a nonparametric, supervised learning algorithm suitable for classification and regression tasks [103]. The input consists of the k closest training data points based on a prespecified distance measure (eg, Euclidean, Manhattan, or Minkowski distance). For classification tasks, the output is a class membership. A test data point is assigned to the class most common among its k -nearest neighbors (if $k=1$, the test data point is assigned to the class of the single nearest neighbor). For regression tasks, the output is the average value of its k -nearest neighbors. KNN should not be confused with k -means. The former is a supervised ML algorithm to determine the class or value of a data point based on its k -nearest neighbors, whereas the latter is an unsupervised ML algorithm to classify data points into k clusters that minimize the distances within clusters while maximizing those between clusters [90]. KNN is a memory-based learning algorithm that requires no training (called a lazy learner) but can become significantly slower when the sample size increases. Wang et al [56] used KNN to predict obesity risk based on features of single-nucleotide polymorphisms. Ramyaa et al [51] performed KNN to predict body weight using physical activity and dietary data.

Support Vector Machines

Support vector machines (SVMs), which are supervised learning models that construct a hyperplane in a high-dimensional space, can be used for classification and regression tasks [104]. SVMs attempt to identify the hyperplane separating different classes while maximizing the distance to any class's nearest training data point (ie, margin). Intuitively, the larger the margin, the

more likely the model's generalizability to new, unseen data. The choice of margin type can be critical for SVMs [86]. Hard-margin SVMs maximize the margin by minimizing the distance from the decision boundary to the training points. However, hard-margin SVMs may lead to overfitting and have no solution if the training data are linearly inseparable. Soft-margin SVMs modify the constraints of the hard-margin SVMs by allowing some data points to violate the margin (ie, misclassified). In practice, data are seldom linearly separable in the original feature space, and kernel methods are applied to map the input space of the data to a higher-dimensional feature space where linear models can be trained [105]. Many kernel functions, such as the Gaussian radial basis, sigmoid, and polynomial kernel, can be chosen. Wang et al [56] used SVM to predict obesity risk based on the features of single-nucleotide polymorphisms. Ramyaa et al [51] applied SVM to predict body weight using physical activity and diet data.

DT Algorithms

DTs are nonparametric supervised learning methods for classification and regression tasks [106]. In DT algorithms, a tree is built by splitting the source set that constitutes the tree's root node into subsets, which comprise the successor children [107]. The splitting is based on a set of rules applied to input features. Different splitting rules exist, such as variance reduction for regression tasks and Gini impurity or information gain for classification tasks. The splitting process is repeated on each derived subset recursively (ie, recursive partitioning). The recursion is completed when all subsets at a node share the same target value or when splitting no longer adds value to the predictions. DTs have several advantages over other ML algorithms, such as high transparency and interpretability and few requirements for data preprocessing [108]. However, DTs can be prone to overfitting (ie, too confident about the rules learned from the training set, which does not generalize well to the test set) and instability (minor variations in the data resulting in a very different tree). Using features extracted from electronic medical records, Hong et al [52] used DTs to predict obesity and 15 other comorbidities. Taghiyev et al [41] performed DTs to identify risk factors associated with obesity onset.

RF Models

Ensemble methods are approaches that aggregate the predictions of a group of models aiming for improved performance in classification or regression tasks [109]. Various ensemble methods exist, such as bagging, pasting, boosting, and stacking [86]. Bagging and pasting use the same training algorithm for every predictor included in the ensemble and train it on different random subsets of the training set. When sampling is performed with replacement, the method is called bagging; when sampling is performed without replacement, it is called pasting. RF is an ensemble of DTs commonly trained via the bagging or pasting method [110]. Specifically, RF fits many DTs on various subsets of the data and uses averaging to improve the predictive accuracy and prevent overfitting. For classification tasks, the RF output is the class selected by most trees; for regression tasks, the mean prediction of the individual trees is used. Some common hyperparameters of RF for fine-tuning include the number of trees in the forest, the maximum number of features considered for splitting a node, the maximum number of

branches in each tree, the minimum number of data points placed in a node before the node is split, the minimum number of data points allowed in a leaf node, and the method for sampling data points (ie, with or without replacement) [86]. RF typically produces more accurate and robust predictions than DTs and is one of the most popular supervised ML algorithms [111]. Using RF models, Hinojosa et al [58] examined the relationship between social and physical school environments and childhood obesity in California, United States. Dunstan et al [46] performed RF to predict national obesity prevalence using food sales data from 79 countries.

Extreme Gradient Boosting

Boosting refers to any ensemble method that combines several weak models into a strong one [112]. The difference between boosting and bagging and pasting is that in boosting, different models are applied to the entire training set sequentially, the new model attempting to address the weaknesses (eg, misclassified targets and residual errors) of the previous model. By contrast, in bagging and pasting, the same models are trained on different random subsets of the training set. A popular boosting algorithm is gradient boosting, in which the new model is trained on the residual errors made by the previous model [113]. Extreme gradient boosting (XGBoost) implements an optimized, parallel-tree gradient boosting algorithm, aiming to be highly efficient, flexible, and portable [114]. XGBoost is considered one of the most powerful ML algorithms, often serving as an essential component of winning entries in ML competitions [86]. A few drawbacks of XGBoost include lacking interpretability and being prone to overfitting. Pang et al [33] used XGBoost to predict early childhood obesity based on electronic health records. Alkutbe et al [27] applied gradient boosting to predict BFP based on cross-sectional health survey data collected in Saudi Arabia.

Multivariate Adaptive Regression Splines

Multivariate adaptive regression splines (MARS) is a nonparametric regression technique that automatically models nonlinearities and interactions among variables by combining ≥ 2 linear regressions using hinge functions [115,116]. A hinge function is a function equal to its argument where that argument is >0 and 0 everywhere else. MARS builds a model using a 2-phase procedure [117]. The forward phase starts with a model consisting of only the intercept term (ie, mean of the target) and repeatedly adds basis functions (ie, constant or hinge function) in pairs to the model that minimizes the squared error loss of the training set. The backward (or pruning) phase usually starts with an overfitted model and removes its least effective term at each step until the best submodel is found. MARS requires little or no data preparation, is easy to understand and interpret, and can address classification and regression tasks. However, it often underperforms boosting ensemble methods. Shao [65] applied MARS to predict BFP using a small-scale health record data set.

DL Models

In the obesity literature reviewed, DL models were applied to 3 distinct data types: tabular data (eg, spreadsheet data), images, and texts. The model architectures differ systematically across these data types.

DL on Tabular Data

Although *shallow* ML models perform well on tabular data sets in most cases, some complex relationships between the features and the target could be more effectively learned by a deep neural network model [118]. A fully connected neural network consists of a series of fully connected layers, with each artificial neuron (ie, node) of a layer linking with all neurons in the following layer [76]. A multilayer perceptron (MLP) is a classic fully connected neural network consisting of at least 3 layers of neurons: an input layer, a hidden layer, and an output layer [119]. One advantage of fully connected neural networks is that they are *structure agnostic*, requiring no specific assumptions about the input. However, neural networks trained on tabular data can sometimes be prone to overfitting [120]. Park and Edington [121] used MLP to identify individuals at elevated diabetic risk. Heydari et al [67] performed MLP to predict obesity status using data from a cross-sectional study of military personnel in Iran.

DL on Images

CV is a field of AI that enables computers to learn from digital images, videos, or other visual inputs and derive meaningful information for decision-making and recommendations [122,123]. Nowadays, most CV applications use DL models, which prove more capable than their *shallow-learning* (ie, ML models) counterparts in representing and revealing high-dimensional, complex nonlinear patterns inherent in image data. Specifically, CNNs consistently outperform the traditional densely connected neural networks (eg, MLP) and achieve human-like or superhuman accuracy in many challenging CV tasks ranging from image classification to object detection and segmentation [124,125]. The main advantages of CNNs over densely connected neural networks are locality, translation invariance, and computational efficiency [126]. Locality refers to the repeated use of small-sized kernels (or filters) in CNNs to identify local patterns at an increasing level of complexity (eg, from basic shapes such as lines and edges to complex objects such as adipose tissue or brain tumor). Translation invariance refers to CNNs' capacity to detect an entity independent of its position in the image. The computational efficiency of CNNs is achieved by using kernels, global pooling, and other techniques, which typically make the models much smaller (ie, fewer learnable parameters) than their densely connected counterparts. Over the past decade, numerous CNN-based DL models were built and adopted to tackle domain-specific CV problems [76,127]. Some landmark models include, but are not limited to, LeNet, AlexNet, VGG, Inception, ResNet, Xception, ResNeXt, and U-Net.

Transfer learning plays a crucial role in modern AI, where a model developed for a task is reused as the starting point for a model on a different but related task [128]; for instance, the ResNet model trained on ImageNet data with >14 million images in approximately 1000 categories (eg, tables and horses) has stored many useful visual patterns in its weights, which can help solve other CV tasks (eg, identifying fat tissues in MRI scans) [129]. Transfer learning can substantially reduce the number of images required to train a model for a particular task and boost model performance compared with models trained from scratch [130].

