



## Machine Learning Techniques in Adaptive and Personalized Systems for Health and Wellness

Oladapo Oyeboode, Jonathon Fowles, Darren Steeves & Rita Orji

**To cite this article:** Oladapo Oyeboode, Jonathon Fowles, Darren Steeves & Rita Orji (2023) Machine Learning Techniques in Adaptive and Personalized Systems for Health and Wellness, International Journal of Human-Computer Interaction, 39:9, 1938-1962, DOI: [10.1080/10447318.2022.2089085](https://doi.org/10.1080/10447318.2022.2089085)

**To link to this article:** <https://doi.org/10.1080/10447318.2022.2089085>



[View supplementary material](#)



[Published online: 27 Jul 2022.](#)



[Submit your article to this journal](#)



[Article views: 1619](#)



[View related articles](#)



[View Crossmark data](#)



[Citing articles: 7 View citing articles](#)



# Machine Learning Techniques in Adaptive and Personalized Systems for Health and Wellness

Oladapo Oyeboade<sup>a</sup> , Jonathon Fowles<sup>b</sup> , Darren Steeves<sup>c</sup>, and Rita Orji<sup>a</sup> 

<sup>a</sup>Faculty of Computer Science, Dalhousie University, Halifax, Canada; <sup>b</sup>School of Kinesiology, Acadia University, Wolfville, Canada; <sup>c</sup>School of Health and Human Performance, Dalhousie University, Halifax, Canada

## ABSTRACT

Traditional health systems mostly rely on rules created by experts to offer adaptive interventions to patients. However, with recent advances in artificial intelligence (AI) and machine learning (ML) techniques, health-related systems are becoming more sophisticated with higher accuracy in providing more personalized interventions or treatments to individual patients. In this paper, we present an extensive literature review to explore the current trends in ML-based adaptive systems for health and well-being. We conduct a systematic search for articles published between January 2011 and April 2022 and selected 87 articles that met our inclusion criteria for review. The selected articles target 18 health and wellness domains including disease management, assistive healthcare, medical diagnosis, mental health, physical activity, dietary management, health monitoring, substance use, smoking cessation, homeopathy remedy finding, patient privacy, mobile health (mHealth) apps finder, clinician knowledge representation for neonatal emergency care, dental and oral health, medication management, disease surveillance, medical specialty recommendation, and health awareness. Our review focuses on five key areas across the target domains: data collection strategies, model development process, ML techniques utilized, model evaluation techniques, as well as adaptive or personalization strategies for health and wellness interventions. We also identified various technical and methodological challenges including data volume constraints, data quality issues, data diversity or variability issues, infrastructure-related issues, and suitability of interventions which offer directions for future work in this area. Finally, we offer recommendations for tackling these challenges, leveraging on technological advances such as multi-modality, Cloud technology, online learning, edge computing, automatic re-calibration, Bluetooth auto-reconnection, feedback pipeline, federated learning, explainable AI, and co-creation of health and wellness interventions.

## 1. Introduction

Personalized medicine (also called precision medicine) which is concerned with tailoring healthcare to individual patients has received increased global attention over the years. For example, the International Consortium for Personalized Medicine (ICPerMed) predicted that personalized medicine will be deployed across healthcare systems by 2030 (Venne et al., 2020; Vicente et al., 2020). This vision is increasingly becoming a reality with recent advances and widespread applications of artificial intelligence. Specifically, machine learning (ML) has been applied in diverse health and wellness domains including detection/diagnosis of medical conditions (e.g., diseases, mental health disorders, adverse events, etc.) (Hung et al., 2019; Shi et al., 2018; Tolkachev et al., 2021), medication adherence (Thyde et al., 2021; L. Wang et al., 2020), medical imaging (Liao et al., 2020; Loram et al., 2020), treatment recommendations (Basu et al., 2020; Zeng, 2020), and healthcare decision making (Liang et al., 2014; Loftus et al., 2020; Ying et al., 2021). Research has shown that access to relevant structured and

unstructured data from diverse modalities contributed significantly to the effectiveness of these ML models (Jiang et al., 2017). Examples of these modalities include, but not limited to, electronic health records (EHR) which contain patients' personal health data (e.g., demographic information, healthcare history including test results, medication profiles, physicians' clinical notes, and diagnostic images); biomedical sensors which capture physiological data and vital signs in real-time; motion and position sensors which track activity and location in real-time; environment sensors which monitor environmental properties (e.g., ambient humidity and ambient pressure); microphone and camera for audio, video, and images; and social media. Hence, ability to identify clinically relevant patterns in single or multi-modal data and stratify patients based on these patterns makes ML models suitable for personalized healthcare (Stafford et al., 2020).

In the past, a number of reviews that focus on applications of machine learning in various health domains have emerged. Existing research discusses opportunities, impact,

and challenges of ML for healthcare (Ben-Israel et al., 2020; Miotto et al., 2018; Waring et al., 2020), as well as accuracy, reliability, and effectiveness of ML algorithms in specific health domains such as disease management (Martin-Isla et al., 2020; Mlodzinski et al., 2020; Thomsen et al., 2020; Zhu et al., 2021), health surveillance (Gupta & Katarya, 2020), mental health (Cho et al., 2019; Le Glaz et al., 2021; Shatte et al., 2019), and other health-related issues (e.g., Qian et al., 2021). Other works focus on the effectiveness of digital interventions aided by ML algorithms without investigating for personalization (Triantafyllidis & Tsanas, 2019) or discuss ML-based disease profiling and personalized treatments for specific diseases only (Buettner et al., 2020). However, none of these works explore how ML techniques are applied to drive adaptive or personalized interventions in the general area of health and wellness on a large scale. Therefore, to the best of our knowledge, our work is the first comprehensive review to investigate advances in ML-based adaptive and personalized systems in the general area of health and wellness.

Our systematic review offers four main contributions in the area of ML-driven adaptive systems in HCI. First, we examine the types of data used for model training and how they are acquired and processed. Second, we explore the conventional and advanced ML models utilized to promote adaptivity, as well as the techniques employed in developing and evaluating the models. Third, we capture model performance using various evaluation metrics and examine how the ML models are integrated into interactive systems to achieve adaptive/personalized health and wellness interventions. Finally, we identify various technical and methodological limitations or challenges of existing systems and offer suggestions on how the challenges could be addressed to improve the performance, reliability, and effectiveness of ML-driven adaptive systems.

This review paper is structured as follows: first, we describe our methodology based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach (Moher et al., 2009). Next, we present our findings, and finally, we discuss our findings and offer practical suggestions for developing ML-based adaptive systems that deliver effective personalized or tailored interventions to individual users to promote their health and well-being.

## 2. Methodology

We conducted our systematic review by searching six popular electronic databases: PubMed, Association for Computing Machinery (ACM) Digital Library, IEEE Xplore, ScienceDirect, Scopus, and Web of Science. PubMed is a credible public and unifying database for biomedical literature from MEDLINE, life science journals, and online books (National Library of Medicine, n.d.), while ACM Digital Library and IEEE Xplore are reputable databases for computing research. ScienceDirect is a leading database for computing, medical, and scientific research. Scopus and Web of Science are huge and unifying databases covering multiple fields including computing and medicine. We focused on peer-reviewed papers of conference proceedings and journals written in English and published in the last 11 years:

between January 2011 and April 2022. The search was conducted on April 6, 2022. Our systematic review followed the PRISMA guidelines (Moher et al., 2009). Figure 1 summarizes our paper selection process.

### 2.1. Paper identification

Our search was based on abstracts, titles, and metadata using specific keywords combined using logical operators (AND/OR). Although the syntax for advanced search differs among the databases, the search query common to all databases (except ScienceDirect) includes (“machine learning” OR “deep learning”) AND (app OR application OR system OR intervention) AND (mobile OR smartphone OR web OR “web based”) AND (health OR healthcare OR wellness OR “well being”) AND (adapt\* OR personaliz\* OR personalis\* OR recommend\* OR tailor\*). The wild character (\*) in the query represents “begins with”; for example, *personaliz\** matches *personalize*, *personalized*, and *personalization*. ScienceDirect only supports up to seven logical operators and does not allow wild characters; hence we adjusted our search query accordingly without losing the main keywords: (“machine learning” OR “deep learning”) AND (app OR system) AND (mobile OR web) AND (health OR wellness OR “well being”). In total, 2133 papers were identified from the six sources and screened by the authors.

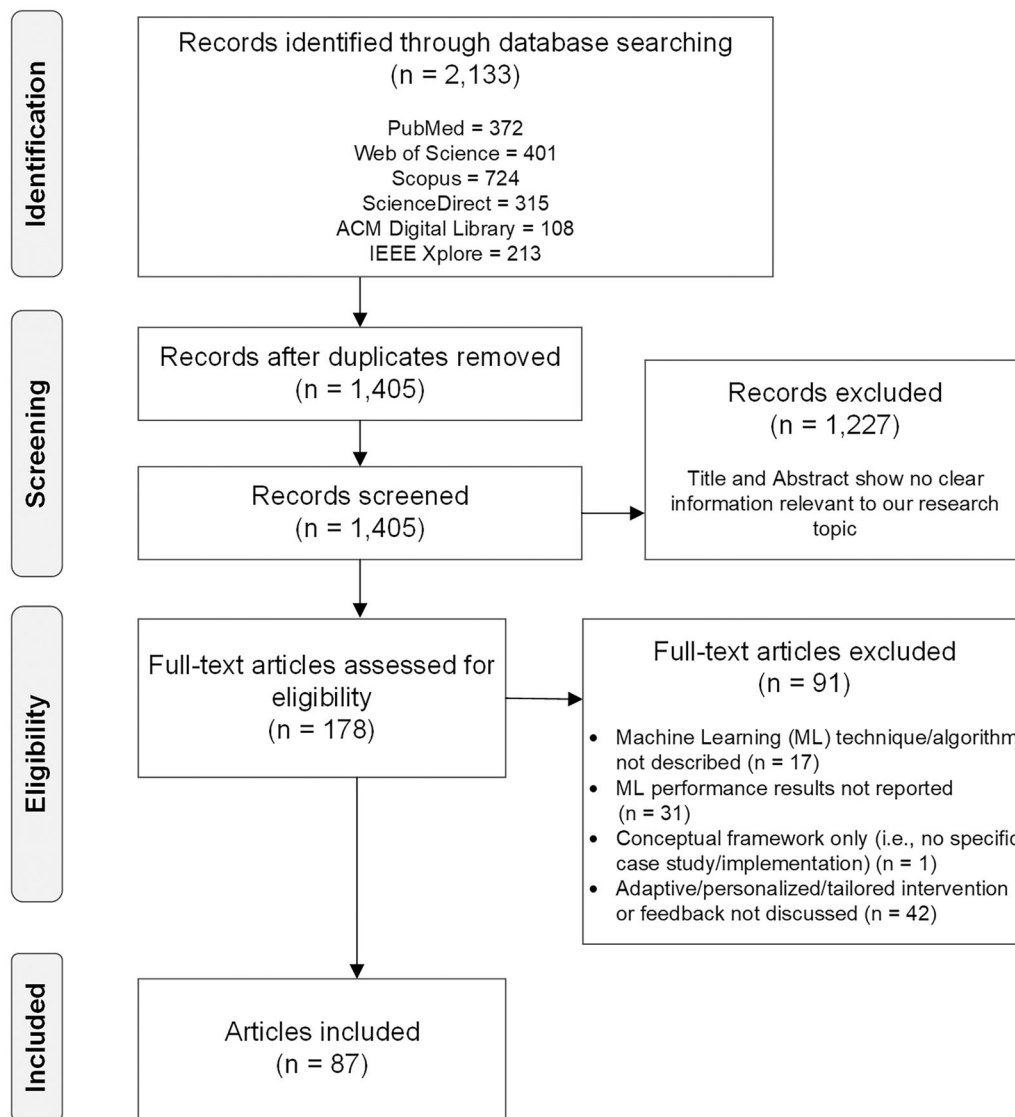
### 2.2. Paper selection

After retrieving initial set of papers, we first removed duplicates and then manually inspected the remaining papers to evaluate them based on our inclusion/exclusion criteria. First, based on independent screening of both title and abstract, 1227 papers were excluded since they discuss topics unrelated to our research objective, such as ML research outside the focus of health and well-being. Next, we screened the full texts of 178 papers and excluded 91 papers because (1) the ML methods were neither mentioned nor discussed; (2) the ML model evaluation results were not reported; (3) only architectural framework of the proposed ML-based system was reported without implementation details; (4) information regarding how ML model output would inform adaptive or personalized intervention/treatment were missing. Finally, we included the remaining 87 papers in the systematic review.

### 2.3. Information extraction

We inspected full texts of the selected papers and extracted key information (see below) which in turn were coded in a coding worksheet that was developed as a collaboration between two researchers. For ease of presentation, we grouped the information extracted into the following categories:

1. *Target Domain*: For each article, we identified the target domain of the work. In case the domain is not explicitly highlighted, we examined the use cases for the system and target audience to infer the target domain.



**Figure 1.** Procedure for Articles Selection based on PRISMA approach.

2. *Data Collection:* Collecting required data is an appropriate step prior to building ML models. Reviewed papers use data of various types and from diverse sources to train their ML models and validate for model generalization. Hence, in this category, we summarize the data sources, types of data, data modalities, and data fusion methods employed by the papers.
3. *Model Development Process:* In this category, we summarized the strategies employed in the development of the ML models including data preprocessing/feature engineering, model training and validation, as well as model testing.
4. *ML Techniques:* We summarized the ML algorithms used for modeling. These include classical ML algorithms, ensemble techniques, and advanced techniques such as deep learning.
5. *Model Evaluation:* This category summarizes the metrics or criteria used for evaluating model performance. Some of the metrics include Accuracy, Precision, Recall, F1-score, AUC (Area Under the ROC Curve), and

RMSE (Root Mean Square Error). We also reported the performance of the models using these metrics.

6. *Adaptive or Personalization Strategies:* Based on the ML models, we summarized the strategies employed in reviewed works to adapt interventions or treatments to individual patients/users.
7. *Limitations and Recommendations:* To inspire future work and advancement in the area of ML-driven adaptive systems for health and wellness domains, we documented and summarized the limitations or challenges reported in the reviewed papers, and offer recommendations on how they could be addressed.

### 3. Results

Of the 87 reviewed papers, 29 were *conference* publications while 58 were *journal* papers, as shown in [Figure 2](#). Majority of the publications (77%) were published between 2018 and 2021.

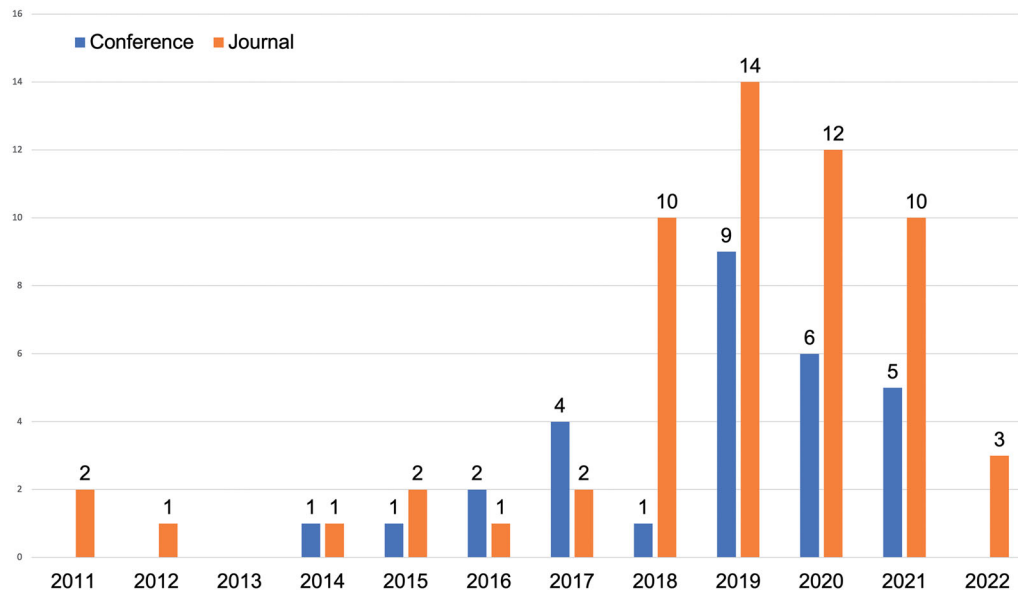


Figure 2. Number of papers per year grouped by publication channel.