Maharana and Nsoesie [57] adopted the VGG model architecture to examine the relationship between obesity prevalence and the built environment measured by Google Maps images (eg, parks, highways, green streets, crosswalks, and diverse housing types). Similarly, Phan et al [42] used the VGG model to assess the link between the statewide prevalence of obesity, physical activity, and chronic disease mortality and the built environment using images from Google Street View. Bhanu et al [38] applied the U-Net model to identify adipose tissues from MRI data. Snehalatha and Sangamithirai [30] applied transfer learning on a pretrained CNN model to detect obesity based on thermal imaging data.

DL on Text

Besides CV, NLP is another field where DL dominates [131]. Early NLP models primarily adopted recurrent neural network (RNN) architecture, demonstrating broad applicability to various NLP tasks such as sentiment analysis, text summarization, language translation, and speech recognition [74,132]. RNN differs from feed-forward MLP in that it takes information from prior inputs (stored as *memories*) to influence the current input and output, which capitalizes on the structure of sequential data where order matters (eg, time series or natural languages) [133]. Some popular RNN models used in NLP tasks include gated recurrent unit and long short-term memory unit [74]. However, in today's NLP landscape, transformers, invented by a team at Google in 2017, have surpassed RNN models such as gated recurrent unit and long short-term memory unit [134-136]. Transformers are encoder-decoder models that use self-attention to process language sequences [137]. An encoder maps an input sequence into state representation vectors. A decoder decodes the state representation vector to generate the target output sequence. The self-attention mechanism is used repeatedly within the encoder and the decoder to help them contextualize the input data. Specifically, the mechanism compares every word in the sentence to every other word, including itself, and reweighs each word's embeddings to incorporate contextual relevance. Popular transformer models such as GPT-3, BERT, XLNet, RoBERTa, and T5 have been widely applied to various NLP tasks and achieved state-of-the-art results [137]. Stephens et al [48] tested the efficacy of pediatric obesity treatment support through Tess, a behavioral coaching chatbot built on NLP models. The study concluded that Tess demonstrated therapeutic values to pediatric patients with obesity and prediabetes, especially outside of office hours, and could be scaled up to serve a larger patient population.

Discussion

Overview

This study conducted a scoping review of the applications of AI to obesity research. A keyword search in digital bibliographic databases identified 46 studies that used diverse ML and DL models to study obesity-related outcomes. In general, the studies found AI models helpful in detecting clinically meaningful patterns of obesity or relationships between specific covariates and weight outcomes. The majority (18/22, 82%) of the studies comparing AI models with conventional statistical approaches found that the AI models achieved higher prediction accuracy

on test data. Some (5/46, 11%) of the studies comparing the performances of different AI models revealed mixed results, likely indicating the high contingency of model performance on the data set and task it was applied to. An accelerating trend of adopting state-of-the-art DL models over standard ML models was observed to address challenging CV and NLP tasks. We concisely introduced the popular ML and DL models and summarized their specific applications in the studies included in the review.

Despite the variety of ML and DL models used in obesity research, it could well be the beginning of the trend for using AI applications in the big data era. Future adoptions of AI in obesity research could be influenced by a broad spectrum of factors, with 3 prominent ones discussed in the following sections.

Artificial General Intelligence

The ML and DL models reviewed in this study were primarily unimodal and task specific: they were built on a single data type (eg, tabular, text, or image) to solve a specific problem such as obesity classification or BMI prediction. Recent advances in AI showcase the feasibility and possibly superior performance of multimodal, multitask ML and DL models that are trained on diverse data types (eg, tabular plus text, image, video, or audio) and can handle many domains of downstream tasks (eg, text generation, object detection, time series prediction, and speech recognition) simultaneously [138-140]. However, it should be noted that the predictive accuracy of AI models may vary across gender and age groups [27] and sex and age groups [59]. Different from BMI, BMI z scores adjust for sex and age differences [141]. Future research may evaluate the potential disparities in AI model performances in their applications to BMI versus BMI z scores as outcome measures. Artificial general intelligence (AGI) refers to the ability of an intelligent agent to understand or learn any intellectual task performed by a human being [142,143]. It is too early to tell whether these multimodal, multitask ML and DL models may lead to AGI (or whether we could ever achieve AGI through technological innovations) [144]. Nevertheless, we may soon witness increasing applications of these models in obesity-related research.

Synthetic Data Generation

Data access is fundamental to any AI model training. Two primary barriers with regard to data are limited sample size and confidentiality concerns [145-148]. ML and DL models are increasingly used to generate synthetic data as an alternative to data collected from the real world [149,150]. Synthetic data do not contain private information requiring human subject review and, therefore, can be shared with other parties or the public without confidentiality concerns [151]. By contrast, synthetic data preserve the original data's mathematical and statistical properties, ensuring that the AI model trained on them can be generalized to real-world data [152]. In addition, given the unrestrained availability of synthetic data (only limited by the computational power of data generation), AI models trained on synthetic data can be robust with regard to data variations [153]. Synthetic data of various types, such as tabular, text, and image, have been generated in massive quantities to train ML and DL

models cost-effectively. Obesity-related data or, more generally, health-related data can be expensive to collect (eg, MRI scans) and contain confidential information (eg, patients' names or residential addresses), which could be addressed by synthetic data generation [154].

Human-in-the-Loop

There have been increasing concerns over AI-related data bias and ethical issues [155,156]. Fundamentally, AI models should facilitate but not replace human judgment and decision-making [157,158]. Human-in-the-loop (HITL) is an AI model that requires human interaction [159,160]. HITL ensures that algorithm biases and potentially destructive model outputs can be identified in a timely manner and corrected to prevent adverse consequences. However, such interactions between humans and machines require thoughtful designs in the data-processing pipeline, model architecture, and personnel management [159]. Data- and model-driven decision-making related to obesity, such as behavioral modifications (eg, diet or physical activity interventions) or medical treatment, can be complex [161]. AI-powered wearables and other digital health platforms can detect change in an individual's physical activity and provide actionable information to improve health outcomes [162-164]. Mobile chemical sensors could offer timely dietary information by monitoring real-time chemical variations upon food consumption, collecting dynamic data based on an individual's metabolic profile and environmental exposure, thus supporting dietary behavior decision-making to improve precise nutrition [165]. HITL may integrate AI model outputs with expert inputs to make informed decisions that capitalize on the strengths of both and maximize patients' chances of health restoration and improvement [166].

Limitations of the Scoping Review and Included Studies

To our knowledge, this study is the first to systematically review AI-related methodologies adopted in the obesity literature and project trends for future technological development and applications. However, several limitations should be noted concerning this review and the included studies. As our review focused on ML and DL methods, study-specific findings (eg, the effectiveness of an intervention and estimated associations between covariates and an outcome) were not synthesized in detail. The included studies were heterogeneous in terms of hypothesis and research question, study design, population sampled, data collection method, sample size, and data quality. The analytic approach chosen was endogenous to these study-specific parameters; therefore, across-study comparisons of model performances may not be reliable. Even within the same study, conclusions about relative model performances (eg, the prediction accuracy of logistic regression vs SVM) may lack generalizability because of the interdependency between data and ML and DL algorithms. AI technologies are rapidly advancing, with innovations and breakthroughs almost daily. A review such as this one will have a short shelf life and warrant periodic updates.

Conclusions

This study reviewed the AI-related methodologies adopted in the obesity literature, particularly ML and DL models applied to tabular, image, and text data for obesity measurement, prediction, and treatment. It aimed to provide researchers and

practitioners with an overview of the AI applications to obesity research, familiarize them with popular ML and DL models, and facilitate their adoption of AI applications. The review also discussed emerging trends such as multimodal and multitask AI models, synthetic data generation, and HITL, which may witness increasing applications in obesity research.

Acknowledgments

This research was partially funded by the Fundamental Research Funds for the Central Universities, China University of Geosciences, Beijing (grant 2-9-2020-036).

Authors' Contributions

RA designed the study and wrote the manuscript. RA and JS jointly designed the search algorithm and screened articles. JS performed data extraction and constructed the summary tables. YX drafted part of the Discussion section. JS and YX revised the manuscript. The co-first authors RA and JS contributed equally.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search algorithm used in PubMed.

[\[DOC File , 12 KB-Multimedia Appendix 1\]](#)

References

1. The double burden of malnutrition. *The Lancet*. 2019 Dec 16. URL: <https://www.thelancet.com/series/double-burden-malnutrition> [accessed 2022-06-18]
2. An R, Ji M, Zhang S. Global warming and obesity: a systematic review. *Obes Rev* 2018 Mar;19(2):150-163. [doi: [10.1111/obr.12624](https://doi.org/10.1111/obr.12624)] [Medline: [28977817](https://pubmed.ncbi.nlm.nih.gov/28977817/)]
3. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014 Aug 30;384(9945):766-781 [FREE Full text] [doi: [10.1016/S0140-6736\(14\)60460-8](https://doi.org/10.1016/S0140-6736(14)60460-8)] [Medline: [24880830](https://pubmed.ncbi.nlm.nih.gov/24880830/)]
4. Obesity and overweight. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> [accessed 2022-06-18]
5. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet* 2016 Apr 02;387(10026):1377-1396 [FREE Full text] [doi: [10.1016/S0140-6736\(16\)30054-X](https://doi.org/10.1016/S0140-6736(16)30054-X)] [Medline: [27115820](https://pubmed.ncbi.nlm.nih.gov/27115820/)]
6. Sivarajah U, Kamal M, Irani Z, Weerakkody V. Critical analysis of Big Data challenges and analytical methods. *J Business Res* 2017 Jan;70:263-286 [FREE Full text] [doi: [10.1016/j.jbusres.2016.08.001](https://doi.org/10.1016/j.jbusres.2016.08.001)]
7. Dash S, Shakyawar S, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. *J Big Data* 2019 Jun 19;6(1):54 [FREE Full text] [doi: [10.1186/s40537-019-0217-0](https://doi.org/10.1186/s40537-019-0217-0)]
8. Agrawal R, Prabakaran S. Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity (Edinb)* 2020 Apr;124(4):525-534 [FREE Full text] [doi: [10.1038/s41437-020-0303-2](https://doi.org/10.1038/s41437-020-0303-2)] [Medline: [32139886](https://pubmed.ncbi.nlm.nih.gov/32139886/)]
9. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, et al. Artificial intelligence: a powerful paradigm for scientific research. *Innovation (Camb)* 2021 Nov 28;2(4):100179 [FREE Full text] [doi: [10.1016/j.xinn.2021.100179](https://doi.org/10.1016/j.xinn.2021.100179)] [Medline: [34877560](https://pubmed.ncbi.nlm.nih.gov/34877560/)]
10. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J Ambient Intell Humaniz Comput* 2022 Jan 13:1-28 [FREE Full text] [doi: [10.1007/s12652-021-03612-z](https://doi.org/10.1007/s12652-021-03612-z)] [Medline: [35039756](https://pubmed.ncbi.nlm.nih.gov/35039756/)]
11. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJ. Artificial intelligence in radiology. *Nat Rev Cancer* 2018 Aug;18(8):500-510 [FREE Full text] [doi: [10.1038/s41568-018-0016-5](https://doi.org/10.1038/s41568-018-0016-5)] [Medline: [29777175](https://pubmed.ncbi.nlm.nih.gov/29777175/)]
12. Syrowatka A, Kuznetsova M, Alsubai A, Beckman AL, Bain PA, Craig KJ, et al. Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *NPJ Digit Med* 2021 Jun 10;4(1):96 [FREE Full text] [doi: [10.1038/s41746-021-00459-8](https://doi.org/10.1038/s41746-021-00459-8)] [Medline: [34112939](https://pubmed.ncbi.nlm.nih.gov/34112939/)]
13. Bin Sawad A, Narayan B, Alnefaie A, Maqbool A, Mckie I, Smith J, et al. A systematic review on healthcare artificial intelligent conversational agents for chronic conditions. *Sensors (Basel)* 2022 Mar 29;22(7):2625 [FREE Full text] [doi: [10.3390/s22072625](https://doi.org/10.3390/s22072625)] [Medline: [35408238](https://pubmed.ncbi.nlm.nih.gov/35408238/)]

14. Goh YS, Ow Yong JQ, Chee BQ, Kuek JH, Ho CS. Machine learning in health promotion and behavioral change: scoping review. *J Med Internet Res* 2022 Jun 02;24(6):e35831 [FREE Full text] [doi: [10.2196/35831](https://doi.org/10.2196/35831)] [Medline: [35653177](https://pubmed.ncbi.nlm.nih.gov/35653177/)]
15. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. In: *Artificial Intelligence in Healthcare*. Cambridge, Massachusetts, United States: Academic Press; 2020.
16. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med* 2022 Jan;28(1):31-38. [doi: [10.1038/s41591-021-01614-0](https://doi.org/10.1038/s41591-021-01614-0)] [Medline: [35058619](https://pubmed.ncbi.nlm.nih.gov/35058619/)]
17. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform* 2018 Nov 27;19(6):1236-1246 [FREE Full text] [doi: [10.1093/bib/bbx044](https://doi.org/10.1093/bib/bbx044)] [Medline: [28481991](https://pubmed.ncbi.nlm.nih.gov/28481991/)]
18. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J* 2019 Jun;6(2):94-98 [FREE Full text] [doi: [10.7861/futurehosp.6-2-94](https://doi.org/10.7861/futurehosp.6-2-94)] [Medline: [31363513](https://pubmed.ncbi.nlm.nih.gov/31363513/)]
19. Aung YY, Wong DC, Ting DS. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *Br Med Bull* 2021 Sep 10;139(1):4-15. [doi: [10.1093/bmb/ldab016](https://doi.org/10.1093/bmb/ldab016)] [Medline: [34405854](https://pubmed.ncbi.nlm.nih.gov/34405854/)]
20. Quinn TP, Senadeera M, Jacobs S, Coghlan S, Le V. Trust and medical AI: the challenges we face and the expertise needed to overcome them. *J Am Med Inform Assoc* 2021 Mar 18;28(4):890-894 [FREE Full text] [doi: [10.1093/jamia/ocaa268](https://doi.org/10.1093/jamia/ocaa268)] [Medline: [33340404](https://pubmed.ncbi.nlm.nih.gov/33340404/)]
21. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019 Oct 29;17(1):195 [FREE Full text] [doi: [10.1186/s12916-019-1426-2](https://doi.org/10.1186/s12916-019-1426-2)] [Medline: [31665002](https://pubmed.ncbi.nlm.nih.gov/31665002/)]
22. Chew HS, Achananuparp P. Perceptions and needs of artificial intelligence in health care to increase adoption: scoping review. *J Med Internet Res* 2022 Jan 14;24(1):e32939 [FREE Full text] [doi: [10.2196/32939](https://doi.org/10.2196/32939)] [Medline: [35029538](https://pubmed.ncbi.nlm.nih.gov/35029538/)]
23. Marmett B, Carvalho RB, Fortes MS, Cazella SC. Artificial Intelligence technologies to manage obesity. *Vittalle J Health Sci* 2018 Sep 27;30(2):73-79. [doi: [10.14295/vittalle.v30i2.7654](https://doi.org/10.14295/vittalle.v30i2.7654)]
24. Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: review of the literature. *J Med Internet Res* 2019 Apr 05;21(4):e12286 [FREE Full text] [doi: [10.2196/12286](https://doi.org/10.2196/12286)] [Medline: [30950797](https://pubmed.ncbi.nlm.nih.gov/30950797/)]
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
26. Abdel-Aal RE, Mangoud AM. Modeling obesity using abductive networks. *Comput Biomed Res* 1997 Dec;30(6):451-471. [doi: [10.1006/cbmr.1997.1460](https://doi.org/10.1006/cbmr.1997.1460)] [Medline: [9466835](https://pubmed.ncbi.nlm.nih.gov/9466835/)]
27. Alkutbe RB, Alruban A, Alturki H, Sattar A, Al-Hazzaa H, Rees G. Fat mass prediction equations and reference ranges for Saudi Arabian Children aged 8-12 years using machine technique method. *PeerJ* 2021;9:e10734 [FREE Full text] [doi: [10.7717/peerj.10734](https://doi.org/10.7717/peerj.10734)] [Medline: [33665006](https://pubmed.ncbi.nlm.nih.gov/33665006/)]
28. Zare S, Thomsen MR, Nayga RM, Goudie A. Use of machine learning to determine the information value of a BMI screening program. *Am J Prev Med* 2021 Mar;60(3):425-433 [FREE Full text] [doi: [10.1016/j.amepre.2020.10.016](https://doi.org/10.1016/j.amepre.2020.10.016)] [Medline: [33483154](https://pubmed.ncbi.nlm.nih.gov/33483154/)]
29. Thamrin SA, Arsyad DS, Kuswanto H, Lawi A, Nasir S. Predicting obesity in adults using machine learning techniques: an analysis of Indonesian basic health research 2018. *Front Nutr* 2021;8:669155 [FREE Full text] [doi: [10.3389/fnut.2021.669155](https://doi.org/10.3389/fnut.2021.669155)] [Medline: [34235168](https://pubmed.ncbi.nlm.nih.gov/34235168/)]
30. U S, K. PT, K S. Computer aided diagnosis of obesity based on thermal imaging using various convolutional neural networks. *Biomed Signal Process Control* 2021 Jan;63:102233 [FREE Full text] [doi: [10.1016/j.bspc.2020.102233](https://doi.org/10.1016/j.bspc.2020.102233)]
31. Rashmi R, Umapathy S, Krishnan P. Thermal imaging method to evaluate childhood obesity based on machine learning techniques. *Int J Imaging Syst Technol* 2021 Mar 20;31(3):1752-1768 [FREE Full text] [doi: [10.1002/ima.22572](https://doi.org/10.1002/ima.22572)]
32. Park HJ, Francisco SC, Pang MR, Peng L, Chi G. Exposure to anti-black lives matter movement and obesity of the black population. *Soc Sci Med* 2021 Jul 28;114265. [doi: [10.1016/j.socscimed.2021.114265](https://doi.org/10.1016/j.socscimed.2021.114265)] [Medline: [34366168](https://pubmed.ncbi.nlm.nih.gov/34366168/)]
33. Pang X, Forrest CB, Lê-Scherban F, Masino AJ. Prediction of early childhood obesity with machine learning and electronic health record data. *Int J Med Inform* 2021 Jun;150:104454 [FREE Full text] [doi: [10.1016/j.ijmedinf.2021.104454](https://doi.org/10.1016/j.ijmedinf.2021.104454)] [Medline: [33866231](https://pubmed.ncbi.nlm.nih.gov/33866231/)]
34. Lin Z, Feng W, Liu Y, Ma C, Arefan D, Zhou D, et al. Machine learning to identify metabolic subtypes of obesity: a multi-center study. *Front Endocrinol (Lausanne)* 2021;12:713592 [FREE Full text] [doi: [10.3389/fendo.2021.713592](https://doi.org/10.3389/fendo.2021.713592)] [Medline: [34335479](https://pubmed.ncbi.nlm.nih.gov/34335479/)]
35. Lee K, Kim HY, Lee SJ, Kwon SO, Na S, Hwang HS, Korean Society of Ultrasound in ObstetricsGynecology Research Group. Prediction of newborn's body mass index using nationwide multicenter ultrasound data: a machine-learning study. *BMC Pregnancy Childbirth* 2021 Mar 02;21(1):172 [FREE Full text] [doi: [10.1186/s12884-021-03660-5](https://doi.org/10.1186/s12884-021-03660-5)] [Medline: [33653299](https://pubmed.ncbi.nlm.nih.gov/33653299/)]
36. Delnevo G, Mancini G, Rocchetti M, Salomoni P, Trombini E, Andrei F. The prediction of body mass index from negative affectivity through machine learning: a confirmatory study. *Sensors (Basel)* 2021 Mar 29;21(7):2361 [FREE Full text] [doi: [10.3390/s21072361](https://doi.org/10.3390/s21072361)] [Medline: [33805257](https://pubmed.ncbi.nlm.nih.gov/33805257/)]
37. Cheng X, Lin S, Liu J, Liu S, Zhang J, Nie P, et al. Does physical activity predict obesity-a machine learning and statistical method-based analysis. *Int J Environ Res Public Health* 2021 Apr 09;18(8):3966 [FREE Full text] [doi: [10.3390/ijerph18083966](https://doi.org/10.3390/ijerph18083966)] [Medline: [33918760](https://pubmed.ncbi.nlm.nih.gov/33918760/)]