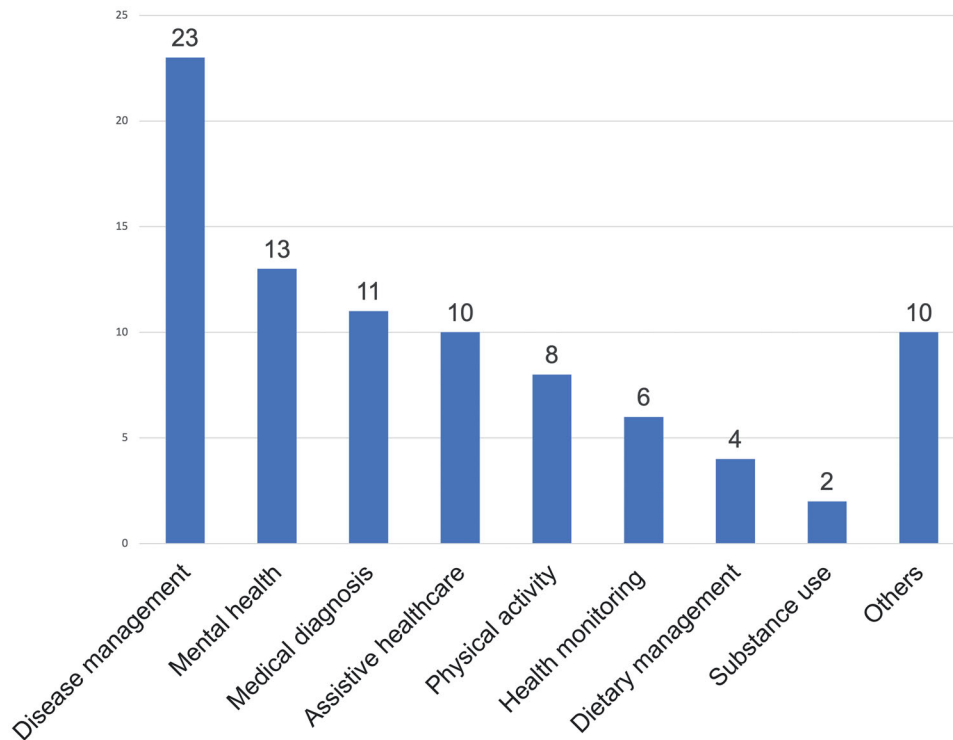


Figure 3. Target health domains and the corresponding number of papers.

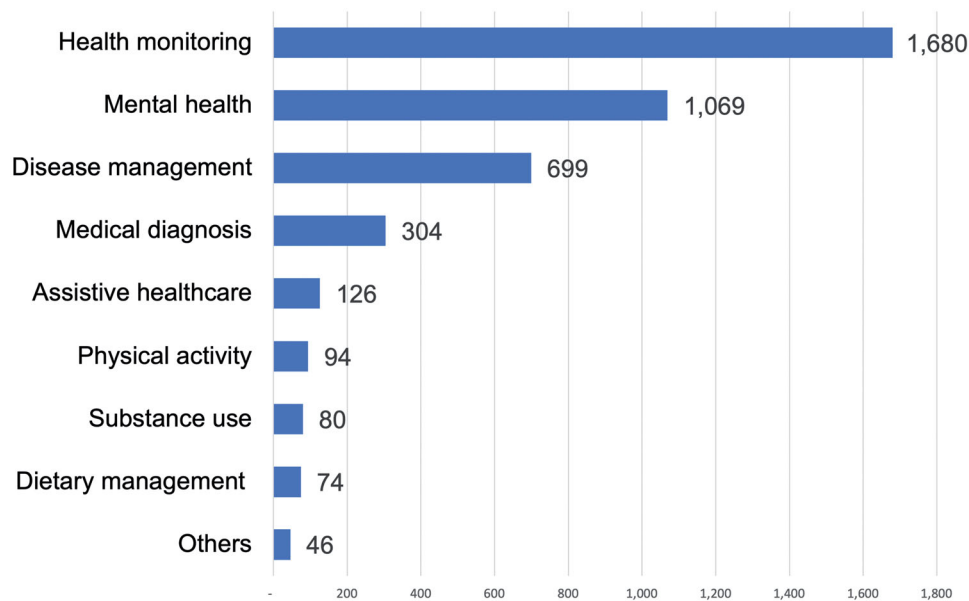
### 3.1. Target domains

This section describes the domains targeted by the papers for the purpose of adapting health and wellness interventions.

The reviewed papers can be broadly grouped into nine (9) categories or domains, as shown in Figure 3. 23 papers (26%) focus on treatments/interventions to support *management of diseases or conditions*, such as diabetes ( $n=5$ ), hypertension ( $n=3$ ), Parkinson's disease ( $n=2$ ), sickle cell ( $n=2$ ), obesity ( $n=3$ ), cardiovascular disease ( $n=1$ ),

epilepsy ( $n=1$ ), heart conditions ( $n=2$ ), osteoarthritis ( $n=1$ ), prostate cancer ( $n=1$ ), asthma ( $n=1$ ), as well as dyslexia, dysgraphia, and dyscalculia ( $n=1$ ). Furthermore, 10 papers (11%) focus on providing *assistive care* that support personal mobility, cognitive functions, and common daily activities. This domain includes human fall detection ( $n=4$ ), human activity recognition ( $n=4$ ), cognitive impairment ( $n=1$ ), and visual impairment ( $n=1$ ). Moreover, 11 papers (13%) target *detection or diagnosis of medical conditions*, with majority of these papers focusing on multiple conditions ( $n=8$ ) while others target specific





**Figure 4.** Citations by category.

conditions (such as cancer (Cheerla & Gevaert, 2017), fetal health (Akbulut et al., 2018), and COVID-19 (Barbosa et al., 2021)). 13 papers (15%) focus on *mental health* issues such as stress ( $n=6$ ), depression ( $n=5$ ), anxiety ( $n=1$ ), and mental wellness ( $n=1$ ). 8 papers (9%) target *physical activity*, while 4 papers (5%) target *dietary management* such as recipes recommendation (Gorbonos et al., 2018), automated food tracking (Rachakonda et al., 2020), and assessment of eating behaviour (Spanakis et al., 2017) and diet (Xu et al., 2019).

Furthermore, six papers (7%) focus on *monitoring body functions* (such as blood pressure (P. Chiang & Dey, 2019; P.-H. Chiang et al., 2021), blood glucose (W. Gu, 2017), postprandial glycemic response (Zeevi et al., 2015)) and *health behaviours* (Kang, 2021; Nosakhare & Picard, 2020) to determine health status and offer interventions. Two papers (2%) target *substance use* such as opioid (Scherzer et al., 2020) and alcohol (Bae et al., 2018). The remaining 10 papers (11%) target *smoking cessation* (T. Chen et al., 2018), *homeopathy remedy finding* (Priyadarshi & Saha, 2020), *patients' location privacy preservation* during healthcare delivery (Abdo et al., 2020), *mHealth apps recommendation* based on health condition (Edara et al., 2020), *representing/modelling healthcare provider's knowledge* (Tuti et al., 2020), *dental and oral health* (Stark & Samarah, 2019), *health awareness* (Ganju et al., 2021), *disease surveillance* (Y. Zhang et al., 2022), *medication management* (Hu et al., 2021), and *medical specialty recommendation* (Lee et al., 2021).

### 3.2. Impact analysis of reviewed papers

We conducted an impact analysis of the reviewed papers based on citation count. Specifically, Figure 4 shows the number of citations per category, with papers in the *health monitoring* category emerging as the most cited ( $n=1680$ ) followed by *mental health* ( $n=1069$ ) and *disease management* categories ( $n=699$ ). Figure 5 shows the top three

papers in terms of the number of citations for each category. Zeevi et al. (2015) which belongs to the *health monitoring* category is the most cited paper overall ( $n=1628$ ), followed by Burns et al. (2011) in the *mental health* category ( $n=635$ ). Rabbi et al. (2015) in the *disease management* category is the third most cited paper overall with 200 citations.

### 3.3. Data sources and data types for ML in adaptive health and wellness systems

We identified eight sources from which data required for ML model training were generated (see Figure 6). Majority of the papers ( $n=50$ , 57%) reported using one or more *sensors* for data collection. The most common are smartphone sensors ( $n=28$ ), followed by wearable sensors ( $n=23$ ), as well as non-wearable sensors ( $n=9$ ) usually found in smart homes (such as mounted cameras, presence sensors, luminosity sensors, beacons, thermometers, etc.).

Furthermore, 12 papers (14%) collected data from target audience using *questionnaires*. Also, *public datasets* (see Supplementary Appendix B) or those that require special permissions to acquire (*private datasets*) were utilized for model training by 20 papers (23%). *Electronic Health Records* (EHRs), also called Electronic Medical Records (EMRs), containing patients' clinical history were utilized by 10 papers (11%). On the other hand, *medical devices* for measuring vital signs (e.g., blood glucose monitors, blood pressure monitors, etc.) were utilized by seven papers (8%), while seven papers (8%) reported the use of automated web search, web crawlers, or social media APIs to retrieve health-related data (e.g., patient reviews (Edara et al., 2020; Zingg et al., 2021), health-related tweets (Y. Zhang et al., 2022), food images (Sowah et al., 2020), homeopathy treatments (Priyadarshi & Saha, 2020), etc.). *Phone/interaction logs* which include call and SMS logs, as well as app usage data, are other sources used in seven papers (8%), while one

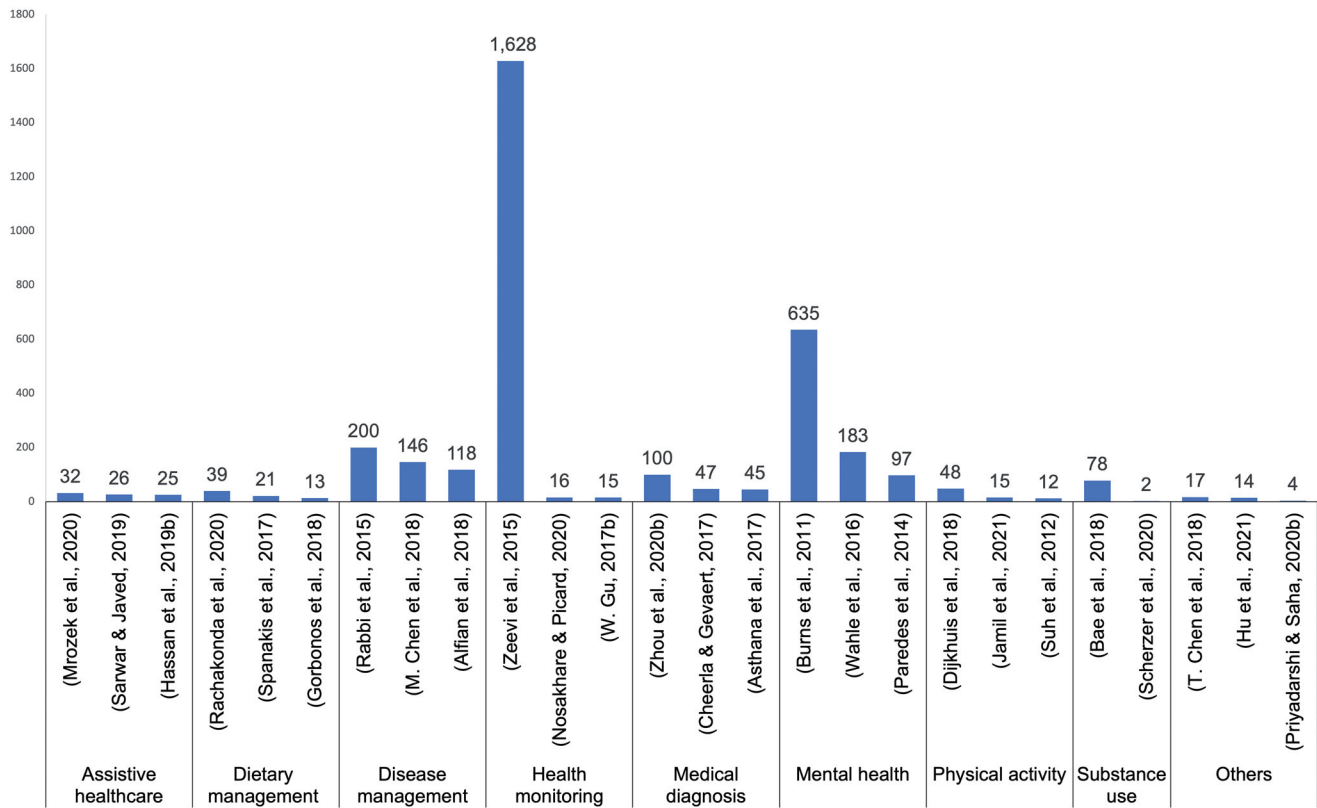


Figure 5. Top three cited papers in each category.

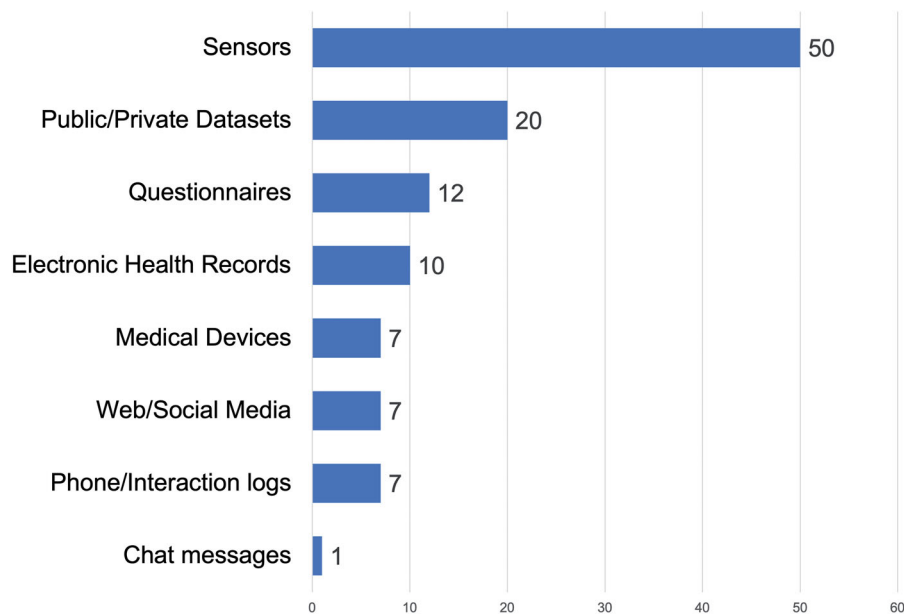
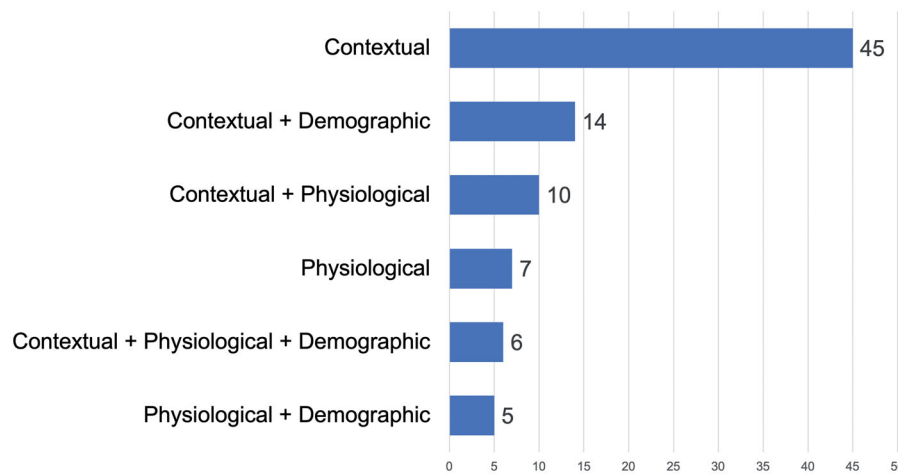


Figure 6. Data sources/modalities utilized by the reviewed papers.

paper (1%) collected *text messages* from a chat-enabled app (Scherzer et al., 2020).

With regard to the types of data collected, **contextual data** which uncover relevant information about patients' values, environment, or behaviours influencing their health and well-being (Holt et al., 2020) are the most common ( $n = 75$ , 86%). Examples include *motion/activity data* (P. Chiang & Dey, 2019; Dijkhuis et al., 2018; Kadri et al., 2020; Z. Li et al., 2019; Sarwar & Javed, 2019; Stamate et al., 2017), *location*

(Alharthi et al., 2019; Ghandeharioun et al., 2019; Paredes et al., 2014; Suh et al., 2012; Wahle et al., 2016), *sleep patterns* (Nosakhare & Picard, 2020; Pelle et al., 2019; Yates & Islam, 2019), *medication intake* (da Silva et al., 2019; Sarwar & Javed, 2019), *dietary/nutrition data* (Forman, Goldstein, Zhang, et al., 2019; Sookrah et al., 2019; Sowah et al., 2020; Xu et al., 2019), *audio sequences* (Barata et al., 2016; Kariyawasam et al., 2019), *images* (Chin et al., 2020; Rabbi et al., 2015; Rachakonda et al., 2020), *emotion/mood log*



**Figure 7.** Types of data utilized by the reviewed papers for training ML models.

(Burns et al., 2011), *clinical history* including medical notes and treatments (M. Chen et al., 2018; Khumrin et al., 2018; Koren et al., 2019; Lopez-Guede et al., 2015), etc. Moreover, 28 papers (32%) reported the use of **physiological data** such as *vital signs* (e.g., body temperature, pulse rate, blood pressure, blood glucose/sugar level, breathing rate, heart rate, etc.) (Abdo et al., 2020; Aujla et al., 2019; Kesavan & Arumugam, 2020; Lopez-Guede et al., 2015), *physiological signals* (e.g., electroencephalogram – EEG (Alharthi et al., 2019; El Barachi et al., 2019; Kesavan & Arumugam, 2020), electrocardiogram – ECG (Delmastro et al., 2020; Hermens et al., 2014; Kesavan & Arumugam, 2020), electromyogram – EMG (Hermens et al., 2014), and electrodermal activity – EDA (Delmastro et al., 2020)), etc. Lastly, 25 papers (29%) reported the use of **demographic data** or user profile such as age, gender, height, weight, etc. (Alfian et al., 2018; Kadri et al., 2020; Khalaf et al., 2016; Neloy et al., 2019; Suh et al., 2012). Figure 7 shows the distribution of reviewed papers based on the data type(s) utilized. The figure reveals that 40% of the papers ( $n=35$ ) combined multiple data types to form their ground truth dataset for model training and evaluation.

### 3.4. Multimodality and data fusion methods

Multimodality refers to an approach in which data is obtained or acquired from more than one modality or source (Lahat et al., 2015). Each modality produces a dataset; hence, a multimodal scenario will more likely generate heterogeneous datasets that need to be combined or *fused*. Data fusion methods have been broadly classified into three levels: *data-level*, *feature-level*, and *decision-level* (J. Zhang, 2010). In data-level fusion, raw data from multiple modalities are combined prior to feature engineering (i.e., the process of extracting features from raw data). Feature-level fusion involves computing or extracting features from each modality and then combining the features prior to model training. In decision-level fusion, features from individual modalities are fed into multiple ML models and then the outputs (decisions) are fused to yield a final output (D'mello & Kory, 2015; Y. Gu et al., 2015).

Based on our findings, 35 papers (40%) reported the use of multimodal data, such as acquisition of data streams from multiple sensors (Delmastro et al., 2020; Javed et al., 2021; Kajiwaru & Kimura, 2019; Kesavan & Arumugam, 2020; Luštrek et al., 2021; Mrozek et al., 2020; Sarwar & Javed, 2019). Other papers used sensor data in conjunction with questionnaires (T. Chen et al., 2018; Ghandeharioun et al., 2019; Paredes et al., 2014), public datasets (Asthana et al., 2017), clinical history from EHR/EMR (Asthana et al., 2017; M. Chen et al., 2018; Lopez-Guede et al., 2015), images (Sowah et al., 2020), vital signs from medical devices (P. Chiang & Dey, 2019; W. Gu, 2017), and call/SMS logs (Burns et al., 2011). In addition, one paper utilized public dataset containing multimodal data including ECG signals, Arterial Blood Pressure (ABP), heart rate, and other vital signs (Popp et al., 2019). Also, another paper combined clinical data obtained from medical devices (such as continuous glucose monitors) and self-reports (via questionnaires) (Zeevi et al., 2015).

Regarding data fusion methods employed by the 35 papers, feature-level fusion emerged as the most utilized ( $n=32$ ), followed by data-level fusion ( $n=3$ ). No paper employed the decision-level fusion method.

### 3.5. Data preprocessing/feature extraction techniques

The reviewed papers utilized various preprocessing techniques to clean or prepare the raw data. These techniques include *data imputation* (Barbosa et al., 2021; Cheerla & Gevaert, 2017; M. Chen et al., 2018; P. Chiang & Dey, 2019; Forman, Goldstein, Zhang, et al., 2019; Jamil et al., 2021; Z. Li et al., 2019), *data transformation or conversion* (Alharthi et al., 2019; Barbosa et al., 2021; M. Chen et al., 2018; da Silva et al., 2019; Dijkhuis et al., 2018; Ghandeharioun et al., 2019; Neloy et al., 2019), *removal of noisy or irrelevant samples* (M. Chen et al., 2018; Lee et al., 2021; Z. Li et al., 2019; Nosakhare & Picard, 2020; Sarwar & Javed, 2019; Sowah et al., 2020), *data normalization or standardization* (Afreen et al., 2019; Cheerla & Gevaert, 2017; M. Chen et al., 2018; Coutinho et al., 2021; Dijkhuis et al., 2018; Ghandeharioun et al., 2019; Hu et al., 2021; Jamil et al., 2021; Javed et al.,



2021; Neloy et al., 2019; Nosakhare & Picard, 2020; Sansrimahachai, 2020a; Stamate et al., 2017), *removal of duplicates* (M. Chen et al., 2018; da Silva et al., 2019; Jamil et al., 2021; Lee et al., 2021; Sowah et al., 2020), *removal of missing values* (P. Chiang & Dey, 2019; Dijkhuis et al., 2018; Lee et al., 2021), *data aggregation (derived attributes)* (Alfian et al., 2018; Dijkhuis et al., 2018; Sansrimahachai, 2020a; Xu et al., 2019), *data augmentation* (Hu et al., 2021), *handling class imbalance* (Forman, Goldstein, Zhang, et al., 2019; Mishra et al., 2021; Y. Zhang et al., 2022), and *spelling error correction* (Jamil et al., 2021; Priyadarshi & Saha, 2020). Some papers also employed the Natural Language Toolkit (NLTK) to convert words to their root form (stemming), remove punctuations, non-English characters, stopwords, and words less than two characters (Edara et al., 2020; Y. Zhang et al., 2022; Zingg et al., 2021).

To extract the features needed to train ML models from raw data, the reviewed papers employed various techniques. For time-series data (such as sensor data), Pathinarupothi et al. (2018) used the *sliding window* technique to look for features at different time segments (5 min, 10 min, 15 min, etc.). Wahle et al. applied the same technique to extract activity-related features using a time window of 2 min from accelerometer data (Wahle et al., 2016), similar to Luštrek et al. (2021). Other papers that used sliding windows approach for feature extraction include (Alharthi et al., 2019; Bae et al., 2018; Delmastro et al., 2020; Elvitigala et al., 2021; Ghandeharioun et al., 2019; Hassan et al., 2019; Lopez-Guede et al., 2015; Mrozek et al., 2020; Prabhu et al., 2018; Sarwar & Javed, 2019; Tuti et al., 2020; Van Woensel et al., 2020; Yates & Islam, 2019). On the other hand, Kesavan & Arumugam (2020) and Stamate et al. (2018) used *deep learning* approaches to extract features from sensor readings. Delmastro et al. used *MATLAB* to extract features from heart rate (HR) and heart rate variability (HRV) time-series data, while *Ledalab* (a MATLAB-based software) was used to extract features from EDA data (Delmastro et al., 2020). *Thresholding* is another feature extraction technique reported in the reviewed literature. For example, Li et al. applied two-fold thresholding such that step counts above the upper threshold were spike features and those below the lower threshold were sedentary features (Z. Li et al., 2019). Also, Cheerla et al. extracted miRNA features within the threshold value of 0.2 (Cheerla & Gevaert, 2017).

Furthermore, features were extracted from images using a *segmentation* approach that involves grouping image pixels to better identify objects in the image (Rachakonda et al., 2020; Xu et al., 2019) or using tools such as OpenPose (Kajiwaru & Kimura, 2019). To extract features from text, Asthana et al. (2017) and Priyadarshi and Saha (2020) used the TF-IDF (Term Frequency – Inverse Document Frequency) which is a *Bag-of-Words* approach for vectorizing texts using a weighting scheme. Koren et al. applied the *Word Embedding* technique (Word2Vec) to identify similar terms (features) and TF-IDF to find features that are highly linked to specific health conditions (Koren et al., 2019). In addition to TF-IDF, Zhou et al. extracted sentence-level features using the *Convolutional Neural Network* (CNN) deep

learning technique (Zhou et al., 2020). Also, Chen et al. applied CNN to extract features from texts and images (M. Chen et al., 2018). In lieu of TF-IDF, Nosakhare et al. utilized One-Hot encoding to vectorize texts (Nosakhare & Picard, 2020).

Prior to model construction, some of the reviewed papers ( $n = 14$ ) applied feature selection techniques to reduce the number of input features by selecting important features that improve model performance while reducing computational cost. Filter methods apply ranking and statistical techniques to score features and then remove features below a threshold (Chandrashekar & Sahin, 2014; Visalakshi & Radha, 2015). Examples of filter methods found in the reviewed literature include *correlation coefficients* (Bae et al., 2018; Cheerla & Gevaert, 2017; Delmastro et al., 2020; Neloy et al., 2019; Stark & Samarah, 2019), *mutual information* (or information gain) (Bae et al., 2018; Delmastro et al., 2020; Priyadarshi & Saha, 2020), *Chi-square test* (Priyadarshi & Saha, 2020), and *variance threshold* (Stark & Samarah, 2019). Wrapper methods evaluate possible combinations of features to find feature subsets that improve the predictive accuracy of the ML algorithms under consideration (Visalakshi & Radha, 2015). An example is the *recursive feature elimination algorithm* used by Cheerla et al. to pick a strong feature subset for cancer diagnosis (Cheerla & Gevaert, 2017). Furthermore, Embedding methods incorporate feature selection as part of the model training process such that features that contribute significantly to prediction performance are selected (Jović et al., 2015; Lee et al., 2021). Akbulut et al. and Chiang et al. used *Tree-based estimators* (Akbulut et al., 2018) and *Random Forest* (P.-H. Chiang et al., 2021; P. Chiang & Dey, 2019), respectively, to determine feature importance and select dominant features. Other techniques include dimensionality reduction using *Principal Component Analysis* (PCA) (Delmastro et al., 2020; Rachakonda et al., 2020; Stark & Samarah, 2019) and *Singular Value Decomposition* (SVD) (Aujla et al., 2019; Zhou et al., 2020) to reduce number of input features, as well as *permutation* method for computing feature importance (Auffenberg et al., 2019) and *constrained vocabulary size* approach for selecting top features in text classification problems (e.g., setting maximum number of features based on word/term frequency) (W. Yang et al., 2020).

### 3.6. Machine learning techniques

Next, we discuss the ML techniques utilized in the literature to adapt health and wellness applications and approaches used to evaluate the models built. We first grouped the papers into four main ML paradigms: (i) supervised learning, (ii) unsupervised learning, (iii) semi-supervised learning, and (iv) reinforcement learning.

#### 3.6.1. Supervised learning

In supervised learning, labelled data is used to train a model which then predicts the label (or class) for new data (Liu, 2011). In other words, the training data contain both the

inputs and the expected outputs; hence the model can learn to predict the actual outputs of new (unseen) inputs. The vast majority of the reviewed papers ( $n = 82$ , 94%) reported the use of supervised learning technique for *classification* and *regression* tasks. For classification tasks, models are used to predict a binary class (such as occurrence or absence of a fall (Mrozek et al., 2020)) or multiclass (e.g., at-risk health conditions (Asthana et al., 2017)); however, for regression tasks, models predict a continuous value (e.g., blood pressure (P. Chiang & Dey, 2019)). To improve readability, we further categorized the supervised learning methods into *classical ML*, *ensemble learning*, and *deep learning*.

Based on our findings, 41 out of the 82 papers (50%) utilized **classical ML** techniques which are based on conventional machine learning algorithms, such as *Decision Tree (DT)* (Abdo et al., 2020; Burns et al., 2011; da Silva et al., 2019; Kadri et al., 2020; Lopez-Guede et al., 2015; Martín et al., 2011; Spanakis et al., 2017; Suh et al., 2012; Van Woensel et al., 2020; Yates & Islam, 2019), *Support Vector Machine (SVM)* (Barata et al., 2016; Cheerla & Gevaert, 2017; El Barachi et al., 2019; Kariyawasam et al., 2019; Pathinarupothi et al., 2018; Prabhu et al., 2018; Priyadarshi & Saha, 2020), *Support Vector Regression (SVR)* (Coutinho et al., 2021), *Logistic Regression (LogisReg)* (Javed et al., 2021; Neloy et al., 2019), *Naïve Bayes (NB)* (Alharthi et al., 2019; Khumrin et al., 2018; Sarwar & Javed, 2019), *Linear Regression (LR)* (Xu et al., 2019; Y. Zhang et al., 2022), *linear-discriminant analysis (LDA)* (Elvitigala et al., 2021), *K-Nearest Neighbors (KNN)* (Kariyawasam et al., 2019), *Hoeffding Tree (HT)* (Javed et al., 2021), *Multiple Regression (MR)* (Hermens et al., 2014), *supervised latent Dirichlet allocation (sLDA)* (Nosakhare & Picard, 2020), and *Artificial Neural Network (ANN)* including *Multilayer Perceptron (MLP)* (Afreen et al., 2019; Alfian et al., 2018; Gorboson et al., 2018; Khalaf et al., 2016; Sansrimahachai, 2020b; Sookrah et al., 2019; Stamate et al., 2017) and the *Levenberg-Marquardt algorithm* in MATLAB (Kang, 2021).

**Ensemble learning** uses multiple ML algorithms or models to achieve better predictive performance, and this technique was used by 30 papers (37%). *Random Forest (RF)* or *Decision Forest (DF)* which refers to an ensemble of many decision trees and an extension of the bagging method (Breiman, 2001) is the most utilized in most of the papers (Akbulut et al., 2018; Auffenberg et al., 2019; Bae et al., 2018; Delmastro et al., 2020; Dijkhuis et al., 2018; Ghandeharioun et al., 2019; Kajiwaru & Kimura, 2019; Z. Li et al., 2019; Neloy et al., 2019; Paredes et al., 2014; Stark & Samarah, 2019; Wahle et al., 2016), followed by boosting techniques such as *Gradient Boosting (GB)* (Popp et al., 2019; Zeevi et al., 2015), *LogitBoost* (Abd et al., 2017), *eXtreme Gradient Boosting (XGBoost)* (Afreen et al., 2019), *Adaptive Boosting (AdaBoost)* (Delmastro et al., 2020; Ghandeharioun et al., 2019), and *Boosted Decision Tree (BDT)* (Mrozek et al., 2020). Other ensemble learning methods include *Bagging SVM* (Neloy et al., 2019), and a combination of other classical or ensemble techniques such as *Bayesian Network (Bayes Net)+XGBoost+LogisReg+ANN* (Koren et al., 2019), *DT+SVM+ANN* (M. Chen et al.,

2018), as well as *LogitBoost+RF+Random-subspace+Bayes Net* (Forman, Goldstein, Zhang, et al., 2019).

**Deep learning** approaches use neural networks to extract features from raw input and learn representations with thousands (up to billions (Rasley et al., 2020)) of parameters to obtain the output. Twenty papers (24%) reported the use of deep learning techniques such as *Long Short-Term Memory (LSTM)* (Alfian et al., 2018; T. Chen et al., 2018; Lee et al., 2021; Mahyari & Pirolli, 2021; Tuti et al., 2020; J. Yang et al., 2020), *Convolutional Neural Network (CNN)* (Aujla et al., 2019; Kariyawasam et al., 2019; Sowah et al., 2020), *Adaptive Deep CNN* (Kesavan & Arumugam, 2020), *MobileNet* (Rachakonda et al., 2020), *Modified SlowFast Network* (Chin et al., 2020), *Multi-task Recurrent Neural Network (MT-RNN)* (W. Gu, 2017), *Bidirectional LSTM (BiLSTM)* (Kadri et al., 2020), and *Hybrid Latent Variable Hierarchical Recurrent Encoder-Decoder (Hybrid-VHRED)* (W. Yang et al., 2020). Different deep learning techniques were also combined to create new architectures such as the *Recurrent Convolutional Neural Network (RCNN)* (Stamate et al., 2018) and *Convolutional Recurrent Neural Network (CRNN)* (Zhou et al., 2020) which combine CNN and RNN. Others include *CNN-LSTM* (a fusion of CNN and LSTM) (Hassan et al., 2019), *DNN-LSTM* which combines Deep Neural Network (DNN) and LSTM (Edara et al., 2020).

### 3.6.2. Unsupervised learning

Unlike supervised learning, unsupervised learning algorithms identify patterns within input data without involving target output labels (Alloghani et al., 2020). Based on the patterns learned, clustering is then performed to group (or classify) similar data samples. Only six papers (9%) reported the use of unsupervised learning algorithms such as *k-Means clustering* (Javed et al., 2021; Lopez-Guede et al., 2015; Zhou et al., 2020) and *Hierarchical clustering* (Z. Li et al., 2019; Spanakis et al., 2017; Van Woensel et al., 2020). Besides using clustering to estimate which activities are performed at a specific point based on user's context (e.g., location) (Javed et al., 2021), most papers used clustering as an initial step prior to supervised learning. For example, Z. Li et al. (2019) and Spanakis et al. (2017) applied clustering to first create user groups based on certain contextual features (such as eating behaviour or daily activities), and then trained supervised ML models using the data in each group or cluster to predict target labels. Other unsupervised learning techniques reported in a few papers include the *Local Outlier Factor (LOF)* algorithm which detects anomalies by comparing the density of data instances around a given instance with the density around its neighbors (Lopez-Guede et al., 2015), as well as the *Non-negative Matrix Factorization (NMF)* technique used by Gorboson et al. (2018) for modelling the interactions between food ingredients and their proportions. Rajasekaran et al. applied the Deep Continuous Deep Belief Network (DC-DBN) to analyze and classify video feeds – collected during physical health monitoring of elderly people – in an unsupervised fashion (Rajasekaran & Kousalya, 2022).