38. Bhanu PK, Arvind CS, Yeow LY, Chen WX, Lim WS, Tan CH. CAFT: a deep learning-based comprehensive abdominal fat analysis tool for large cohort studies. *MAGMA* 2022 Apr;35(2):205-220. [doi: [10.1007/s10334-021-00946-9](https://doi.org/10.1007/s10334-021-00946-9)] [Medline: [34338926](https://pubmed.ncbi.nlm.nih.gov/34338926/)]
39. Yao Y, Song L, Ye J. Motion-to-BMI: using motion sensors to predict the body mass index of smartphone users. *Sensors (Basel)* 2020 Feb 19;20(4):1134 [FREE Full text] [doi: [10.3390/s20041134](https://doi.org/10.3390/s20041134)] [Medline: [32093013](https://pubmed.ncbi.nlm.nih.gov/32093013/)]
40. Xiao Y, Zhang Y, Sun Y, Tao P, Kuang X. Does green space really matter for residents' obesity? A new perspective from Baidu street view. *Front Public Health* 2020;8:332 [FREE Full text] [doi: [10.3389/fpubh.2020.00332](https://doi.org/10.3389/fpubh.2020.00332)] [Medline: [32850579](https://pubmed.ncbi.nlm.nih.gov/32850579/)]
41. Taghiyev A, Altun A, Caglar S. A hybrid approach based on machine learning to identify the causes of obesity. *J Control Eng Applied Informatic* 2020;22(2):56-66.
42. Phan L, Yu W, Keralis JM, Mukhija K, Dwivedi P, Brunisholz KD, et al. Google street view derived built environment indicators and associations with state-level obesity, physical activity, and chronic disease mortality in the united states. *Int J Environ Res Public Health* 2020 May 22;17(10):3659 [FREE Full text] [doi: [10.3390/ijerph17103659](https://doi.org/10.3390/ijerph17103659)] [Medline: [32456114](https://pubmed.ncbi.nlm.nih.gov/32456114/)]
43. Park B, Chung C, Lee MJ, Park H. Accurate neuroimaging biomarkers to predict body mass index in adolescents: a longitudinal study. *Brain Imaging Behav* 2020 Oct;14(5):1682-1695. [doi: [10.1007/s11682-019-00101-y](https://doi.org/10.1007/s11682-019-00101-y)] [Medline: [31065926](https://pubmed.ncbi.nlm.nih.gov/31065926/)]
44. Kibble M, Khan SA, Ammad-Ud-Din M, Bollepalli S, Palviainen T, Kaprio J, et al. An integrative machine learning approach to discovering multi-level molecular mechanisms of obesity using data from monozygotic twin pairs. *R Soc Open Sci* 2020 Oct;7(10):200872 [FREE Full text] [doi: [10.1098/rsos.200872](https://doi.org/10.1098/rsos.200872)] [Medline: [33204460](https://pubmed.ncbi.nlm.nih.gov/33204460/)]
45. Fu Y, Gou W, Hu W, Mao Y, Tian Y, Liang X, et al. Integration of an interpretable machine learning algorithm to identify early life risk factors of childhood obesity among preterm infants: a prospective birth cohort. *BMC Med* 2020 Jul 10;18(1):184 [FREE Full text] [doi: [10.1186/s12916-020-01642-6](https://doi.org/10.1186/s12916-020-01642-6)] [Medline: [32646442](https://pubmed.ncbi.nlm.nih.gov/32646442/)]
46. Dunstan J, Aguirre M, Bastías M, Nau C, Glass TA, Tobar F. Predicting nationwide obesity from food sales using machine learning. *Health Informatics J* 2020 Mar;26(1):652-663 [FREE Full text] [doi: [10.1177/1460458219845959](https://doi.org/10.1177/1460458219845959)] [Medline: [31106648](https://pubmed.ncbi.nlm.nih.gov/31106648/)]
47. Blanes-Selva V, Tortajada S, Vilar R, Valdivieso B, García-Gómez JM. Machine learning-based identification of obesity from positive and unlabelled electronic health records. *Stud Health Technol Inform* 2020 Jun 16;270:864-868. [doi: [10.3233/SHTI200284](https://doi.org/10.3233/SHTI200284)] [Medline: [32570505](https://pubmed.ncbi.nlm.nih.gov/32570505/)]
48. Stephens TN, Joerin A, Rauws M, Werk LN. Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. *Transl Behav Med* 2019 May 16;9(3):440-447. [doi: [10.1093/tbm/ibz043](https://doi.org/10.1093/tbm/ibz043)] [Medline: [31094445](https://pubmed.ncbi.nlm.nih.gov/31094445/)]
49. Shin S, Lee J, Choe S, Yang HI, Min J, Ahn K, et al. Dry electrode-based body fat estimation system with anthropometric data for use in a wearable device. *Sensors (Basel)* 2019 May 10;19(9):2177 [FREE Full text] [doi: [10.3390/s19092177](https://doi.org/10.3390/s19092177)] [Medline: [31083445](https://pubmed.ncbi.nlm.nih.gov/31083445/)]
50. Scheinker D, Valencia A, Rodriguez F. Identification of factors associated with variation in US county-level obesity prevalence rates using epidemiologic vs machine learning models. *JAMA Netw Open* 2019 Apr 05;2(4):e192884 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.2884](https://doi.org/10.1001/jamanetworkopen.2019.2884)] [Medline: [31026030](https://pubmed.ncbi.nlm.nih.gov/31026030/)]
51. Ramyaa R, Hosseini O, Krishnan GP, Krishnan S. Phenotyping women based on dietary macronutrients, physical activity, and body weight using machine learning tools. *Nutrients* 2019 Jul 22;11(7):1681 [FREE Full text] [doi: [10.3390/nu11071681](https://doi.org/10.3390/nu11071681)] [Medline: [31336626](https://pubmed.ncbi.nlm.nih.gov/31336626/)]
52. Hong N, Wen A, Stone DJ, Tsuji S, Kingsbury PR, Rasmussen LV, et al. Developing a FHIR-based EHR phenotyping framework: a case study for identification of patients with obesity and multiple comorbidities from discharge summaries. *J Biomed Inform* 2019 Nov;99:103310 [FREE Full text] [doi: [10.1016/j.jbi.2019.103310](https://doi.org/10.1016/j.jbi.2019.103310)] [Medline: [31622801](https://pubmed.ncbi.nlm.nih.gov/31622801/)]
53. Hammond R, Athanasiadou R, Curado S, Aphinyanaphongs Y, Abrams C, Messito MJ, et al. Predicting childhood obesity using electronic health records and publicly available data. *PLoS ONE* 2019 Apr 22;14(4):e0215571 [FREE Full text] [doi: [10.1371/journal.pone.0215571](https://doi.org/10.1371/journal.pone.0215571)] [Medline: [31009509](https://pubmed.ncbi.nlm.nih.gov/31009509/)]
54. Gerl MJ, Klose C, Surma MA, Fernandez C, Melander O, Männistö S, et al. Machine learning of human plasma lipidomes for obesity estimation in a large population cohort. *PLoS Biol* 2019 Oct;17(10):e3000443 [FREE Full text] [doi: [10.1371/journal.pbio.3000443](https://doi.org/10.1371/journal.pbio.3000443)] [Medline: [31626640](https://pubmed.ncbi.nlm.nih.gov/31626640/)]
55. Duran I, Martakis K, Rehberg M, Semler O, Schoenau E. Diagnostic performance of an artificial neural network to predict excess body fat in children. *Pediatr Obes* 2019 Feb;14(2):e12494. [doi: [10.1111/ijpo.12494](https://doi.org/10.1111/ijpo.12494)] [Medline: [30590878](https://pubmed.ncbi.nlm.nih.gov/30590878/)]
56. Wang H, Chang S, Lin W, Chen C, Chiang S, Huang K, et al. Machine learning-based method for obesity risk evaluation using single-nucleotide polymorphisms derived from next-generation sequencing. *J Comput Biol* 2018 Dec;25(12):1347-1360. [doi: [10.1089/cmb.2018.0002](https://doi.org/10.1089/cmb.2018.0002)] [Medline: [30204480](https://pubmed.ncbi.nlm.nih.gov/30204480/)]
57. Maharana A, Nsoesie EO. Use of deep learning to examine the association of the built environment with prevalence of neighborhood adult obesity. *JAMA Netw Open* 2018 Aug 03;1(4):e181535 [FREE Full text] [doi: [10.1001/jamanetworkopen.2018.1535](https://doi.org/10.1001/jamanetworkopen.2018.1535)] [Medline: [30646134](https://pubmed.ncbi.nlm.nih.gov/30646134/)]
58. Ortega Hinojosa AM, MacLeod KE, Balmes J, Jerrett M. Influence of school environments on childhood obesity in California. *Environ Res* 2018 Oct;166:100-107. [doi: [10.1016/j.envres.2018.04.022](https://doi.org/10.1016/j.envres.2018.04.022)] [Medline: [29883903](https://pubmed.ncbi.nlm.nih.gov/29883903/)]