### 3.6.3. Reinforcement learning

Reinforcement learning (RL) is about mapping situations to actions to maximize a reward (Sutton & Barto, 2018). An RL agent not only exploits past actions that led to a reward but also explores new actions that could yield better reward. One of the common RL algorithms is the *Multi-armed Bandit (MAB)* and interestingly the only RL algorithm found in our review. Two papers applied MAB to (i) dynamically learn and influence user behaviours by suggesting actions (such as walking to work or visiting the gym) that maximize the chance of achieving calorie loss (Rabbi et al., 2015), and (ii) propose challenging and achievable goals relevant for the treatment of osteoarthritis (Pelle et al., 2019). Wang et al. formulated an RL model as a Markov decision process in which the agent represents a mobile health system that learns the optimal strategy to interact with a target user (S. Wang et al., 2021). The reward indicates the physical activity performed after the users act on a reminder sent by the RL model.

### 3.6.4. Semi-supervised learning

As regards semi-supervised learning which combines labelled and unlabelled data to train a model, two papers (2%) reported using this technique for retraining cancer diagnosis and treatment recommendation models (Cheerla & Gevaert, 2017), as well as for anomaly detection in patients' vitals (Arpaia et al., 2022).

## 3.7. Evaluation of machine learning models

As previously mentioned, supervised learning which involves classification and regression tasks using labelled data is the most prominent in our review. Supervised ML models are usually evaluated by measuring how well they perform in predicting the correct labels for test or "holdout" data (which are completely different from the training data). The model evaluation techniques applied in the reviewed papers include *k-fold cross-validation* (Abdo et al., 2020; Akbulut et al., 2018; Alfian et al., 2018; Bae et al., 2018; Burns et al., 2011; Cheerla & Gevaert, 2017; P. Chiang & Dey, 2019; da Silva et al., 2019; Delmastro et al., 2020; Dijkhuis et al., 2018; El Barachi et al., 2019; Ghandeharioun et al., 2019; Javed et al., 2021; Kajiwarra & Kimura, 2019; Lopez-Guede et al., 2015; Mrozek et al., 2020; Neloy et al., 2019; Nosakhare & Picard, 2020; Pathinarupothi et al., 2018; Prabhu et al., 2018; Priyadarshi & Saha, 2020; Sarwar & Javed, 2019; Van Woensel et al., 2020) and *leave-one-fold-out cross-validation* (Forman, Goldstein, Zhang, et al., 2019; Javed et al., 2021; Stamate et al., 2017, 2018; Wahle et al., 2016; Zeevi et al., 2015), *train/test split* (e.g., 70/30 split which means 30% of the test data is used to evaluate the model) (Afreen et al., 2019; Akbulut et al., 2018; Auffenberg et al., 2019; Burns et al., 2011; P. Chiang & Dey, 2019; Dijkhuis et al., 2018; El Barachi et al., 2019; W. Gu, 2017; Rachakonda et al., 2020; Sansrimahachai, 2020b; Scherzer et al., 2020; Stark & Samarah, 2019; Tuti et al., 2020; J. Yang et al., 2020), and *train/test/validation split* (e.g., 70/20/10 split which involves performing additional evaluation using

10% of the data, called "validation set," to assess the generalization ability of the model) (Bae et al., 2018; Ghandeharioun et al., 2019; Neloy et al., 2019; Sowah et al., 2020).

Furthermore, standard metrics for evaluating the performance of classification and regression models were employed in the literature. For classification tasks, majority of the papers ( $n=53$ , 61%) reported measure of *accuracy* (acc) which is the ratio of the number of input samples correctly classified to the total number of samples, while 37% ( $n=32$ ) reported either *precision* or *recall* or both. Precision is a measure of a model's usefulness (low false positive rate), while recall measures a model's completeness (low false negative rate) (Thieme et al., 2020). Recall (also called *sensitivity* or *true positive rate* – TPR) can also be described as the proportion of actual positive samples that are correctly predicted as positive; hence a low false negative rate implies a high recall. Moreover, 24 papers (28%) reported *F1-score* or *F-measure* which is the harmonic mean of precision and recall, and commonly viewed as the preferred evaluation metric for classifiers. Another metric – *specificity* which is the proportion of actual negative samples that are correctly predicted as negative – was reported (Barbosa et al., 2021; Cheerla & Gevaert, 2017; Forman, Goldstein, Crochiere, et al., 2019; Forman, Goldstein, Zhang, et al., 2019; Kesavan & Arumugam, 2020; Pelle et al., 2019; Wahle et al., 2016). Other metrics include *AUC* or *AUROC* (i.e., "Area Under the ROC Curve" which is a measure or degree of separability of distinct classes) (Abd et al., 2017; Akbulut et al., 2018; Arpaia et al., 2022; Auffenberg et al., 2019; Bae et al., 2018; Javed et al., 2021; Z. Li et al., 2019; Stamate et al., 2018; Tuti et al., 2020), *AUPRC* ("Area Under the Precision-Recall Curve" which is more appropriate for imbalanced data, compared to AUC) (Delmastro et al., 2020), *false positive rate/ratio* (FPR) which is the proportion of negative samples incorrectly identified as positive) (Rajasekaran & Kousalya, 2022), and *test loss* (Chin et al., 2020).

As regards regression tasks, the common evaluation metrics include *correlation coefficient* ( $r$ ) which measures the strength of the linear relationship or dependence between two variables (i.e., predicted and actual values) (Alfian et al., 2018; Kang, 2021; Popp et al., 2019; Zeevi et al., 2015), as well as the error metrics for measuring the proximity of predicted values to actual values such as the *Root Mean Square Error* (RMSE) (Afreen et al., 2019; Alfian et al., 2018; Asthana et al., 2017; Aujla et al., 2019; P. Chiang & Dey, 2019; P.-H. Chiang et al., 2021; Coutinho et al., 2021; Khalaf et al., 2016), *Mean Absolute Error* (MAE) (P. Chiang & Dey, 2019; Kang, 2021; Khalaf et al., 2016; Sansrimahachai, 2020b), *Mean Squared Error* (MSE) (Kang, 2021; Khalaf et al., 2016; Sansrimahachai, 2020b), and *Mean Absolute Percentage Error* (MAPE) (Aujla et al., 2019; P. Chiang & Dey, 2019; Khalaf et al., 2016). Other regression metrics are *R-squared* ( $r^2$ ) which is a measure of the variation in the dependent variable (i.e., target or output variable) explained by the model (Afreen et al., 2019; Ganju et al., 2021; Y. Zhang et al., 2022) and *Explained Variance score* (EVar)



which is similar to  $r^2$  but also incorporates mean error (Afreen et al., 2019).

Finally, a few papers used less common metrics to evaluate model performance such as the *BLEU score* (i.e., “Bi-Lingual Evaluation Understudy” score which is a measure of how model-generated response is similar to the human-written response) (W. Yang et al., 2020), *Cohen-d* (Rabbi et al., 2015), and *mean quadratic error (MQE)* (Martín et al., 2011).

### 3.8. Applying ML to adapt health and wellness systems or interventions

Across all papers, the end goal is to personalize, tailor, or adapt health and wellness systems/interventions to individuals based on the output of ML models.

#### 3.8.1. Mental health

One of such adaptive systems is the *MOSS* (Mobile Sensing and Support) system (Wahle et al., 2016) which offers just-in-time cognitive behaviour therapy tailored to individual patients depending on their depression level, as predicted by two ML classifiers – SVM and RF with accuracies of 59.4% and 61.5% respectively. Likewise, Yates and Islam (2019) developed an adaptive mobile application called *Mindful* that predicts patients’ depression level using a high-performing DT model (acc = 96.6%), and then provides a proactive warning (in form of notification alerts) about the change in their mental health state followed by suggestions on lifestyle changes to improve well-being. *Mobilyze* application by Burns et al. (2011) also leveraged a DT model (acc: 60–91%) to predict the contextual states (e.g., location, mood/emotion, activities, social context, etc.) of depressive disorder patients, and based on the prediction, personalized interventions including lessons, tools, and graphical feedback on individual patients’ states are provided. In addition, a coach is notified via email to offer direct therapeutic support. To help college students cope with anxiety, Yang et al.’s mobile chatbot (*Mental Mentor*) employed the Hybrid-VHRED deep learning model (BLEU = 51.12%) to generate tailored responses based on students’ utterance and feelings/emotions (W. Yang et al., 2020). Each response includes interventions such as guided exercises and study tips. *POLYHYMNIA Mood* (Coutinho et al., 2021) is an adaptive web application that empowers people to use music effectively for coping with depression in everyday life. The application employed two ML models – SVR (RMSE = 0.047) and RF (RMSE = 0.110) – to estimate the emotions expressed by each track in the users’ music libraries. Based on the estimated emotion (valence and arousal), a personalized playlist comprising 14 tracks (about 45 min of music) was automatically generated to elevate individual users’ mood and reduce depressive symptoms. Similarly, the *Digilego* framework applied ML to analyze social media posts to inform digital features of a mobile intervention (MomMind) that promotes Peripartum Depression (PPD) prevention and self-management (Zingg et al., 2021). An RF model (average F1-score = 77.2%) was used to predict one of four

categories per post: *social support*, *symptom disclosure*, *medication*, *family and friends*, and *breastfeeding*. Based on the predicted category, appropriate digital features (e.g., medication list, instructions, reminders, action planning, pharmacological support, etc.) are delivered within MomMind using behaviour change techniques.

In the area of stress management, Delmastro et al. (2020) developed an adaptive mobile application that detects the stress level of older adults during a cognitive training session using RF and AdaBoost models with accuracy of 85.3% and 85.5% respectively, and then supports them in the definition of a personalized training activity that helps to reduce the stress level (if above expected limit) while improving cognitive ability. Similarly, Alharthi et al.’s context-aware acute stress prediction system predicts users’ stress state (*relaxed*, *normal*, or *stressed*) using the NB classifier (acc = 78.3%) and automatically triggers relief interventions (such as relaxation techniques) if the predicted state of a user is *stressed* (Alharthi et al., 2019). Paredes et al. developed a recommender system aimed at matching interventions to the personal traits of individual users and their temporal context using RF model (Paredes et al., 2014). The model predicts the expected stress reduction of each intervention for an individual in a given context, and based on these estimates, the recommender selects the best interventions that promote self-awareness of stress, lower depression-related symptoms, and knowledge of stress coping strategies for the user. The system’s effectiveness is 97% (i.e., percentage of interventions adopted by the users overall). Ghandeharioun et al., however, focused on improving people’s mental well-being via an ML-driven and emotion-aware personal assistant (Ghandeharioun et al., 2019). The agent predicts user’s emotional state (valence and arousal) using RF and AdaBoost regression models (acc: 82.4% and 67% respectively) and, based on the predicted state, suggests micro-interventions including individual or social short activities that fall into one of the following psychotherapy categories: positive psychology, cognitive behavioural, meta-cognitive, or somatic interventions. Moreover, Hermens et al. (2014) developed an ML-based coaching system which adapt a stress detection algorithm based on MR to individual users to achieve the most reliable estimation of stress ( $r^2 = 0.481$  for one of the users). Nosakhare and Picard (2020) predicted three categories of students’ well-being – stress, mood, and physical health – in their recommender system using the sLDA algorithm (F1-scores: *Stressed-Calm* = 72%, *Sad-Happy* = 68%, *Sick-Healthy* = 68%). Students receive personalized interventions (behavioural suggestions) that are both achievable and that might improve a future outcome, such as changing the bedtime, length of sleep, and social interactions planned for the following day. Furthermore, StressShoe – developed by Elvitigala et al. (2021) – estimates stress level by sensing behavioural changes based on sitting posture using sensor. The application delivers just-in-time interventions if the stress level estimated by a linear-discriminant analysis model (acc = 84.3%) exceeds a user-defined threshold. Mahyari et al. aimed to reduce perceived stress by recommending exercises that have a high probability of

being pursued and achieved for each individual on each day using two interconnected RNNs ( $\text{acc} = 80\%$ ) (Mahyari & Pirolli, 2021).

### 3.8.2. Disease management

To achieve personalized diabetes self-management, the *uHealthFit* application (Alfian et al., 2018) utilized two ML models – MLP ( $\text{acc} = 77\%$ ) and LSTM ( $r \leq 0.999$ ) – for predicting presence/absence of diabetes in individual users and their blood glucose level, respectively. Depending on the prediction output, the application presents tailored suggestions which include maintaining a healthy diet, losing weight, or exercising regularly to improve the quality of health. Based on the diabetic profile of a patient, Afreen et al. (2019) applied XGBoost regressor ( $r^2 = 0.9799$ ) to predict the total calorie consumption while MLP ( $r^2 = 0.9522$ ) is used to predict the percentage of carbohydrates, proteins, and fats in the daily diet. Within this range of values, a recommender system generates a personalized diet chart for the patient. Furthermore, M. Chen et al. (2018) employed an ensemble model ( $\text{acc} = 94\%$ ) comprising DT, SVM, and ANN to determine if a patient has a higher risk of diabetes or not. If diabetes risk is confirmed, the patient receives tailored interventions such as suggestions for breakfast, lunch, and dinner, as well as reminders to take oral hypoglycemic drugs or receive insulin treatment based on the patient's blood glucose index. Similarly, Sowah et al. (2020) built an adaptive system that uses a CNN model (based on GoogLeNet or Inception network (Szegedy et al., 2014)) to recognize food ( $\text{acc} > 95\%$ ) and KNN algorithm to suggest the right meals for breakfast, lunch, and dinner based on patients' profile and a knowledge base. For their personal nutrition project, Popp et al. (2019) developed a GB regression model ( $r = 0.6267$ ) to predict PPGR (postprandial glycemic response) for each diabetic patient. In other words, using the model, personalized PPGRs are calculated for every meal and snack based on their nutrient composition and calorie-adjusted quintile cutoffs of PPGR are used to create meal ratings of "excellent," "good," "medium," "bad," and "very bad." Patients are advised to make different choices or food substitutions if their meal is neither excellent nor good.

Besides diabetes, ML-driven adaptive systems were also developed to manage other diseases. For instance, the *MobileCoach* application (Barata et al., 2016) supports asthmatic patients by detecting coughs (via sounds obtained from their smartphone's microphone) using an SVM model ( $\text{acc} = 83.3\%$ ) and then providing personalized intervention to the patient based on the amount of coughs. Similarly, to support patients who suffered a cardiac event and are in Phase III of the recovery process, Prabhu et al. (2018) developed the *MedFit* system to recognize up to 14 local muscular endurance (LME) exercises completed by the patient using a high-performing SVM classifier ( $\text{acc} \geq 98\%$ ). The system also uses a repetition counting algorithm to determine the number of repetitions for each exercise detected. Based on the exercise completed and its repetition count, the system offers personalized video exercise classes

and feedback including motivational messages, progress visualizations, and repetition count report. Furthermore, El Barachi et al. (2019) addressed epilepsy by introducing an adaptive application – *EpiSense* – that can detect epileptic seizures in real-time by means of an SVM model ( $\text{acc} = 86\%$ ), and then alerts the patient's caretakers of the occurrence of a seizure for quick action. In a bid to establish an early warning mechanism and rescue strategy of hospital-home linkages for heart attack condition, J. Yang et al. (2020) built an adaptive system named *HEARTlistener* which employs an LSTM regression model (validation loss = 0.0054) to assess individual patients' current physiological state (breath rate and heartbeat) and predict their health condition. When an abnormal condition or heart attack is detected, the system triggers the first-aid strategy which involves contacting an emergency doctor online, requesting an ambulance from the medical service centre, and notifying the patient's household. For at-risk condition with no obvious abnormality, the responsible doctor first remotely checks the patient's physiological data and prediction results, and then decides whether to trigger first-aid strategy or not. Similarly, the *HeartMan* mobile application (Luštrek et al., 2021) is aimed at monitoring patients with congestive heart failure (CHF) by utilizing ML models for continuous blood pressure (BP) estimation – systolic BP (SBP) and diastolic BP (DBP), physical activity monitoring, and psychological profile recognition. RF and SVM were used to estimate BP (MAE: SBP = 9.0, DBP = 7.0) and SVM ( $\text{acc} = 88.6\%$ ) for predicting *motivated*, *anxious*, and *depressed* profiles. Based on model prediction/estimation, the application offers personalized exercise programs (such as endurance and resistance exercises), nutrition advice, CBT for anxiety and depression management, and mindfulness-related contents.

da Silva et al. focused on increasing treatment adherence of hypertensive older adults by developing an adaptive system that applied a DT model ( $\text{acc} = 95.1\%$ ) to predict adherence or non-adherence to prescribed medications (da Silva et al., 2019). In case of non-adherence, the system triggers personalized messages alerting individual patients, as well as doctors, caregivers, and relatives that the prescribed treatment is not being followed. Sookrah et al., however, aimed to reduce the amount of salt (sodium) intake by recommending personalized diet plans to patients who are at risk of hypertension (Sookrah et al., 2019). The recommender system classifies food-related parameters into *positive* or *negative* using MLP ( $\text{acc} = 99\%$ ), where positive means the food can be recommended. The system then applies content-based filtering technique to use all the foods classified as *positive* to recommend a meal plan with the highest score based on standard DASH serving sizes for the food category and the number of days since the user last consumed that meal plan to avoid repetitive recommendations. Arpaia et al. developed a telemonitoring system that provides an alert when patients' vitals exceed certain thresholds which are automatically personalized for individual patients (Arpaia et al., 2022). An LSTM Autoencoder ( $\text{acc} = 93\%$ ) predicts a score that quantifies the risk of hypertension



which in turn is checked against the threshold by the system. Rabbi et al. addressed obesity issues by developing an adaptive system that utilize the MAB reinforcement learning algorithm (cohen-d  $\leq 0.84$ ) to dynamically learn and influence user's physical activity and dietary behaviours by suggesting actions that maximize the chances of achieving calorie loss goals (Rabbi et al., 2015). Forman et al.'s *OnTrack* application aimed to prevent obesity by using a continuously improving ensemble model (sensitivity: 69.2%, specificity: 83.8%) of multiple algorithms (RF, Bayes Net, LogitBoost, and Random-subspace) to predict the risk of dietary lapse and deliver tailored interventions when risk is elevated (Forman, Goldstein, Crochiere, et al., 2019). A previous study on the application achieved 80% specificity (Forman, Goldstein, Zhang, et al., 2019). To enhance self-management and optimize non-surgical health care utilization in patients with knee and/or hip osteoarthritis, Pelle et al. (2019) created the *dr. Bart* application which employs the MAB dynamic model (sensitivity: 75%, specificity: 89%) to suggest tailored pre-formulated goals (top 5) to individual users based on their personal and contextual data. Goals were automatically chosen from a list of 72 goals covering physical activity/exercises, vitality/sleep quality, and nutrition. Interventions to motivate users to achieve their goals were also delivered using three behaviour change strategies – reminders, rewards, and self-monitoring – to enhance motivation, app engagement, and intervention effect.

Furthermore, Stamate et al. (2018) developed an adaptive and mobile-based system (cloudUPDRS) to help individuals affected by Parkinson's disease adhere to prescribed movements and to reduce test duration for each patient. This is achieved by using RCNN model (acc = 78%) to identify failures to follow the movement protocol (e.g., tapping the screen or holding the phone on the knee) and using an ensemble of randomized DTs to create personalized quick tests by selecting a subset of observations that closely estimate the motor performance of a particular patient. The system also provides audio, video, and textual media to guide patients and their caregivers to conduct the tests at home and in the community. A similar work by Stamate et al. (2017) used MLP (acc = 76.9%) instead of RCNN to support Parkinson's disease patients. Furthermore, Auffenberg et al. (2019) utilized clinicopathologic and demographic characteristics of prostate cancer patients to provide an individualized prediction of treatment. Specifically, they developed a system that applied an RF model (AUC = 0.81) to predict the probability of receiving a given primary treatment (radical prostatectomy, radiation therapy, primary androgen deprivation therapy, active surveillance, and watchful waiting). Using the prediction result, the system generates personalized treatment decisions based on similar patients from the clinical registry.

For sickle cell patients, Khalaf et al. (2016) developed a web-based system that leverages an MLP model (MAPE = 0.1345) to predict the correct amount of medication (Hydroxycarbamide drugs/liquid) or dose with the aim of providing accurate personalized therapeutic recommendations. Abd et al. (2017) attempted to classify patients into

those *with sickle cell trait* and those *without the trait* using the best performing model: LogitBoost (acc = 99.6%). For patients with sickle cell trait, the system further analyzed their clinical data (such as blood tests) to determine if the situation is critical or not; if critical, the patient receives personalized recommendations and treatment directly. Otherwise, the system contacts the physician directly to suggest the proper action that the patient should follow. Kariyawasam et al.'s *Pubudu* system aimed to enhance reading, writing and mathematical skills of children with dyslexia, dysgraphia, and dyscalculia conditions by providing tailored interventions (Kariyawasam et al., 2019). For dyslexia, the system applied CNN model to determine whether letters are pronounced correctly or not (acc = 65%) and KNN to predict whether a child has dyslexia or not. For dysgraphia, CNN model determines whether letters are written correctly or not (acc = 85%) while RF classifier (acc = 90%) was used to check the correctness of hand-written numbers. In addition, SVM classifier was used to determine if a child has the disease or not. For dyscalculia, an SVM model was used to detect the existence of the disease with high accuracy of 90%. Appropriate interventions are automatically triggered by the system to support the child. For example, if a child was predicted to be letter dysgraphic, he/she is trained on how to write letters in proceeding path with animations.

### 3.8.3. Medical diagnosis

In the area of medical diagnosis, Pathinarupothi et al. developed the RASPRO (Rapid Active Summarization for effective PROgnosis) system that provides *personalized* patient monitoring, *precision* diagnostics, and *preventive* criticality alerts (Pathinarupothi et al., 2018). The system employed SVM model to make diagnostic prediction of Acute Hypotensive Episodes (AHE) with a high F1-score (> 88%). Next the system determines severity level to prioritize individual patients based on their urgency for physicians' interventional attention. Alerts are automatically triggered for critical conditions. Similarly, in the decision-making phase of their healthcare monitoring system, Kesavan and Arumugam (2020) classified patients' health conditions into *normal*, *sensitive*, and *critical* severity levels using the ADCNN classifier in conjunction with Levy Flight-based Grey Wolf Optimization (LFGWO) algorithm to optimize the weights (F1-score = 95%). If a patient's condition is not normal, the system notifies the doctor or practitioner immediately. Also, Neloy et al.'s Critical Patient Management System (CPMS) used several ensemble ML models including Bagging SVM which achieved the highest accuracy of 92%, to predict the health condition of a patient (Neloy et al., 2019). If the condition is worse, CPMS sends an SMS to the appropriate health professionals (e.g., doctor/nurse on duty) for immediate treatment of the patient. Asthana et al. developed an adaptive framework that applies a DT model (RMSE = 0.1066) to predict a user's at-risk health conditions, and then suggests appropriate wearable devices and measurements that can help the user to evaluate health risks and monitor them (Asthana et al., 2017). Readings from the

recommended devices are captured in a feedback loop to retrain the model to update the measurements and possibly the wearables previously suggested. Zhou et al. (2020) also created an adaptive recommendation system based on CRNN classification model (acc = 88.63%) that provides patients with automatic clinical guide and pre-diagnosis suggestions. The system is able to guide a patient to the appropriate healthcare department according to his/her question or suggests the names of the drugs according to the patient's symptoms.

Furthermore, Khumrin et al. (2018) utilized an NB model (acc = 60%) to support the development of problem-solving and diagnostic decision-making skills in medical students. As students work through a clinical scenario (or virtual cases), the model monitors and analyzes their decision path by predicting the target diagnoses given a set of patient information (including demographics and symptoms) and the system (called *DrKnow*) then formulates personalized feedback to help students review and reflect on their diagnostic rationale and inform next steps. Koren et al. developed an adaptive diagnostic system that prompts questions about patient's symptoms and then used an ensemble of BayesNet, XGBoost, LogisReg, and ANN (acc = 83.9%) to determine the cohort of similar people with a similar set of symptoms (Koren et al., 2019). Afterwards, the patient is shown the cohort's path to treatment including the various conditions with which they were diagnosed, along with the full course of action of the cohort (such as the types of medical professionals seen, tests ordered, medications prescribed, and expected recovery time). A feedback loop continues to update the model using patients' response to follow-up questions (probing for their eventual diagnosis and treatment after seeing a doctor). In their adaptive recommender system, Aujla et al. (2019) used a DT model (RMSE = 0.424, MAPE = 4.64%) to first classify individual patients' health data into one of  $k$  diseases, and then recommend a ranked list of doctors that can provide the right treatments for the disease using CNN. Each patient can choose a nearby or remote doctor based on his/her ranking. Cheerla et al.'s system focused on diagnosing different cancer types using patients' miRNA data (Cheerla & Gevaert, 2017). They applied an SVM model to classify different cancer types with a high overall accuracy of 97.2%, and then with the help of a prognosis model (also based on SVM with an overall accuracy of 85%) developed a recommendation module that suggests three personalized treatment regimens to each patient. Also, a semi-supervised algorithm was created to periodically retrain the models using new clinical and miRNA samples uploaded via a dedicated system interface. To support fetal health and well-being, Akbulut et al. (2018) developed an adaptive system that employs a DF model (acc = 89.5%) to predict fetal anomaly status based on maternal and clinical data, and then recommends suitable physical activities (with corresponding schedule) to perform during pregnancy depending on the predicted status. Barbosa et al. developed the HegIA system (Barbosa et al., 2021) to detect the presence or absence of COVID-19 virus from clinical data (e.g., blood tests) using RF model (acc = 92.9%).

The model is also used to predict/recommend the best type of hospitalization for the patient: *regular ward*, *semi-ICU* or *ICU* (acc > 99%).

### 3.8.4. Assistive healthcare

To assist people with cognitive impairment, Javed et al. (2021) created the PP-SPA framework to detect the current activity or task a user is performing using HT and LogisReg models both of which achieved an overall F1-score of 90.2%. The current task is then recorded in a digital diary and a personalized real-time support (e.g., prompts showing functional aid) is provided to help the user complete the task. For the visually impaired, *Follow Me!* application (Kajiwarra & Kimura, 2019) provides smart navigation support using an RF classifier (acc = 92%) to detect object type (such as "same" representing people walking in the same direction as the visually impaired user, "oncoming" for oncoming pedestrians, "steps," and "unknown"), and then recommending a safe route to the user. Furthermore, Sarwar and Javed (2019) developed an ambient system that utilizes an NB model (acc = 90%) to recognize the current activity being performed by a user (e.g., sleeping, walking, etc.). Based on the recognized activity, the system suggests an optimal care plan for health improvements. Moreover, the *iLocate* framework created by Van Woensel et al. (2020) applied a DT classifier (acc = 84%) to accurately identify individuals' discrete location (room-level estimation) and hierarchical clustering (acc = 90%) for region-level estimation, together with knowledge-based techniques (based on ontology concepts) to supply the associated semantics of identified locations. The eHealth system integrated with *iLocate* delivered context-sensitive care activities in line with the inferred current location, as well as pathfinding aids and health warnings in care facilities. Another assistive and adaptive system is *Lynx* (Lopez-Guede et al., 2015) which monitors an older adult's habit or behaviour to detect deviations from his/her normal daily tasks (wake-up times, sleep habits, diary strolls, etc.) and health situations. To achieve this, the system uses a DT model to predict the user's current state (acc = 81.8%) and LOF algorithm to detect anomalies. Once an unusual situation is detected, the system sends a real-time alarm/notification to the family, care centre, or medical agents.

ML's utilization also extends to the area of human fall detection to drive personalized interventions, especially for the elderly. For example, Martín et al. (2011) developed a mobile-based adaptive multiagent system capable of detecting if a fall has occurred or not using a DT regression model (MQE = 0.16), and then placing an automatic emergency call to the closest medical centre based on user location/profile if a fall is detected. An SMS is also sent to the contact person configured by the user. A similar system targeting older adults is *Whoops* (Mrozek et al., 2020) which is a mobile application that uses a BDT classification model deployed on the edge (i.e., on user's phone with 99.2% accuracy) and on the cloud (acc = 99.8%) to detect falls. If a fall is detected, the patient is expected to confirm if a fall truly occurred. Once confirmed, an emergency call is placed to the caregiver automatically. Chin et al. developed an

adaptive fall detection robot that relies on a lightweight deep learning model – SlowFast Network (test loss = 0.477) – deployed on an iPhone mounted on the robot to detect human falls (Chin et al., 2020). The same model is also deployed on a 2D camera installed in a room to track and detect a fall. Once a fall is detected or human is not in-view, the camera triggers the assistive robot to locate the human and assess for a fall. If a fall is confirmed, the caregiver is notified for quick rescue. Likewise, Hassan et al. (2019) used the CNN-LSTM deep learning model (F1-score = 97%) running on a smartphone to detect human falls in real-time. If a fall is detected by the model, the system triggers an indoor sound alert to family members through a wireless access point at home and an outdoor SMS alert is sent to a hospital or caregiver through a mobile network base station. Rajasekaran et al. developed a virtual nursing system using Deep Continuous Deep Belief Network (DC-DBN) with Restricted Boltzmann Machines (RBM) that processes video streams to track elderly patients' well-being (TPR = 95.7%, FPR = 0.04%) (Rajasekaran & Kousalya, 2022). An alarm is transmitted to the patient's relatives via the system if the patient's motions alter abruptly (such as during a fall).

### 3.8.5. Physical activity

To foster physical activity, Dijkhuis et al. (2018) developed an adaptive system that makes hourly estimations of the probability of users meeting their individual daily physical activity goals using RF model with a mean accuracy of 93%. Based on this estimation, each user receives personalized feedback throughout the day to aid goal achievement. Similarly, Z. Li et al. (2019) developed an adaptive system that generates an hour-by-hour activity plan (steps goal) based on the user's probability of adhering to the plan. Pretrained models (based on RF algorithm;  $AUC \geq 0.8$ ) which improve as step counts collected incrementally over time are fed into them were deployed on users' smartphone to predict the likelihood of reaching their individual activity target. If the predicted probability is lower than a threshold, the system suggests an alternative plan to the user; otherwise, the daily target is adjusted and a new activity plan is recommended. Suh et al. built an adaptive system that generates a personalized exercise schedule based on individuals' exercise sessions using a DT model (acc = 88.71%) (Suh et al., 2012). In addition, the system recommends conditions allowing the user to maximize the amount of exercise within a given heart rate range while avoiding fatigue. Kadri et al. applied a DT classifier (F1-score = 62%) to predict physical activity behaviour of users (jogging, sitting, standing, upstairs, downstairs, and walking) (Kadri et al., 2020). A BiLSTM deep learning model was also used but achieved a lesser performance (F1-score = 59%), compared to the DT model. Based on the predicted activities, the recommendation system informs individual users about their health behaviour including calorie suggestions. Furthermore, the *WalkPal* system (Sansrimahachai, 2020b) for older adults predicts weekly walking exercise minutes based on contextual and health information using an

ANN regression model (MAE = 16.78, MSE = 15.3), and then generates a challenging but realistic personalized walking plan for each elder based on the predicted minutes. Ally+ is a chat-based mobile application that employed two ML models (SVM and LR) trained using contextual information about a person (time, battery state and level, phone state, and activity) to predict when a person is more receptive and then triggered interventions at that moment (Mishra et al., 2021). Jamil et al. developed a proof-of-concept, blockchain-based fitness application that collects fitness data via IoT devices and then applied an SVM model (acc = 92.1%) in its inference engine to recommend personalized diet and workout plan to individuals (Jamil et al., 2021). Wang et al. created a reinforcement learning agent (PAUL) to adaptively select the optimal strategy for delivering physical activity reminders with respect to the momentary context of this user (time and calendar) (S. Wang et al., 2021). 83.3% of reminders sent at adaptive moments were able to elicit user reaction within 50 min.

### 3.8.6. Dietary management

To encourage healthy eating, Rachakonda et al. (2020) developed an adaptive system called *iLog* to monitor users' eating behaviour in real-time and provide personalized interventions. Specifically, the system utilized a MobileNet deep learning model (acc = 98%) running on the edge (smartphone and single-board computer) to detect, classify, and quantify food objects on the plate of the user. Based on the quantification, the system determines if the user is eating normally (normal-eating) or eating under uncontrollable cravings (stress-eating). In case of stress-eating, remedial interventions are offered such as suggestions on when, what, and how much to eat. Xu et al. targets children's dietary health by developing a system that employs LR model (Precision > 80%) to estimate food weight and nutrient load (Xu et al., 2019). Based on these estimations, the user receives personalized dietary recommendations. Likewise, Spanakis et al. aimed to assess eating behaviour via their *ThinkSlim* mobile application which employs DT classification model to extract significant rules that indicate what combinations of states (e.g., scoring high on food craving + being at home + low positive feelings + negative feelings) are predictive of unhealthy or healthy eating (Spanakis et al., 2017). Based on these rules, the application warns individual users (via adaptive messaging) prior to a probable unhealthy eating event. The application was found to be effective in increasing the number of users ( $n=8$ ) eating healthily in first four weeks of intervention to 20 in the last two weeks of intervention. Gorbonos et al., however, developed a system (*NutRec*) which finds recipes that best fit a set of ingredients while following healthy eating guidelines (Gorbonos et al., 2018). The system utilized ANN model (acc = 62.63%) and NMF technique to model the interactions between ingredients and their proportions within recipes for the purpose of offering appropriate recommendations.



### 3.8.7. Health monitoring

Adapting interventions based on vital sign tracking outcome is another interesting area targeted by a few papers. For instance, P. Chiang and Dey (2019) created an adaptive framework which applied an RF-based model, called *RF with Feature Selection* (RFFS), to predict systolic BP (SBP) and diastolic BP (DBP) of individual users. Based on this prediction, each user receives personalized health behaviour recommendations such as increasing exercise or going to bed earlier. Also, the RFFS model (MAE: 5.18 and 4.30 for SBP and DBP respectively) is continuously refined (online learning) using the online weighted-resampling technique. Furthermore, W. Gu (2017) designed and developed *BGMonitor*, a personalized smartphone-based monitoring system that detects abnormal blood glucose (BG) events based on contextual data such as meal, medications and insulin intake, physical activity, and sleep quality using the Multi-RNN deep learning model (82.14%), and then reminds the user to double-check using a clinical continuous glucose monitoring (CGM) device or a finger pricking method. Moreover, Zeevi et al. (2015) developed an adaptive framework to predict individual's postprandial glycemic response (PPGR) to real-life meals using a GB regression model ( $r \leq 0.80$ ). Based on the predicted PPGRs, personalized dietary suggestions are provided. Kang developed a human performance management system to monitor the fitness status of warfighters in real-time using physiological data (e.g., heart rate, respiration rate, skin temperature and blood pressure) and Levenberg-Marquardt algorithm – a neural network-based on MATLAB R2021b – (acc = 95.6%) to identify and predict fatigue, potential injury, or physical strength (Kang, 2021). The system offers personalized fitness training based on individuals' predicted health index and fitness level. Chiang et al. developed a BP prediction and recommendation system that uses ML model to predict a user's current BP level using his/her historical BP readings as well as activity, sleep and heart rate data (P.-H. Chiang et al., 2021). The system also identifies the most important lifestyle features/factors that impact the individual's BP trend, and then recommend the next top feature to the user. For example, if higher walking speed is associated with a lower BP, the system suggests that the user increases his/her walking/running speed.