59. Seyednasrollah F, Mäkelä J, Pitkänen N, Juonala M, Hutri-Kähönen N, Lehtimäki T, et al. Prediction of adulthood obesity using genetic and childhood clinical risk factors in the cardiovascular risk in young finns study. *Circ Cardiovasc Genet* 2017 Jun;10(3):e001554 [FREE Full text] [doi: [10.1161/CIRCGENETICS.116.001554](https://doi.org/10.1161/CIRCGENETICS.116.001554)] [Medline: [28620069](https://pubmed.ncbi.nlm.nih.gov/28620069/)]
60. Lingren T, Thaker V, Brady C, Namjou B, Kennebeck S, Bickel J, et al. Developing an algorithm to detect early childhood obesity in two tertiary pediatric medical centers. *Appl Clin Inform* 2016 Jul 20;7(3):693-706 [FREE Full text] [doi: [10.4338/ACI-2016-01-RA-0015](https://doi.org/10.4338/ACI-2016-01-RA-0015)] [Medline: [27452794](https://pubmed.ncbi.nlm.nih.gov/27452794/)]
61. Almeida SM, Furtado JM, Mascarenhas P, Ferraz ME, Silva LR, Ferreira JC, et al. Anthropometric predictors of body fat in a large population of 9-year-old school-aged children. *Obes Sci Pract* 2016 Sep;2(3):272-281 [FREE Full text] [doi: [10.1002/osp4.51](https://doi.org/10.1002/osp4.51)] [Medline: [27708844](https://pubmed.ncbi.nlm.nih.gov/27708844/)]
62. Nau C, Ellis H, Huang H, Schwartz BS, Hirsch A, Bailey-Davis L, et al. Exploring the forest instead of the trees: an innovative method for defining obesogenic and obesoprotective environments. *Health Place* 2015 Sep;35:136-146 [FREE Full text] [doi: [10.1016/j.healthplace.2015.08.002](https://doi.org/10.1016/j.healthplace.2015.08.002)] [Medline: [26398219](https://pubmed.ncbi.nlm.nih.gov/26398219/)]
63. Dugan T, Mukhopadhyay S, Carroll A, Downs S. Machine learning techniques for prediction of early childhood obesity. *Appl Clin Inform* 2017 Dec 19;06(03):506-520 [FREE Full text] [doi: [10.4338/aci-2015-03-ra-0036](https://doi.org/10.4338/aci-2015-03-ra-0036)]
64. Chen H, Yang B, Liu D, Liu W, Liu Y, Zhang X, et al. Using blood indexes to predict overweight statuses: an extreme learning machine-based approach. *PLoS One* 2015;10(11):e0143003 [FREE Full text] [doi: [10.1371/journal.pone.0143003](https://doi.org/10.1371/journal.pone.0143003)] [Medline: [26600199](https://pubmed.ncbi.nlm.nih.gov/26600199/)]
65. Shao YE. Body fat percentage prediction using intelligent hybrid approaches. *ScientificWorldJournal* 2014;2014:383910 [FREE Full text] [doi: [10.1155/2014/383910](https://doi.org/10.1155/2014/383910)] [Medline: [24723804](https://pubmed.ncbi.nlm.nih.gov/24723804/)]
66. Kupusinac A, Stokić E, Doroslovački R. Predicting body fat percentage based on gender, age and BMI by using artificial neural networks. *Comput Methods Programs Biomed* 2014 Feb;113(2):610-619. [doi: [10.1016/j.cmpb.2013.10.013](https://doi.org/10.1016/j.cmpb.2013.10.013)] [Medline: [24275480](https://pubmed.ncbi.nlm.nih.gov/24275480/)]
67. Heydari ST, Ayatollahi SM, Zare N. Comparison of artificial neural networks with logistic regression for detection of obesity. *J Med Syst* 2012 Aug;36(4):2449-2454. [doi: [10.1007/s10916-011-9711-4](https://doi.org/10.1007/s10916-011-9711-4)] [Medline: [21556898](https://pubmed.ncbi.nlm.nih.gov/21556898/)]
68. Zhang S, Tjortjts C, Zeng X, Qiao H, Buchan I, Keane J. Comparing data mining methods with logistic regression in childhood obesity prediction. *Inf Syst Front* 2009 Feb 24;11(4):449-460 [FREE Full text] [doi: [10.1007/s10796-009-9157-0](https://doi.org/10.1007/s10796-009-9157-0)]
69. Yang H, Spasic I, Keane JA, Nenadic G. A text mining approach to the prediction of disease status from clinical discharge summaries. *J Am Med Inform Assoc* 2009;16(4):596-600 [FREE Full text] [doi: [10.1197/jamia.M3096](https://doi.org/10.1197/jamia.M3096)] [Medline: [19390098](https://pubmed.ncbi.nlm.nih.gov/19390098/)]
70. Ergün U. The classification of obesity disease in logistic regression and neural network methods. *J Med Syst* 2009 Feb;33(1):67-72. [doi: [10.1007/s10916-008-9165-5](https://doi.org/10.1007/s10916-008-9165-5)] [Medline: [19238898](https://pubmed.ncbi.nlm.nih.gov/19238898/)]
71. Positano V, Cusi K, Santarelli MF, Sironi A, Petz R, Defronzo R, et al. Automatic correction of intensity inhomogeneities improves unsupervised assessment of abdominal fat by MRI. *J Magn Reson Imaging* 2008 Aug;28(2):403-410. [doi: [10.1002/jmri.21448](https://doi.org/10.1002/jmri.21448)] [Medline: [18666138](https://pubmed.ncbi.nlm.nih.gov/18666138/)]
72. Haenlein M, Kaplan A. A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. *California Manag Rev* 2019 Jul 17;61(4):5-14. [doi: [10.1177/0008125619864925](https://doi.org/10.1177/0008125619864925)]
73. Artificial Intelligence. *Stanford Encyclopedia of Philosophy*. URL: <https://plato.stanford.edu/> [accessed 2022-06-18]
74. Chollet F. *Deep Learning with Python, Second Edition*. Shelter Island, New York, United States: Manning; 2021.
75. Samuel A. Some studies in machine learning using the game of checkers. *IBM J Res Dev* 1959 Jul;3(3):210-229 [FREE Full text] [doi: [10.1147/rd.33.0210](https://doi.org/10.1147/rd.33.0210)]
76. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 2021;8(1):53 [FREE Full text] [doi: [10.1186/s40537-021-00444-8](https://doi.org/10.1186/s40537-021-00444-8)] [Medline: [33816053](https://pubmed.ncbi.nlm.nih.gov/33816053/)]
77. Chauhan N, Singh K. A review on conventional machine learning vs deep learning. In: *Proceedings of the International Conference on Computing, Power and Communication Technologies (GUCON)*. 2018 Presented at: International Conference on Computing, Power and Communication Technologies (GUCON); Sep 28-29, 2018; Greater Noida, India. [doi: [10.1109/gucon.2018.8675097](https://doi.org/10.1109/gucon.2018.8675097)]
78. Rajula HS, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina (Kaunas)* 2020 Sep 08;56(9):455 [FREE Full text] [doi: [10.3390/medicina56090455](https://doi.org/10.3390/medicina56090455)] [Medline: [32911665](https://pubmed.ncbi.nlm.nih.gov/32911665/)]
79. Bennett M, Hayes K, Kleczyk E, Mehta R. Similarities and differences between machine learning and traditional advanced statistical modeling in healthcare analytics. *arXiv* 2022 [FREE Full text] [doi: [10.5772/intechopen.105116](https://doi.org/10.5772/intechopen.105116)]
80. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: making sense of the differences. *Knee Surg Sports Traumatol Arthrosc* 2022 Mar;30(3):753-757. [doi: [10.1007/s00167-022-06896-6](https://doi.org/10.1007/s00167-022-06896-6)] [Medline: [35106604](https://pubmed.ncbi.nlm.nih.gov/35106604/)]
81. KhosrowHassibi. Machine learning vs. traditional statistics: different philosophies, different approaches. *Data Science Central*. 2016 Oct 28. URL: <https://www.datasciencecentral.com/machine-learning-vs-traditional-statistics-different-> [accessed 2022-06-18]
82. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv*. 2018 Nov. URL: <https://arxiv.org/pdf/1811.12808.pdf> [accessed 2022-06-18]