### 3.8.8. Substance use

To discourage smoking behaviour, T. Chen et al. (2018) developed an adaptive system based on an LSTM deep learning model (F1-score = 86.3%) to detect smoking motion (out of six possible motions) in real-time. Once smoking motion is confirmed, the system triggers a request to an internet message platform (*Trumpia*) to send an alert message that includes quit plan information to the user and subscribers (such as doctors or family members). Bae et al. focused on dissuading alcohol use in young adults (Bae et al., 2018). Particularly, they developed a system to classify time periods as *non-drinking*, *low-risk*, and *high-risk drinking* using an RF model (AUC = 96%). The system has the potential to trigger just-in-time behavioural interventions

once a high-drinking is detected (such as delivering supportive messaging or contacting supportive friends or family members), though yet to be implemented according to the authors. To combat opioid-use disorder (OUD), Scherzer et al. (2020) developed a mobile-based adaptive system (*Marigold*) to detect text messaging content that may signal relapse or impending relapse in patients recovering from OUD. Basically, the system checks for two red flags within peer chats or messages: (i) implied intent to harm self or others, and (ii) malicious conduct or "trolling," using a ML model (nicknamed *Marigold model*) based on natural language processing. The model predicts whether a given message needs moderator intervention (in cases of self-harm, harm to others, risk of relapse) or not (F1-score  $\geq 85\%$ ), and also predicts the severity of the message on a scale from 1 to 5 (F1-score  $\approx 70\%$ ). Contents flagged for suicidal or homicidal ideation are automatically sent to a clinician for review in real time, while others are sent to a moderator. With such adaptive assessment of patient relapse risk and subsequent notification of providers when a patient needs more intensive care resources, the system provides a safe and accessible peer support to individuals.

### 3.8.9. Other domains

To promote good dental and oral health, Stark and Samarah (2019) designed and tested a prototype of an adaptive and IoT-enabled toothbrush (paired with a mobile application) that detects the tooth and surface brushed via an RF model (acc: 99.71% and 99.63% for *left-handed* and *right-handed* users respectively). The application provides comprehensive and real-time feedback to individual users, informing them about their brushing behaviours including suggestions to improve tooth brushing. Users are also reminded to floss and clean their tongue, as part of good oral hygiene practices. Furthermore, Tuti et al. developed a mobile-based gamified system to provide personalized training on emergency neonatal care, specifically on early recognition and treatment of new-born babies who need urgent care and hospitalization (Tuti et al., 2020). To achieve this, the system employs LSTM model (acc  $\leq 88.32\%$ ) to predict learners' future performance and forgetting curves by feeding the model sequence embeddings of learning task attempts from specific healthcare providers. Based on the prediction outcome, the system provides timely interventions that support self-regulated personalized learning at scale. In the domain of homeopathy, Priyadarshi and Saha (2020) developed an adaptive system that analyzes patients' question using two SVM classifiers to determine whether the question is seeking remedy (acc = 99.04%) and part of five shortlisted diseases (acc = 93.42%), respectively, and then forms relevant queries to search the web for homeopathy remedy (medicines) related to the disease. The system recommends top-ranked medicine name to the user. Moreover, Edara et al. (2020) created an adaptive framework for recommending mobile health applications with positive reviews to individual users based on their health conditions. The framework applied a hybrid deep learning model (F1-score = 97.9%) comprising both DNN and LSTM to generate useful recommendations.

In the area of medication management, Hu et al.'s prescription system used CNN and LDA to predict and recommend appropriate herbs for individuals based on their tongue image (Hu et al., 2021). Zhang et al. developed a web-based COVID-19 surveillance system that utilized a fine-tuned BERT model for classifying tweets into COVID-19-related or otherwise (F1-score = 98.0%) (Y. Zhang et al., 2022). Classified tweets are then analyzed to forecast epidemic using linear regression model ( $R^2 > 0.90$ ). The system provides early warning messages based on the forecast. Abdo et al. designed an adaptive mobile-based healthcare monitoring framework that preserves the location privacy of patients without negatively impacting quality of service (Abdo et al., 2020). The framework leverages a high-performing DT model (acc = 95.4%) to classify users' health state into one of the following: *urgency*, *illness*, and *healthy*. In case of urgency, location is sent accurately to the nearest medical centre for immediate dispatch of an ambulance. If healthy, the user's location will not be shared. If illness is predicted, location privacy-preservation methods (perturbation and obfuscation) are applied based on the privacy level preferred by the user or patient. To sustain user engagement and promote long-term effectiveness of *Saathealth* (a mobile application that provide interactive contents on children's health, nutrition, and development), Ganju et al. applied RF model (acc = 93.0%, RMSE: 25.09,  $R^2 = 0.91$ ) to predict user churn and user lifetimes (i.e., number of days a user will stay on the application before uninstalling). Model output was used to incentivize users with optimized and more personalized/targeted offers and omni-channel nudges, as well as augmented in-app experiences.

Finally, Lee et al. used patients' actual descriptions of their symptoms on social media to develop an LSTM model (F1-score = 73.9%) for medical specialty classification (Lee et al., 2021). A web-based chatbot was then developed to help patients understand which medical specialty is appropriate for the treatment of their current symptoms and then make an appointment with the appropriate specialist.

## 4. Discussion

ML is a hotspot in artificial intelligence and health and wellness systems have recently tapped into its capability to learn individuals' peculiarities and conditions with the aim of adapting interventions or treatments. Figure 2 shows that most existing work in this area were from recent years, thereby confirming that ML-based adaptive health and wellness systems are emerging and will continue to advance. However, there are challenges or limitations that should be addressed to improve the accuracy and reliability of the tailored interventions they offer. In this section, we discuss these challenges and how to address them.

### 4.1. Challenges of ML-driven adaptive systems for health and wellness

#### 4.1.1. Data availability, data diversity, and data quality issues

Obtaining required data for training ML models remains a challenge. Shortage of data may be due to stringent privacy

policies associated with health-related or clinical data, hence hampering data sharing for research purposes. Also, obtaining labelled data for supervised learning tasks are generally time-consuming and expensive. Consequently, insufficient dataset for each target class would cause unbalanced classes which in turn could bias a supervised model toward the majority class(es) and ultimately lead to poor performance. Small dataset can also limit the ability of advanced ML models, such as deep learning models, to effectively learn complex patterns required for adaptation. Various works reported issues in model performance due to insufficient training data (Bae et al., 2018; Burns et al., 2011; Delmastro et al., 2020; Kadri et al., 2020; Khumrin et al., 2018; Mahyari & Pirolli, 2021; Nosakhare & Picard, 2020; Spanakis et al., 2017; Stark & Samarah, 2019; Tuti et al., 2020) including Auffenberg et al. (2019) whose model had difficulty discriminating between active surveillance and watchful waiting since the model was not trained with enough data to accurately separate the two classes. Second, there is a challenge of variability or diversity in the target population or cases observed by many papers (see [Supplementary Appendix A](#)), thereby reducing model generalizability. For example, Alfian et al. (2018) limited their dataset to Pima Indian women, Pelle et al. (2019) focused on knee/hip osteoarthritis, Popp et al. (2019) targeted English- and Hebrew-speaking audiences, Bae et al. (2018) targeted young adults only, and so on.

Another challenge reported in the literature is data quality issue, such as missing values (Bae et al., 2018; Ghandeharioun et al., 2019; Khumrin et al., 2018) and inconsistency in data labelling (Burns et al., 2011; Nosakhare & Picard, 2020). The latter has to do with different human annotators assigning dissimilar labels to similar samples due to individual bias/subjectivity. Missing values could mean losing important information that might have provided further context to guide a model's learning curve. Therefore, if not addressed, these data quality issues would render a model unreliable due to incorrect predictions that could be disastrous for a patient if a wrong intervention or treatment is suggested, for example.

#### 4.1.2. Infrastructure-related issues

Furthermore, infrastructure-related issues – such as high latency while retrieving data from external sources (e.g., sensors, websites, etc.) or transmitting data to a remote server, CPU and memory consumption, battery drainage, unstable internet connectivity, and mobile data usage – were reported in the literature (Alfian et al., 2018; Bae et al., 2018; Burns et al., 2011; Pathinarupothi et al., 2018). People in low-resource communities or areas are more susceptible to connectivity issues, and even incur internet charges they may not be able to afford (Salemink et al., 2017). Consequently, system functions would be disrupted including consistent delivery of interventions that could have improved users' health conditions or save their lives. Battery drainage, on the other hand, has a generic impact (geography-independent), hence would have far-reaching consequences including increased dropout rate whereby many users abandon the system altogether (Murnane et al., 2015). Also, scalability



issues should be addressed since model training consumes much compute and memory resources. Cases of poor calibration of sensors and bluetooth-related issues were also reported in the literature (Elvitigala et al., 2021).

#### 4.1.3. Suitability of interventions

Although many papers achieved a considerably high model performance (Supplementary Appendix A), majority of the reviewed literature (78%) did not evaluate the effectiveness of their systems to determine whether the health and wellness interventions are actually suitable to individual users and achieve the desired effect, as well as whether user preferences have changed so the model can self-update. For instance, some of the works offer therapeutic treatments that similar patients have received (Auffenberg et al., 2019; Nosakhare & Picard, 2020); however, treatment may not be medically appropriate for an individual even if similar patients receive it. Also, offering tailored interventions based on user groups (as found in Spanakis et al. (2017)) does not necessarily translate to a personalized experience since individuals within the same group may have dissimilar preferences and thus could find some interventions unsuitable.

#### 4.2. Recommendations and future research

Recent advances in ML have revolutionized health and wellness services, as individuals can receive tailored or personalized interventions/treatments in context and in real-time. Yet, the challenges discussed earlier need to be addressed to maximize the above benefits/opportunities and also sustain user engagement with ML-based adaptive health and wellness systems while building trust and confidence.

First, with the advances and popularity of Internet of Things (IoT), physiological and contextual data can be collected in large volumes from the target population (e.g., patients) via sensors embedded in wearables and smartphones. IoT has been shown to be a structured and well-established technique for handling the health and wellness needs of patients based on remote monitoring and mobile health (Kesavan & Arumugam, 2020). Therefore, rather than relying on subjective sources (such as self-reports or questionnaires) which require more time/effort and usually result in small datasets, objective data from sensors should also be captured to increase the data volume and also improve data variability as more diverse people are added to the pool of subjects. In addition, data capture can be upscaled to target more health conditions by using additional and appropriate sensors. Also, public datasets are constantly being shared on relevant online platforms (see Supplementary Appendix B) and can be leveraged by researchers and developers to augment existing datasets. As the data volume increases, low quality samples can then be removed from the training set with minimal impact on model performance. Hence, by using multimodal data collection approach with appropriate data fusion techniques, more actionable and meaningful physiological and contextual features could be injected into the ML model, thereby improving the overall model

prediction/classification accuracy. Based on our findings (see Supplementary Appendix A), 77% ( $n=27/35$ ) of the reviewed papers that used multimodal data reported good model performance of **at least 80%**, while 69% ( $n=36/52$ ) of papers that utilized single-modal data achieved similar performance. Some of the unimodal papers already have future plans for multimodal data acquisition (El Barachi et al., 2019; Forman, Goldstein, Crochiere, et al., 2019; Hassan et al., 2019; Neloy et al., 2019; Spanakis et al., 2017).

Second, to address infrastructure-related issues associated with ML-driven adaptive systems, the following could be considered:

- i. Cloud-based Platform-as-a-Service (PaaS) solutions guarantee faster model training times by providing needed resources on demand, especially compute power and memory, hence balancing scalability and cost. This offers significant benefits over the on-premises option which is usually difficult to auto-scale and may even be more costly to maintain. PaaS solutions are available on popular Cloud technology platforms such as Azure (Copeland et al., 2015), Google Cloud (Bisong, 2019; Krishnan & Gonzalez, 2015), Amazon Web Services (AWS) (Wittig & Wittig, 2018), and IBM Cloud (IBM, n.d.). Mrozek et al. (2020) and Neloy et al. (2019) are sample papers that used Azure and IBM Cloud to train their ML models.
- ii. Online learning using pretrained models (i.e., models trained offline using Cloud environments, for example) deployed on the edge (e.g., smartphones) helps to minimize battery consumption since data transmission to a remote server in terms of size and frequency (network traffic) have been drastically reduced. This also reduces mobile data usage and cost since the model resides on the user's phone instead of a remote server. Researchers and developers can utilize ML frameworks such as TensorFlow Lite (Google Inc., n.d.-b), Google ML Kit (Google Inc., n.d.-a), and Apple Core ML (Apple Inc., n.d.) to train and deploy their models on smartphones. Research has shown that edge-based model achieves similar predictive performance as cloud-based model (Mrozek et al., 2020).
- iii. Although most adaptive systems rely on notifications to suggest interventions or keep users on track with their therapeutic commitments, excessive notifications drain users' battery faster than normal. Priority notifications such as those suggesting interventions or actions to be taken could be sent more frequently than notifications that merely inform users of their progress or performance. The latter should be sent at reasonably spaced intervals and not frequently.
- iv. Procedure for automatically re-calibrating sensors could be integrated into the overall architecture such that re-calibration occurs at the appropriate time/interval without user intervention to ensure continuous collection of accurate readings. Also, providing Bluetooth auto-reconnection feature within the system would prevent data loss and the burden imposed by manually

reconnecting to Bluetooth each time a disconnection occurs.

Third, the essence of interventions is to bring about improved health condition and well-being in target audience. Hence, it is imperative for researchers/developers to implement a mechanism (such as experience sampling) within the system to collect feedback about users' health status and preferences at specific time intervals after applying the interventions to their lives or situations. This feedback should (automatically) be used to improve the performance of ML models used by the adaptive systems. This process of incorporating user feedback into a model's learning process is known as the *Feedback Loop*. With feedback loop, a system can almost accurately determine whether to replace an intervention, stop an intervention, or supplement an intervention with others to make it more effective. However, most existing systems do not utilize feedback loops. Based on our review findings, only two papers incorporated feedback loop into their system architecture (Asthana et al., 2017; Koren et al., 2019). This gap should be addressed by researchers/developers to enhance their models' predictive performance and also future-proof their systems with respect to sustained user engagement and continuous delivery of effective/reliable health and wellness interventions.

Fourth, to address the privacy and security concerns linked to data transfer from users' devices to Cloud servers, researchers/developers of adaptive health and wellness systems could implement advanced and sophisticated approaches such as Federated Learning to train their ML models on multiple decentralized edge devices using local data without transferring or exchanging them (T. Li et al., 2020; Rieke et al., 2020). This training should be done during off-peak periods (i.e., when users do not actively engage with their devices) or when the devices are connected to a power source to avoid battery drainage. Also, researchers can gain users' trust and acceptance by leveraging Explainable AI (XAI) methods and tools (Rai, 2020; Shin, 2021) to describe (in a comprehensible fashion) how a model reaches its final decision including potential biases and expected impact.

Finally, researchers and developers should consider co-creating interventions with target users and health professionals such that they are evidence-based and more impactful. Also, the right techniques for delivering these interventions should be discussed with experts, depending on the target health and wellness domain. For example, persuasive and behaviour change strategies such as rewards, reminders, simulation, suggestion, goal setting, etc. could be operationalized in adaptive systems to reinforce and sustain interventions, hence making them more effective (Baumeister et al., 2019). We also suggest that researchers should evaluate the efficacy of their systems in real-world settings by conducting long-term field trials to uncover insights as to whether the interventions lead to improved health and well-being in target audience/users. In addition, conducting usability studies would reveal issues that could inhibit system effectiveness and user engagement.

## 5. Conclusion

In this paper, we conducted an extensive review of 11 years research (January 2011–April 2022) to explore current trend in adaptive and personalized systems for health and well-being using machine learning techniques. We searched for articles systematically and selected 87 articles that met inclusion criteria for review. The selected articles target various health domains which are *disease management, assistive healthcare, medical diagnosis, mental health, physical activity, dietary management, health monitoring, and substance use*. Other domains include *smoking cessation, homeopathy remedy finding, patient privacy, mHealth app finder, clinician knowledge representation* for emergency neonatal care, *dental and oral health, medication management, disease surveillance, medical specialty recommendation, and health awareness*. We summarized key findings in the area of data collection, model development process, ML techniques, model evaluation, and adaptive/personalization strategies. Specifically, across the target domains, we explored data modalities, data types, data fusion methods, feature extraction/data preprocessing techniques, as well as ML techniques including supervised learning, unsupervised learning, and reinforcement learning. We also discussed evaluation of ML models using various metrics, as well as how models were applied in health and wellness systems to achieve adaptivity or personalization. Several challenges were also identified from the literature such as data volume constraints, data quality issues, data diversity issues, infrastructure constraints, and suitability of interventions.

Future research could address these challenges by leveraging advances in multimodality, Cloud technology, online learning, edge computing, automatic re-calibration, Bluetooth auto-reconnection, feedback pipeline, federated learning, explainable AI, and co-creation of health and wellness interventions. We anticipate further growth and widespread adoption of ML-driven adaptive systems in health and wellness domains to promote tailored, contextual, and real-time interventions/support and motivate better adherence and continuous user engagement.

## Disclosure statement

There are no financial or non-financial competing interests to report.

## Funding

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program. We acknowledge the support of Mitacs Canada and the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Grant.

## ORCID

Oladapo Oyeboade  <http://orcid.org/0000-0002-5797-7790>

Jonathon Fowles  <http://orcid.org/0000-0002-0268-0762>

Rita Orji  <http://orcid.org/0000-0001-6152-8034>

## References

Abd, D., Alwan, J. K., Ibrahim, M., & Naeem, M. B. (2017). *The utilisation of machine learning approaches for medical data classification*

- and personal care system management for sickle cell disease [Paper presentation]. 2017 Annual Conference on New Trends in Information and Communications Technology Applications, NTICT 2017 (pp. 213–218). <https://doi.org/10.1109/NTICT.2017.7976147>
- Abdo, M. A., Abdel-Hamid, A. A., & Elzouka, H. A. (2020). A cloud-based mobile healthcare monitoring framework with location privacy preservation [Paper presentation]. 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT) (pp. 1–8). <https://doi.org/10.1109/3ICT51146.2020.9311999>
- Afreen, N., Singh, S., & Kumar, S. (2019). Individualized patient-centered type 2 diabetes recommender system. In *International Conference on Advances in Computing and Data Sciences (ICACDS)* (pp. 45–54). Springer. [https://doi.org/10.1007/978-981-13-9939-8\\_5](https://doi.org/10.1007/978-981-13-9939-8_5)
- Akbulut, A., Ertugrul, E., & Topcu, V. (2018). Fetal health status prediction based on maternal clinical history using machine learning techniques. *Computer Methods and Programs in Biomedicine*, 163, 87–100. <https://doi.org/10.1016/j.cmpb.2018.06.010>
- Alfian, G., Syafrudin, M., Ijaz, M. F., Syaekhoni, M. A., Fitriyani, N. L., & Rhee, J. (2018). A personalized healthcare monitoring system for diabetic patients by utilizing BLE-based sensors and real-time data processing. *Sensors*, 18(7), 2183. <https://doi.org/10.3390/s18072183>
- Alharthi, R., Alharthi, R., Guthier, B., & El Saddik, A. (2019). CASP: Context-aware stress prediction system. *Multimedia Tools and Applications*, 78(7), 9011–9031. <https://doi.org/10.1007/s11042-017-5246-0>
- Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., & Aljaaf, A. J. (2020). A systematic review on supervised and unsupervised machine learning algorithms for data science. In M. W. Berry (Ed.), *Supervised and unsupervised learning for data science* (pp. 3–21). Springer. [https://doi.org/10.1007/978-3-030-22475-2\\_1](https://doi.org/10.1007/978-3-030-22475-2_1)
- Apple Inc. (n.d.). *Core ML overview – Machine learning – Apple developer*. Retrieved October 31, 2021, from <https://developer.apple.com/machine-learning/core-ml/>
- Arpaia, P., Crauso, F., de Benedetto, E., Duraccio, L., Improta, G., & Serino, F. (2022). Soft transducer for patient's vitals telemonitoring with deep learning-based personalized anomaly detection. *Sensors*, 22(2), 536. <https://doi.org/10.3390/s22020536>
- Asthana, S., Megahed, A., & Strong, R. (2017). A recommendation system for proactive health monitoring using IoT and wearable technologies. In M. Z.-H. Tata S. (Ed.), *Proceedings - 2017 IEEE 6th International Conference on AI and Mobile Services, AIMS 2017* (pp. 14–21). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/AIMS.2017.11>
- Auffenberg, G. B., Ghani, K. R., Ramani, S., Usoro, E., Denton, B., Rogers, C., Stockton, B., Miller, D. C., & Singh, K. (2019). askMUSIC: Leveraging a clinical registry to develop a new machine learning model to inform patients of prostate cancer treatments chosen by similar men. *European Urology*, 75(6), 901–907. <https://doi.org/10.1016/j.eururo.2018.09.050>
- Aujla, G. S., Jindal, A., Chaudhary, R., Kumar, N., Vashist, S., Sharma, N., & Obaidat, M. S. (2019). DLRS: Deep learning-based recommender system for smart healthcare ecosystem [Paper presentation]. ICC 2019 - 2019 IEEE International Conference on Communications (ICC) (pp. 1–6). <https://doi.org/10.1109/ICC.2019.8761416>
- Bae, S., Chung, T., Ferreira, D., Dey, A. K., & Suffoletto, B. (2018). Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addictive Behaviors*, 83(SI), 42–47. <https://doi.org/10.1016/j.addbeh.2017.11.039>
- Barata, F., Kowatsch, T., Tinschert, P., & Filler, A. (2016). *Personal MobileCoach: Tailoring behavioral interventions to the needs of individual participants* [Paper presentation]. UbiComp 2016 Adjunct - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 1089–1094). <https://doi.org/10.1145/2968219.2972713>
- Barbosa, V. A. D. F., Gomes, J. C., de Santana, M. A., de Lima, C. L., Calado, R. B., Bertoldo Júnior, C. R., Albuquerque, J. E., de, A., de Souza, R. G., de Araújo, R. J. E., Mattos Júnior, L. A. R., de Souza, R. E., & dos Santos, W. P. (2021). Covid-19 rapid test by combining a Random Forest-based web system and blood tests. *Journal of Biomolecular Structure and Dynamics*, 1–20. <https://doi.org/10.1080/07391102.2021.1966509>
- Basu, D., Kashid, S., Pawar, S., & Datta, D. (2020). An integrated detection and treatment recommendation framework for breast cancer using convolutional neural networks and TOPSIS [Paper presentation]. 2020 IEEE 17th India Council International Conference, INDICON 2020, December 10. <https://doi.org/10.1109/INDICON49873.2020.9342266>
- Baumeister, H., Kraft, R., Baumel, A., Pryss, R., & Messner, E.-M. (2019). Persuasive E-health design for behavior change. In *Digital phenotyping and mobile sensing* (pp. 261–276). Springer. [https://doi.org/10.1007/978-3-030-31620-4\\_17](https://doi.org/10.1007/978-3-030-31620-4_17)
- Ben-Israel, D., Jacobs, W. B., Casha, S., Lang, S., Ryu, W. H. A., de Lotbiniere-Bassett, M., & Cadotte, D. W. (2020). The impact of machine learning on patient care: A systematic review. *Artificial Intelligence in Medicine*, 103, 101785. <https://doi.org/10.1016/j.artmed.2019.101785>
- Bisong, E. (2019). *Building machine learning and deep learning models on Google cloud platform: A comprehensive guide for beginners*. Apress.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Buettner, R., Klenk, F., & Ebert, M. (2020). A systematic literature review of machine learning-based disease profiling and personalized treatment [Paper presentation]. Proceedings - 2020 IEEE 44th Annual Computers, Software, and Applications Conference, COMPSAC 2020 (pp. 1673–1678). <https://doi.org/10.1109/COMPSAC48688.2020.00-15>
- Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*, 13(3), e55. <https://doi.org/10.2196/jmir.1838>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1), 16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Cheerla, N., & Gevaert, O. (2017). MicroRNA based pan-cancer diagnosis and treatment recommendation. *BMC Bioinformatics*, 18(1), 32. <https://doi.org/10.1186/s12859-016-1421-y>
- Chen, M., Yang, J., Zhou, J., Hao, Y., Zhang, J., & Youn, C.-H. (2018). 5G-smart diabetes: Toward personalized diabetes diagnosis with healthcare big data clouds. *IEEE Communications Magazine*, 56(4), 16–23. <https://doi.org/10.1109/MCOM.2018.1700788>
- Chen, T., Zhang, X., Jiang, H., Asaeikheybari, G., Goel, N., Hooper, M. W., & Huang, M.-C. (2018). Are you smoking? Automatic alert system helping people keep away from cigarettes. *Smart Health*, 9–10, 158–169. <https://doi.org/10.1016/j.smhl.2018.07.008>
- Chiang, P., & Dey, S. (2019). Offline and online learning techniques for personalized blood pressure prediction and health behavior recommendations. *IEEE Access*, 7, 130854–130864. <https://doi.org/10.1109/ACCESS.2019.2939218>
- Chiang, P.-H., Wong, M., & Dey, S. (2021). Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure. *IEEE Journal of Translational Engineering in Health and Medicine*, 9, 2700513. <https://doi.org/10.1109/JTEHM.2021.3098173>
- Chin, W. H., Nuo Wi Tay, N., Kubota, N., & Loo, C. K. (2020). A lightweight neural-net with assistive mobile robot for human fall detection system [Paper presentation]. Proceedings of the International Joint Conference on Neural Networks (pp. 1–6). <https://doi.org/10.1109/IJCNN48605.2020.9206637>
- Cho, G., Yim, J., Choi, Y., Ko, J., & Lee, S. H. (2019). Review of machine learning algorithms for diagnosing mental illness. *Psychiatry Investigation*, 16(4), 262–269. <https://doi.org/10.30773/pi.2018.12.21.2>
- Copeland, M., Soh, J., Puca, A., Manning, M., & Gollob, D. (2015). *Microsoft azure*. Apress.
- Coutinho, E., Alshukri, A., de Berardinis, J., & Dowrick, C. (2021). *POLYHYMNIA mood-empowering people to cope with depression*



- through music listening [Paper presentation]. UbiComp/ISWC 2021 - Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers (pp. 188–193). <https://doi.org/10.1145/3460418.3479334>
- da Silva, V. J., da Silva Souza, V., Guimarães da Cruz, R., Mesquita Vidal Martinez de Lucena, J., Jazdi, N., & Ferreira de Lucena Junior, V. (2019). Commercial devices-based system designed to improve the treatment adherence of hypertensive patients. *Sensors*, 19(20), 4539. <https://doi.org/10.3390/s19204539>
- Delmastro, F., Martino, F. D., & Dolciotti, C. (2020). Cognitive training and stress detection in MCI frail older people through wearable sensors and machine learning. *IEEE Access*, 8, 65573–65590. <https://doi.org/10.1109/ACCESS.2020.2985301>
- Dijkhuis, T. B., Blaauw, F. J., van Ittersum, M. W., Velthuisen, H., & Aiello, M. (2018). Personalized physical activity coaching: A machine learning approach. *Sensors*, 18(2), 623. <https://doi.org/10.3390/s18020623>
- D'mello, S. K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys*, 47(3), 1–36. <https://doi.org/10.1145/2682899>
- Edara, D. C., Sistla, V., & Kolli, V. K. K. (2020). Health app recommendation system using ensemble multimodel deep learning. *Journal of Engineering Science and Technology Review*, 13(5), 7–19. <https://doi.org/10.25103/jestr.135.03>
- El Barachi, M., Oroumchian, F., Rauf, R., Khan, U., Al Hassooni, B., Al Basosi, A., & Kazia, S. (2019). *EpiSense: Towards a smart solution for epileptic patients' care* [Paper presentation]. 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech) (pp. 1–6). <https://doi.org/10.23919/SpliTech.2019.8783034>
- Elvitigala, D. S., Scholl, P. M., Suriyaarachchi, H., Dissanayake, V., & Nanayakkara, S. (2021). *StressShoe: A DIY toolkit for just-in-time personalised stress interventions for office workers performing sedentary tasks* [Paper presentation]. Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction. <https://doi.org/10.1145/3447526.3472023>
- Forman, E. M., Goldstein, S. P., Crochiere, R. J., Butryn, M. L., Juarascio, A. S., Zhang, F., & Foster, G. D. (2019). Randomized controlled trial of OnTrack, a just-in-time adaptive intervention designed to enhance weight loss. *Translational Behavioral Medicine*, 9(6), 989–1001. <https://doi.org/10.1093/tbm/ibz137>
- Forman, E. M., Goldstein, S. P., Zhang, F., Evans, B. C., Manasse, S. M., Butryn, M. L., Juarascio, A. S., Abichandani, P., Martin, G. J., & Foster, G. D. (2019). OnTrack: Development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Translational Behavioral Medicine*, 9(2), 236–245. <https://doi.org/10.1093/tbm/iby016>
- Ganju, A., Satyan, S., Tanna, V., & Menezes, S. R. (2021). AI for improving children's health: A community case study. *Frontiers in Artificial Intelligence*, 3, 544972. <https://doi.org/10.3389/frai.2020.544972>
- Ghandeharioun, A., McDuff, D., Czerwinski, M., & Rowan, K. (2019). *EMMA: An emotion-aware wellbeing chatbot* [Paper presentation]. 2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019 (pp. 15–21). <https://doi.org/10.1109/ACII.2019.8925455>
- Google Inc. (n.d.-a). *ML Kit | Google developers*. Retrieved October 31, 2021, from <https://developers.google.com/ml-kit>
- Google Inc. (n.d.-b). *TensorFlow Lite | ML for mobile and edge devices*. Retrieved October 31, 2021, from <https://www.tensorflow.org/lite/>
- Gorbonos, E., Liu, Y., & Hoang, C. T. (2018). *NutRec: Nutrition oriented online recipe recommender* [Paper presentation]. 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2018) (pp. 25–32). <https://doi.org/10.1109/WI.2018.0-111>
- Gu, W. (2017). *PhD forum abstract: Non-intrusive blood glucose monitor by multi-task deep learning* [Paper presentation]. Proceedings - 2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN 2017 (pp. 249–250). <https://doi.org/10.1145/3055031.3055049>
- Gu, Y., Yang, M., Shi, Z., & Wu, Z. (2015). The applications of decision-level data fusion techniques in the field of multiuser detection for DS-UWB systems. *Sensors*, 15(10), 24771–24790. <https://doi.org/10.3390/s151024771>
- Gupta, A., & Katarya, R. (2020). Social media based surveillance systems for healthcare using machine learning: A systematic review. *Journal of Biomedical Informatics*, 108, 103500. <https://doi.org/10.1016/j.jbi.2020.103500>
- Hassan, M. M., Gumaei, A., Aloji, G., Fortino, G., & Zhou, M. (2019). A smartphone-enabled fall detection framework for elderly people in connected home healthcare. *IEEE Network*, 33(6), 58–63. <https://doi.org/10.1109/MNET.001.1900100>
- Hermens, H., Op den Akker, H., Tabak, M., Wijsman, J., & Vollenbroek, M. (2014). Personalized coaching systems to support healthy behavior in people with chronic conditions. *Journal of Electromyography and Kinesiology*, 24(6), 815–826. <https://doi.org/10.1016/j.jelekin.2014.10.003>
- Holt, J. M., Cusatis, R., Asan, O., Williams, J., Nukuna, S., Flynn, K. E., Moore, J., & Crotty, B. H. (2020). Incorporating patient-generated contextual data into care: Clinician perspectives using the consolidated framework for implementation science. *Healthcare*, 8(1), 100369. <https://doi.org/10.1016/j.hjdsi.2019.100369>
- Hu, Y., Wen, G., Liao, H., Wang, C., Dai, D., & Yu, Z. (2021). Automatic construction of chinese herbal prescriptions from tongue images using CNNs and auxiliary latent therapy topics. *IEEE Transactions on Cybernetics*, 51(2), 708–721. <https://doi.org/10.1109/TCYB.2019.2909925>
- Hung, C. Y., Lin, C. H., Chang, C. S., Li, J. L., & Lee, C. C. (2019). *Predicting gastrointestinal bleeding events from multimodal in-hospital electronic health records using deep fusion networks* [Paper presentation]. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS (pp. 2447–2450). <https://doi.org/10.1109/EMBC.2019.8857244>
- IBM. (n.d.). *IBM Cloud*. Retrieved October 31, 2021, from <https://www.ibm.com/cloud>
- Jamil, F., Kahng, H. K., Kim, S., & Kim, D.-H. (2021). Towards secure fitness framework based on IoT-enabled blockchain network integrated with machine learning algorithms. *Sensors*, 21(5), 1640. <https://doi.org/10.3390/s21051640>
- Javed, A. R., Sarwar, M. U., Ur Rehman, S., Khan, H. U., Al-Otaibi, Y. D., & Alnumay, W. S. (2021). PP-SPA: Privacy preserved smartphone-based personal assistant to improve routine life functioning of cognitive impaired individuals. *Neural Processing Letters*, 1–18. <https://doi.org/10.1007/s11063-020-10414-5>
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>
- Jović, A., Brkić, K., & Bogunović, N. (2015). *A review of feature selection methods with applications* [Paper presentation]. 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings (pp. 1200–1205). <https://doi.org/10.1109/MIPRO.2015.7160458>
- Kadri, N., Ellouze, A., & Ksantini, M. (2020). *Recommendation system for human physical activities using smartphones* [Paper presentation]. 2020 2nd International Conference on Computer and Information Sciences, ICCIS 2020. <https://doi.org/10.1109/ICCIS49240.2020.9257671>
- Kajiwar, Y., & Kimura, H. (2019). Object identification and safe route recommendation based on human flow for the visually impaired. *Sensors*, 19(24), 5343. <https://doi.org/10.3390/s19245343>
- Kang, J. J. (2021). *A military human performance management system design using machine learning algorithms* [Paper presentation]. 2021 31st International Telecommunication Networks and Applications Conference (ITNAC) (pp. 13–18). <https://doi.org/10.1109/ITNAC53136.2021.9652140>
- Kariyawasam, R., Nadeeshani, M., Hamid, T., Subasinghe, I., Samarasinghe, P., & Ratnayake, P. (2019). *Pubudu: Deep learning based screening and intervention of dyslexia, dysgraphia and*