83. Ding J, Tarokh V, Yang Y. Model selection techniques: an overview. arXiv. URL: <https://arxiv.org/pdf/> [accessed 2022-06-18]
84. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf A. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. Cham: Springer; 2019.
85. Wittek P. Quantum Machine Learning What Quantum Computing Means to Data Mining. Boston: Academic Press; 2014.
86. Géron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems. Sebastopol, California, United States: O'Reilly Media; 2017.
87. Addi A, Tarik A, Fatima G. Comparative survey of association rule mining algorithms based on multiple-criteria decision analysis approach. In: Proceedings of the 3rd International Conference on Control, Engineering & Information Technology (CEIT). 2015 Presented at: 3rd International Conference on Control, Engineering & Information Technology (CEIT); May 25-27, 2015; Tlemcen, Algeria. [doi: [10.1109/ceit.2015.7233078](https://doi.org/10.1109/ceit.2015.7233078)]
88. Velliangiri S, Alagumuthukrishnan S, Thankumar Joseph SI. A review of dimensionality reduction techniques for efficient computation. *Procedia Comput Sci* 2019;165:104-111. [doi: [10.1016/j.procs.2020.01.079](https://doi.org/10.1016/j.procs.2020.01.079)]
89. Singh A, Thakur N, Sharma A. A review of supervised machine learning algorithms. In: Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom). 2016 Presented at: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom); Mar 16-18, 2016; New Delhi, India.
90. Ashabi A, Sahibuddin S, Haghighi MS. The systematic review of K-means clustering algorithm. In: Proceedings of the 2020 The 9th International Conference on Networks, Communication and Computing. 2020 Presented at: ICNCC 2020: 2020 The 9th International Conference on Networks, Communication and Computing; Dec 18 - 20, 2020; Tokyo Japan. [doi: [10.1145/3447654.3447657](https://doi.org/10.1145/3447654.3447657)]
91. Gosain A, Dahiya S. Performance analysis of various fuzzy clustering algorithms: a review. *Procedia Comput Sci* 2016;79:100-111. [doi: [10.1016/j.procs.2016.03.014](https://doi.org/10.1016/j.procs.2016.03.014)]
92. Arora J, Khatter K, Tushir M. Fuzzy c-Means Clustering Strategies: A Review of Distance Measures. Singapore: Springer; 2018.
93. Zhang M, Zhang W, Sicotte H, Yang P. A new validity measure for a correlation-based fuzzy c-means clustering algorithm. In: Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2009 Presented at: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Sep 03-06, 2009; Minneapolis, MN, USA. [doi: [10.1109/iembs.2009.5332582](https://doi.org/10.1109/iembs.2009.5332582)]
94. Tavakol M, Wetzel A. Factor analysis: a means for theory and instrument development in support of construct validity. *Int J Med Educ* 2020 Nov 06;11:245-247 [FREE Full text] [doi: [10.5116/ijme.5f96.0f4a](https://doi.org/10.5116/ijme.5f96.0f4a)] [Medline: [33170146](https://pubmed.ncbi.nlm.nih.gov/33170146/)]
95. Klami A, Virtanen S, Leppäaho E, Kaski S. Group factor analysis. *IEEE Trans Neural Netw Learn Syst* 2015 Sep;26(9):2136-2147. [doi: [10.1109/TNNLS.2014.2376974](https://doi.org/10.1109/TNNLS.2014.2376974)] [Medline: [25532193](https://pubmed.ncbi.nlm.nih.gov/25532193/)]
96. Steyerberg E. Clinical Prediction Models A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
97. Dasgupta A, Sun YV, König IR, Bailey-Wilson JE, Malley JD. Brief review of regression-based and machine learning methods in genetic epidemiology: the Genetic Analysis Workshop 17 experience. *Genet Epidemiol* 2011;35 Suppl 1:S5-11 [FREE Full text] [doi: [10.1002/gepi.20642](https://doi.org/10.1002/gepi.20642)] [Medline: [22128059](https://pubmed.ncbi.nlm.nih.gov/22128059/)]
98. Gosiewska A, Kozak A, Biecek P. Simpler is better: lifting interpretability-performance trade-off via automated feature engineering. *Decision Support Syst* 2021 Nov;150:113556. [doi: [10.1016/j.dss.2021.113556](https://doi.org/10.1016/j.dss.2021.113556)]
99. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proc Natl Acad Sci U S A* 2019 Aug 06;116(32):15849-15854 [FREE Full text] [doi: [10.1073/pnas.1903070116](https://doi.org/10.1073/pnas.1903070116)] [Medline: [31341078](https://pubmed.ncbi.nlm.nih.gov/31341078/)]
100. Bewick V, Cheek L, Ball J. Statistics review 14: logistic regression. *Crit Care* 2005 Feb;9(1):112-118 [FREE Full text] [doi: [10.1186/cc3045](https://doi.org/10.1186/cc3045)] [Medline: [15693993](https://pubmed.ncbi.nlm.nih.gov/15693993/)]
101. Rifkin R, Klautau A. In defense of one-vs-all classification. *J Mach Learn Res* 2004;5:101-141.
102. Wickramasinghe I, Kalutarage H. Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation. *Soft Comput* 2020 Sep 09;25(3):2277-2293. [doi: [10.1007/s00500-020-05297-6](https://doi.org/10.1007/s00500-020-05297-6)]
103. Taunk K, De S, Verma S, Swetapadma A. A brief review of nearest neighbor algorithm for learning and classification. In: Proceedings of the 2019 International Conference on Intelligent Computing and Control Systems (ICCS). 2019 Presented at: 2019 International Conference on Intelligent Computing and Control Systems (ICCS); May 15-17, 2019; Madurai, India URL: <https://doi.org/10.1109/ICCS45141.2019.9065747> [doi: [10.1109/iccs45141.2019.9065747](https://doi.org/10.1109/iccs45141.2019.9065747)]
104. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: applications, challenges and trends. *Neurocomputing* 2020 Sep;408:189-215. [doi: [10.1016/j.neucom.2019.10.118](https://doi.org/10.1016/j.neucom.2019.10.118)]
105. Ponte P, Melko RG. Kernel methods for interpretable machine learning of order parameters. *Phys Rev B* 2017 Nov 27;96(20). [doi: [10.1103/physrevb.96.205146](https://doi.org/10.1103/physrevb.96.205146)]
106. Kotsiantis SB. Decision trees: a recent overview. *Artif Intell Rev* 2011 Jun 29;39(4):261-283. [doi: [10.1007/s10462-011-9272-4](https://doi.org/10.1007/s10462-011-9272-4)]

107. Podgorelec V, Zorman M. Decision tree learning. In: Encyclopedia of Complexity and Systems Science. Berlin, Heidelberg: Springer; 2015.
108. Somvanshi M, Chavan P, Tambade S, Shinde S. A review of machine learning techniques using decision tree and support vector machine. In: Proceedings of the 2016 International Conference on Computing Communication Control and automation (ICCUBEA). 2016 Presented at: 2016 International Conference on Computing Communication Control and automation (ICCUBEA); Aug 12-13, 2016; Pune, India. [doi: [10.1109/iccubea.2016.7860040](https://doi.org/10.1109/iccubea.2016.7860040)]
109. Re M, Valentini G. Ensemble methods: a review. In: Advances in Machine Learning and Data Mining for Astronomy. London, United Kingdom: Chapman & Hall; 2012.
110. Parmar A, Katariya R, Patel V. A review on random forest: an ensemble classifier. In: International Conference on Intelligent Data Communication Technologies and Internet of Things. Cham: Springer; 2018.
111. Talekar B. A detailed review on decision tree and random forest. Biosci Biotech Res Comm 2020 Dec 28;13(14):245-248. [doi: [10.21786/bbrc/13.14/57](https://doi.org/10.21786/bbrc/13.14/57)]
112. Ferreira A, Figueiredo M. Boosting algorithms: a review of methods, theory, and applications. In: Ensemble Machine Learning. Boston, MA: Springer; 2012.
113. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. Artif Intell Rev 2020 Aug 24;54(3):1937-1967. [doi: [10.1007/s10462-020-09896-5](https://doi.org/10.1007/s10462-020-09896-5)]
114. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13 - 17, 2016; San Francisco California USA. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
115. Prihastuti Yasmirullah SD, Otok BW, Trijoyo Purnomo JD, Prastyo DD. Modification of Multivariate Adaptive Regression Spline (MARS). J Phys Conf Ser 2021 Mar 01;1863(1):012078. [doi: [10.1088/1742-6596/1863/1/012078](https://doi.org/10.1088/1742-6596/1863/1/012078)]
116. Friedman J. Multivariate adaptive regression splines. Ann Statist 1991 Mar 1;19(1):1-67 [FREE Full text] [doi: [10.1214/aos/1176347963](https://doi.org/10.1214/aos/1176347963)]
117. Zhang W, Goh A. Multivariate adaptive regression splines for analysis of geotechnical engineering systems. Comput Geotechnics 2013 Mar;48:82-95. [doi: [10.1016/j.compgeo.2012.09.016](https://doi.org/10.1016/j.compgeo.2012.09.016)]
118. Zhong G, Ling X, Wang L. From shallow feature learning to deep learning: benefits from the width and depth of deep architectures. WIREs Data Min Knowl 2018 Mar 28;9(1):e1255 [FREE Full text] [doi: [10.1002/widm.1255](https://doi.org/10.1002/widm.1255)]
119. Gardner M, Dorling S. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric Environ 1998 Aug;32(14-15):2627-2636. [doi: [10.1016/s1352-2310\(97\)00447-0](https://doi.org/10.1016/s1352-2310(97)00447-0)]
120. Rynkiewicz J. On overfitting of multilayer perceptrons for classification. In: ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. 2019 Presented at: ESANN 2019 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning; Apr 24-26, 2019; Bruges, Belgium URL: <https://www.esann.org/sites/default/files/proceedings/legacy/es2019-80.pdf> [doi: [10.1007/978-1-4615-3210-1_4](https://doi.org/10.1007/978-1-4615-3210-1_4)]
121. Park J, Edington DW. Application of a prediction model for identification of individuals at diabetic risk. Methods Inf Med 2004;43(3):273-281. [Medline: [15227557](https://pubmed.ncbi.nlm.nih.gov/15227557/)]
122. Voulodimos A, Doulamis N, Bebis G, Stathaki T. Recent developments in deep learning for engineering applications. Comput Intell Neurosci 2018;2018:8141259-8142018 [FREE Full text] [doi: [10.1155/2018/8141259](https://doi.org/10.1155/2018/8141259)] [Medline: [29861713](https://pubmed.ncbi.nlm.nih.gov/29861713/)]
123. Chai J, Zeng H, Li A, Ngai EW. Deep learning in computer vision: a critical review of emerging techniques and application scenarios. Mach Learn Application 2021 Dec;6:100134. [doi: [10.1016/j.mlwa.2021.100134](https://doi.org/10.1016/j.mlwa.2021.100134)]
124. Dhillon A, Verma GK. Convolutional neural network: a review of models, methodologies and applications to object detection. Prog Artif Intell 2019 Dec 20;9(2):85-112. [doi: [10.1007/s13748-019-00203-0](https://doi.org/10.1007/s13748-019-00203-0)]
125. Khan A, Sohail A, Zahoor U, Qureshi AS. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 2020 Apr 21;53(8):5455-5516. [doi: [10.1007/s10462-020-09825-6](https://doi.org/10.1007/s10462-020-09825-6)]
126. Stevens E, Antiga L, Viehmann T. Deep Learning with PyTorch Build, Train, and Tune Neural Networks Using Python Tools. Shelter Island, New York, United States: Manning; 2020.
127. Aloysius N, Geetha M. A review on deep convolutional neural networks. In: Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP). 2017 Presented at: 2017 International Conference on Communication and Signal Processing (ICCSP); Apr 06-08, 2017; Chennai, India. [doi: [10.1109/iccsp.2017.8286426](https://doi.org/10.1109/iccsp.2017.8286426)]
128. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. Proc IEEE 2021 Jan;109(1):43-76. [doi: [10.1109/jproc.2020.3004555](https://doi.org/10.1109/jproc.2020.3004555)]
129. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016; Las Vegas, NV, USA. [doi: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90)]
130. Weiss K, Khoshgoftaar T, Wang D. A survey of transfer learning. J Big Data 2016 May 28;3(1):1-40 [FREE Full text] [doi: [10.1186/s40537-016-0043-6](https://doi.org/10.1186/s40537-016-0043-6)]
131. Li H. Deep learning for natural language processing: advantages and challenges. National Sci Rev 2017;5(1):24-26 [FREE Full text]

132. Le Glaz A, Haralambous Y, Kim-Dufor D, Lenca P, Billot R, Ryan TC, et al. Machine learning and natural language processing in mental health: systematic review. *J Med Internet Res* 2021 May 04;23(5):e15708 [[FREE Full text](#)] [doi: [10.2196/15708](https://doi.org/10.2196/15708)] [Medline: [33944788](https://pubmed.ncbi.nlm.nih.gov/33944788/)]
133. Lipton ZC, Berkowitz J, Elkan C. A critical review of recurrent neural networks for sequence learning. *arXiv* 2015 Jun 5 [[FREE Full text](#)]
134. Chernyavskiy A, Ilvovsky D, Nakov P. Transformers: “the end of history” for natural language processing? In: *Machine Learning and Knowledge Discovery in Databases*. Cham: Springer; 2021.
135. Lin T, Wang Y, Liu X, Qiu X. A survey of transformers. *AI Open* 2022;3:111-132 [[FREE Full text](#)] [doi: [10.1016/j.aiopen.2022.10.001](https://doi.org/10.1016/j.aiopen.2022.10.001)]
136. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN. Attention is all you need. In: *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*. 2017 Presented at: 31st Conference on Neural Information Processing Systems (NIPS 2017); Dec 4 - 9, 2017; Long Beach, CA, USA URL: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
137. Tunstall L, von WL, Wolf T. *Natural Language Processing with Transformers, Revised Edition*. Sebastopol, California, United States: O'Reilly Media; 2022.
138. Summaira J, Li X, Shoib AM, Bourahla O, Songyuan L, Abdul J. Recent advances and trends in multimodal deep learning: a review. *arXiv* 2021 May [[FREE Full text](#)]
139. Bayoudh K, Knani R, Hamdaoui F, Mtibaa A. A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets. *Vis Comput* 2022;38(8):2939-2970 [[FREE Full text](#)] [doi: [10.1007/s00371-021-02166-7](https://doi.org/10.1007/s00371-021-02166-7)] [Medline: [34131356](https://pubmed.ncbi.nlm.nih.gov/34131356/)]
140. Zhang Y, Yang Q. An overview of multi-task learning. *National Sci Rev* 2018;5(1):30-43. [doi: [10.1093/nsr/nwx105](https://doi.org/10.1093/nsr/nwx105)]
141. Fagerberg P, Charmandari E, Diou C, Heimeier R, Karavidopoulou Y, Kassari P, et al. Fast eating is associated with increased BMI among high-school students. *Nutrients* 2021 Mar 09;13(3):880 [[FREE Full text](#)] [doi: [10.3390/nu13030880](https://doi.org/10.3390/nu13030880)] [Medline: [33803093](https://pubmed.ncbi.nlm.nih.gov/33803093/)]
142. Nyalapelli V, Gandhi M, Bhargava S, Dhanare R, Bothe S. Review of progress in artificial general intelligence and human brain inspired cognitive architecture. In: *Proceedings of the 2021 International Conference on Computer Communication and Informatics (ICCCI)*. 2021 Presented at: 2021 International Conference on Computer Communication and Informatics (ICCCI); Jan 27-29, 2021; Coimbatore, India. [doi: [10.1109/iccci50826.2021.9402446](https://doi.org/10.1109/iccci50826.2021.9402446)]
143. Long L, Cotner C. A review and proposed framework for artificial general intelligence. In: *Proceedings of the 2019 IEEE Aerospace Conference*. 2019 Presented at: 2019 IEEE Aerospace Conference; Mar 02-09, 2019; Big Sky, MT, USA. [doi: [10.1109/aero.2019.8742125](https://doi.org/10.1109/aero.2019.8742125)]
144. Fjelland R. Why general artificial intelligence will not be realized. *Humanit Soc Sci Commun* 2020 Jun 17;7(1). [doi: [10.1057/s41599-020-0494-4](https://doi.org/10.1057/s41599-020-0494-4)]
145. Bae H, Jang J, Jung D, Jang H, Ha H, Lee H, et al. Security and privacy issues in deep learning. *arXiv* 2021 Mar [[FREE Full text](#)]
146. Ha T, Dang TK, Le H, Truong TA. Security and privacy issues in deep learning: a brief review. *Sn Comput Sci* 2020 Aug 06;1(5). [doi: [10.1007/s42979-020-00254-4](https://doi.org/10.1007/s42979-020-00254-4)]
147. Keshari R, Ghosh S, Chhabra S, Vatsa M, Singh R. Unravelling small sample size problems in the deep learning world. In: *Proceedings of the 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*. 2020 Presented at: 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM); Sep 24-26, 2020; New Delhi, India. [doi: [10.1109/bigmm50055.2020.00028](https://doi.org/10.1109/bigmm50055.2020.00028)]
148. Liu B, Wei Y, Zhang Y, Yang Q. Deep neural networks for high dimension, low sample size data. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*. 2017 Presented at: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17); Aug 19-25, 2017; Melbourne, Australia URL: <https://www.ijcai.org/Proceedings/2017/0318.pdf> [doi: [10.24963/ijcai.2017/318](https://doi.org/10.24963/ijcai.2017/318)]
149. Goncalves A, Ray P, Soper B, Stevens J, Coyle L, Sales AP. Generation and evaluation of synthetic patient data. *BMC Med Res Methodol* 2020 May 07;20(1):108 [[FREE Full text](#)] [doi: [10.1186/s12874-020-00977-1](https://doi.org/10.1186/s12874-020-00977-1)] [Medline: [32381039](https://pubmed.ncbi.nlm.nih.gov/32381039/)]
150. Chen RJ, Lu MY, Chen TY, Williamson DF, Mahmood F. Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 2021 Jun;5(6):493-497 [[FREE Full text](#)] [doi: [10.1038/s41551-021-00751-8](https://doi.org/10.1038/s41551-021-00751-8)] [Medline: [34131324](https://pubmed.ncbi.nlm.nih.gov/34131324/)]
151. Rankin D, Black M, Bond R, Wallace J, Mulvenna M, Epelde G. Reliability of supervised machine learning using synthetic data in health care: model to preserve privacy for data sharing. *JMIR Med Inform* 2020 Jul 20;8(7):e18910 [[FREE Full text](#)] [doi: [10.2196/18910](https://doi.org/10.2196/18910)] [Medline: [32501278](https://pubmed.ncbi.nlm.nih.gov/32501278/)]
152. El Emam K, Mosquera L, Hoptroff R. *Practical Synthetic Data Generation Balancing Privacy and the Broad Availability of Data*. Sebastopol, California, United States: O'Reilly Media; 2020.
153. Sergey I N. *Synthetic Data for Deep Learning*. Cham: Springer International Publishing; 2021.
154. Moya-Sáez E, Peña-Nogales Ó, Luis-García R, Alberola-López C. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. *Comput Methods Programs Biomed* 2021 Oct;210:106371 [[FREE Full text](#)] [doi: [10.1016/j.cmpb.2021.106371](https://doi.org/10.1016/j.cmpb.2021.106371)] [Medline: [34525411](https://pubmed.ncbi.nlm.nih.gov/34525411/)]

155. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv* 2021 Jul;54(6):1-35. [doi: [10.1145/3457607](https://doi.org/10.1145/3457607)]
156. Zhou N, Zhang Z, Nair V, Singhal H, Chen J. Bias, fairness and accountability with artificial intelligence and machine learning algorithms. *Int Statistical Rev* 2022 Apr 10;90(3):468-480 [FREE Full text] [doi: [10.1111/insr.12492](https://doi.org/10.1111/insr.12492)]
157. Zerilli J, Knott A, Maclaurin J, Gavaghan C. Algorithmic decision-making and the control problem. *Minds Mach* 2019 Dec 11;29(4):555-578. [doi: [10.1007/s11023-019-09513-7](https://doi.org/10.1007/s11023-019-09513-7)]
158. Lepri B, Oliver N, Pentland A. Ethical machines: the human-centric use of artificial intelligence. *iScience* 2021 Mar 19;24(3):102249 [FREE Full text] [doi: [10.1016/j.isci.2021.102249](https://doi.org/10.1016/j.isci.2021.102249)] [Medline: [33763636](https://pubmed.ncbi.nlm.nih.gov/33763636/)]
159. Monarch R. *Human-in-the-Loop Machine Learning Active Learning and Annotation for Human-centered AI*. Shelter Island, New York, United States: Manning Publications; 2021.
160. Wu X, Xiao L, Sun Y, Zhang J, Ma T, He L. A survey of human-in-the-loop for machine learning. *Future Gen Comput Syst* 2022 Oct;135:364-381 [FREE Full text] [doi: [10.1016/j.future.2022.05.014](https://doi.org/10.1016/j.future.2022.05.014)]
161. Timmins KA, Green MA, Radley D, Morris MA, Pearce J. How has big data contributed to obesity research? A review of the literature. *Int J Obes (Lond)* 2018 Dec;42(12):1951-1962 [FREE Full text] [doi: [10.1038/s41366-018-0153-7](https://doi.org/10.1038/s41366-018-0153-7)] [Medline: [30022056](https://pubmed.ncbi.nlm.nih.gov/30022056/)]
162. Sapci AH, Sapci HA. Innovative assisted living tools, remote monitoring technologies, artificial intelligence-driven solutions, and robotic systems for aging societies: systematic review. *JMIR Aging* 2019 Nov 29;2(2):e15429 [FREE Full text] [doi: [10.2196/15429](https://doi.org/10.2196/15429)] [Medline: [31782740](https://pubmed.ncbi.nlm.nih.gov/31782740/)]
163. Guisado-Fernández E, Giunti G, Mackey LM, Blake C, Caulfield BM. Factors influencing the adoption of smart health technologies for people with dementia and their informal caregivers: scoping review and design framework. *JMIR Aging* 2019 Apr 30;2(1):e12192 [FREE Full text] [doi: [10.2196/12192](https://doi.org/10.2196/12192)] [Medline: [31518262](https://pubmed.ncbi.nlm.nih.gov/31518262/)]
164. Wilmlink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. *JMIR Aging* 2020 Sep 10;3(2):e19554 [FREE Full text] [doi: [10.2196/19554](https://doi.org/10.2196/19554)] [Medline: [32723711](https://pubmed.ncbi.nlm.nih.gov/32723711/)]
165. Sempionatto J, Montiel V, Vargas E, Teymourian H, Wang J. Wearable and mobile sensors for personalized nutrition. *ACS Sens* 2021 May 28;6(5):1745-1760 [FREE Full text] [doi: [10.1021/acssensors.1c00553](https://doi.org/10.1021/acssensors.1c00553)] [Medline: [34008960](https://pubmed.ncbi.nlm.nih.gov/34008960/)]
166. Patel BN, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J, et al. Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *NPJ Digit Med* 2019 Nov 18;2(1):111 [FREE Full text] [doi: [10.1038/s41746-019-0189-7](https://doi.org/10.1038/s41746-019-0189-7)] [Medline: [31754637](https://pubmed.ncbi.nlm.nih.gov/31754637/)]

Abbreviations

- AGI:** artificial general intelligence
- AI:** artificial intelligence
- BFP:** body fat percentage
- CNN:** convolutional neural network
- CV:** computer vision
- DL:** deep learning
- DT:** decision tree
- GFA:** group factor analysis
- HITL:** human-in-the-loop
- KNN:** k-nearest neighbor
- LASSO:** least absolute shrinkage and selection operator
- MARS:** multivariate adaptive regression splines
- ML:** machine learning
- MLP:** multilayer perceptron
- MRI:** magnetic resonance imaging
- NB:** naïve Bayes
- NLP:** natural language processing
- PCA:** principal component analysis
- PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- PRISMA-ScR:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
- RF:** random forest
- RNN:** recurrent neural network
- SVM:** support vector machine
- WC:** waist circumference
- WHR:** waist-to-hip ratio
- XGBoost:** extreme gradient boosting

Edited by R Kukafka; submitted 28.06.22; peer-reviewed by N Maglaveras, B Puladi; comments to author 30.08.22; revised version received 05.10.22; accepted 01.11.22; published 07.12.22

Please cite as:

An R, Shen J, Xiao Y

Applications of Artificial Intelligence to Obesity Research: Scoping Review of Methodologies

J Med Internet Res 2022;24(12):e40589

URL: <https://www.jmir.org/2022/12/e40589>

doi: [10.2196/40589](https://doi.org/10.2196/40589)

PMID: [36476515](https://pubmed.ncbi.nlm.nih.gov/36476515/)

©Ruopeng An, Jing Shen, Yunyu Xiao. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 07.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.