- dyscalculia [Paper presentation]. 2019 IEEE 14th International Conference on Industrial and Information Systems: Engineering for Innovations for Industry 4.0, ICIIIS 2019 - Proceedings (pp. 476–481). <https://doi.org/10.1109/ICIIIS47346.2019.9063301>
- Kesavan, R., & Arumugam, S. (2020). Adaptive deep convolutional neural network-based secure integration of fog to cloud supported Internet of Things for health monitoring system. *Transactions on Emerging Telecommunications Technologies*, 31(10), e4104. <https://doi.org/10.1002/ett.4104>
- Khalaf, M., Hussain, A. J., Al-Jumeily, D., Keenan, R., Keight, R., Fergus, P., & Idowu, I. O. (2016). *Applied difference techniques of machine learning algorithm and web-based management system for sickle cell disease* [Paper presentation]. Proceedings - 2015 International Conference on Developments in eSystems Engineering, DeSE 2015 (pp. 231–235). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/DeSE.2015.18>
- Khumrin, P., Ryan, A., Juddy, T., Verspoor, K. (2018). DrKnow: A diagnostic learning tool with feedback from automated clinical decision support. *AMIA Annual Symposium Proceedings. AMIA Symposium*, 2018, 1348–1357.
- Koren, G., Souroujon, D., Shaul, R., Bloch, A., Leventhal, A., Lockett, J., & Shalev, V. (2019). “A patient like me” – An algorithm-based program to inform patients on the likely conditions people with symptoms like theirs have. *Medicine*, 98(42), e17596. <https://doi.org/10.1097/MD.00000000000017596>
- Krishnan, S. P. T., & Gonzalez, J. L. U. (2015). *Building your next big thing with Google Cloud Platform*. Apress. <https://doi.org/10.1007/978-1-4842-1004-8>
- Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal data fusion: An overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9), 1449–1477. <https://doi.org/10.1109/JPROC.2015.2460697>
- Le Glaz, A., Haralambous, Y., Kim-Dufor, D. H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., DeVlyder, J., Walter, M., Berrouiguet, S., & Lemey, C. (2021). Machine learning and natural language processing in mental health: Systematic review. *Journal of Medical Internet Research*, 23(5), e15708. <https://doi.org/10.2196/15708>
- Lee, H., Kang, J., & Yeo, J. (2021). Medical specialty recommendations by an artificial intelligence chatbot on a smartphone: Development and deployment. *Journal of Medical Internet Research*, 23(5), e27460. <https://doi.org/10.2196/27460>
- Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60. <https://doi.org/10.1109/MSP.2020.2975749>
- Li, Z., Das, S., Codella, J., Hao, T., Lin, K., Maduri, C., & Chen, C.-H. (2019). An adaptive, data-driven personalized advisor for increasing physical activity. *IEEE Journal of Biomedical and Health Informatics*, 23(3), 999–1010. <https://doi.org/10.1109/JBHI.2018.2879805>
- Liang, Z., Zhang, G., Huang, J. X., & Hu, Q. V. (2014). *Deep learning for healthcare decision making with EMRs* [Paper presentation]. Proceedings - 2014 IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2014 (pp. 556–559). <https://doi.org/10.1109/BIBM.2014.6999219>
- Liao, W. X., He, P., Hao, J., Wang, X. Y., Yang, R. L., An, D., & Cui, L. G. (2020). Automatic identification of breast ultrasound image based on supervised block-based region segmentation algorithm and features combination migration deep learning model. *IEEE Journal of Biomedical and Health Informatics*, 24(4), 984–993. <https://doi.org/10.1109/JBHI.2019.2960821>
- Liu, B. (2011). Supervised learning. In *Web data mining* (pp. 63–132). Springer. [https://doi.org/10.1007/978-3-642-19460-3\\_3](https://doi.org/10.1007/978-3-642-19460-3_3)
- Loftus, T. J., Filiberto, A. C., Li, Y., Balch, J., Cook, A. C., Tighe, P. J., Efron, P. A., Upchurch, G. R., Rashidi, P., Li, X., & Bihorac, A. (2020). Decision analysis and reinforcement learning in surgical decision-making. *Surgery*, 168(2), 253–266. <https://doi.org/10.1016/j.surg.2020.04.049>
- Lopez-Guede, J. M., Moreno-Fernandez-De-Leceta, A., Martinez-Garcia, A., & Graña, M. (2015). Lynx: Automatic elderly behavior prediction in home telecare. *BioMed Research International*, 2015, 201939. <https://doi.org/10.1155/2015/201939>
- Loram, I., Siddique, A., Sanchez, M. B., Harding, P., Silverdale, M., Kobylecki, C., & Cunningham, R. (2020). Objective analysis of neck muscle boundaries for cervical dystonia using ultrasound imaging and deep learning. *IEEE Journal of Biomedical and Health Informatics*, 24(4), 1016–1027. <https://doi.org/10.1109/JBHI.2020.2964098>
- Luštrek, M., Bohanec, M., Cavero Barca, C., Ciancarelli, M. C., Clays, E., Dawodu, A. A., Derboven, J., De Smedt, D., Dovgan, E., Lampe, J., Marino, F., Mlakar, M., Pioggia, G., Puddu, P. E., Rodríguez, J. M., Schiariti, M., Slapničar, G., Slegers, K., Tartarisco, G., Valič, J., & Vodopija, A. (2021). A personal health system for self-management of congestive heart failure (HeartMan): Development, technical evaluation, and proof-of-concept randomized controlled trial. *JMIR Medical Informatics*, 9(3), e24501. <https://doi.org/10.2196/24501>
- Mahyari, A., & Pirolli, P. (2021). *Physical exercise recommendation and success prediction using interconnected recurrent neural networks* [Paper presentation]. 2021 IEEE International Conference on Digital Health (ICDH) (pp. 148–153). <https://doi.org/10.1109/ICDH52753.2021.00027>
- Martín, P., Sánchez, M., Álvarez, L., Alonso, V., & Bajo, J. (2011). Multi-agent system for detecting elderly people falls through mobile devices. *Advances in Intelligent and Soft Computing*, 92, 93–99. [https://doi.org/10.1007/978-3-642-19937-0\\_12](https://doi.org/10.1007/978-3-642-19937-0_12)
- Martin-Isla, C., Campello, V. M., Izquierdo, C., Raisi-Estabragh, Z., Baeßler, B., Petersen, S. E., & Lekadir, K. (2020). Image-based cardiac diagnosis with machine learning: A review. *Frontiers in Cardiovascular Medicine*, 7(1), 1. <https://doi.org/10.3389/fcvm.2020.00001>
- Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236–1246. <https://doi.org/10.1093/bib/bbx044>
- Mishra, V., Künzler, F., Kramer, J.-N., Fleisch, E., Kowatsch, T., & Kotz, D. (2021). Detecting receptivity for mHealth interventions in the natural environment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2), 1–24. <https://doi.org/10.1145/3463492>
- Mlodzinski, E., Stone, D. J., & Celi, L. A. (2020). Machine learning for pulmonary and critical care medicine: A narrative review. *Pulmonary Therapy*, 6(1), 67–77. <https://doi.org/10.1007/s41030-020-00110-z>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mrozek, D., Koczur, A., & Małysiak-Mrozek, B. (2020). Fall detection in older adults with mobile IoT devices and machine learning in the cloud and on the edge. *Information Sciences*, 537, 132–147. <https://doi.org/10.1016/j.ins.2020.05.070>
- Murnane, E. L., Huffaker, D., & Kossinets, G. (2015). *Mobile health apps: Adoption, adherence, and abandonment* [Paper presentation]. UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015 ACM International Symposium on Wearable Computers (pp. 261–264). <https://doi.org/10.1145/2800835.2800943>
- National Library of Medicine. (n.d.). *PubMed*. Retrieved October 31, 2021, from <https://pubmed.ncbi.nlm.nih.gov/>
- Neloy, A. A., Alam, S., Bindu, R. A., & Moni, N. J. (2019). *Machine learning based health prediction system using IBM cloud as PaaS* [Paper presentation]. Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019 (pp. 444–450). <https://doi.org/10.1109/ICOEI.2019.8862754>
- Nosakhare, E., & Picard, R. (2020). Toward assessing and recommending combinations of behaviors for improving health and well-being. *ACM Transactions on Computing for Healthcare*, 1(1), 1–29. <https://doi.org/10.1145/3368958>



- Paredes, P., Gilad-Bachrach, R., Czerwinski, M., Roseway, A., Rowan, K., & Hernandez, J. (2014). *PopTherapy: Coping with stress through pop-culture* [Paper presentation]. Proceedings - PERVASIVEHEALTH 2014: 8th International Conference on Pervasive Computing Technologies for Healthcare (pp. 109–117). <https://doi.org/10.4108/icst.pervasivehealth.2014.255070>
- Pathinarupothi, R. K., Durga, P., & Rangan, E. S. (2018). Data to diagnosis in global health: A 3P approach. *BMC Medical Informatics and Decision Making*, 18(1), 78. <https://doi.org/10.1186/s12911-018-0658-y>
- Pelle, T., Bevers, K., van der Palen, J., van den Hoogen, F. H. J., & van den Ende, C. H. M. (2019). Development and evaluation of a tailored e-self-management intervention (dr. Bart app) for knee and/or hip osteoarthritis: Study protocol. *BMC Musculoskeletal Disorders*, 20(1), 398. <https://doi.org/10.1186/s12891-019-2768-9>
- Popp, C. J., St-Jules, D. E., Hu, L., Ganguzza, L., Illiano, P., Curran, M., Li, H., Schoenthaler, A., Bergman, M., Schmidt, A. M., Segal, E., Godneva, A., & Sevick, M. A. (2019). The rationale and design of the personal diet study, a randomized clinical trial evaluating a personalized approach to weight loss in individuals with pre-diabetes and early-stage type 2 diabetes. *Contemporary Clinical Trials*, 79, 80–88. <https://doi.org/10.1016/j.cct.2019.03.001>
- Prabhu, G., Kuklyte, J., Gualano, L., Venkataraman, K., Ahmadi, A., Duff, O., Walsh, D., Woods, C., O'Connor, N. E., & Moran, K. (2018). Design and development of the medFit app: A mobile application for cardiovascular disease rehabilitation. In *International Conference on Wireless Mobile Communication and Healthcare (MobiHealth)* (pp. 20–28). Springer. [https://doi.org/10.1007/978-3-319-98551-0\\_3](https://doi.org/10.1007/978-3-319-98551-0_3)
- Priyadarshi, A., & Saha, S. K. (2020). Web information extraction for finding remedy based on a patient-authored text: A study on homoeopathy. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 9(1), 9. <https://doi.org/10.1007/s13721-019-0216-2>
- Qian, K., Janott, C., Schmitt, M., Zhang, Z., Heiser, C., Hemmert, W., Yamamoto, Y., & Schuller, B. W. (2021). Can machine learning assist locating the excitation of snore sound? A review. *IEEE Journal of Biomedical and Health Informatics*, 25(4), 1233–1246. <https://doi.org/10.1109/JBHI.2020.3012666>
- Rabbi, M., Aung, M. H., Zhang, M., & Choudhury, T. (2015). *MyBehavior: Automatic personalized health feedback from user behaviors and preferences using smartphones* [Paper presentation]. UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 707–718). <https://doi.org/10.1145/2750858.2805840>
- Rachakonda, L., Mohanty, S. P., & Kougianos, E. (2020). ILog: An intelligent device for automatic food intake monitoring and stress detection in the IoMT. *IEEE Transactions on Consumer Electronics*, 66(2), 115–124. <https://doi.org/10.1109/TCE.2020.2976006>
- Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48(1), 137–141. <https://doi.org/10.1007/s11747-019-00710-5>
- Rajasekaran, S., & Kousalya, G. (2022). Virtual nursing using deep belief networks for elderly people (DBN-EP). *Computer Systems Science and Engineering*, 42(3), 985–1000. <https://doi.org/10.32604/csse.2022.022234>
- Rasley, J., Rajbhandari, S., Ruwase, O., & He, Y. (2020). *DeepSpeed: System optimizations enable training deep learning models with over 100 billion parameters* [Paper presentation]. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 3505–3506). <https://doi.org/10.1145/3394486.3406703>
- Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3(1), 1–7. <https://doi.org/10.1038/s41746-020-00323-1>
- Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies*, 54, 360–371. <https://doi.org/10.1016/j.jrurstud.2015.09.001>
- Sansrimahachai, W. (2020a). *Personalized walking exercise support system for elderly based on machine learning* [Paper presentation]. JCSSE 2020 - 17th International Joint Conference on Computer Science and Software Engineering (pp. 6–11). <https://doi.org/10.1109/JCSSE49651.2020.9268327>
- Sansrimahachai, W. (2020b). *Personalized walking exercise support system for elderly based on machine learning* [Paper presentation]. JCSSE 2020 - 17th International Joint Conference on Computer Science and Software Engineering (pp. 6–11). <https://doi.org/10.1109/JCSSE49651.2020.9268327>
- Sarwar, M. U., & Javed, A. R. (2019). *Collaborative health care plan through crowdsourcing data using ambient application* [Paper presentation]. Proceedings - 22nd International Multitopic Conference, INMIC 2019. <https://doi.org/10.1109/INMIC48123.2019.9022684>
- Scherzer, C. R., Ranney, M. L., Jain, S., Bommaraju, S. P., Patena, J., Langdon, K., Nimaja, E., Jennings, E., & Beaudoin, F. L. (2020). Mobile peer-support for opioid use disorders: Refinement of an innovative machine learning tool. *Journal of Psychiatry and Brain Science*, 5(1), e200001. <https://doi.org/10.20900/jpbs.20200001>
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 49(9), 1426–1448. <https://doi.org/10.1017/S0033291719000151>
- Shi, J., Zheng, X., Li, Y., Zhang, Q., & Ying, S. (2018). Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*, 22(1), 173–183. <https://doi.org/10.1109/JBHI.2017.2655720>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Sookrah, R., Dhowtal, J. D., & Nagowah, S. D. (2019). *A DASH diet recommendation system for hypertensive patients using machine learning* [Paper presentation]. 2019 7th International Conference on Information and Communication Technology, ICoICT 2019 (pp. 1–6). <https://doi.org/10.1109/ICoICT.2019.8835323>
- Sowah, R. A., Bampoe-Addo, A. A., Armoo, S. K., Saalia, F. K., Gatsi, F., & Sarkodie-Mensah, B. (2020). Design and development of diabetes management system using machine learning. *International Journal of Telemedicine and Applications*, 2020, 8870141. <https://doi.org/10.1155/2020/8870141>
- Spanakis, G., Weiss, G., Boh, B., Lemmens, L., & Roefs, A. (2017). Machine learning techniques in eating behavior e-coaching: Balancing between generalization and personalization. *Personal and Ubiquitous Computing*, 21(4), 645–659. <https://doi.org/10.1007/s00779-017-1022-4>
- Stafford, I. S., Kellermann, M., Mossotto, E., Beattie, R. M., MacArthur, B. D., & Ennis, S. (2020). A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digital Medicine*, 3(1), 1–11. <https://doi.org/10.1038/s41746-020-0229-3>
- Stamate, C., Magoulas, G. D., Kueppers, S., Nomikou, E., Daskalopoulos, I., Jha, A., Pons, J. S., Rothwell, J., Luchini, M. U., Moussouri, T., Iannone, M., & Roussos, G. (2018). The cloudUPDRS app: A medical device for the clinical assessment of Parkinson's disease. *Pervasive and Mobile Computing*, 43, 146–166. <https://doi.org/10.1016/j.pmcj.2017.12.005>
- Stamate, C., Magoulas, G. D., Kueppers, S., Nomikou, E., Daskalopoulos, I., Luchini, M. U., Moussouri, T., & Roussos, G. (2017). *Deep learning Parkinson's from smartphone data* [Paper presentation]. 2017 IEEE International Conference on Pervasive Computing and Communications, PerCom 2017 (pp. 31–40). <https://doi.org/10.1109/PERCOM.2017.7917848>
- Stark, B., & Samarah, M. (2019). *Mac7: Adaptive Smart Toothbrush* [Paper presentation]. Proceedings - 2018 International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2018 (pp. 153–158). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/SDPC.2018.8664828>

- Suh, M., Nahapetian, A., Woodbridge, J., Rofouei, M., & Sarrafzadeh, M. (2012). Machine learning-based adaptive wireless interval training guidance system. *Mobile Networks and Applications*, 17(2), 163–177. <https://doi.org/10.1007/s11036-011-0331-5>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. (2014). *Going deeper with convolutions* [Paper presentation]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June-2015 (pp. 1–9).
- Thieme, A., Belgrave, D., & Doherty, G. (2020). Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer-Human Interaction*, 27(5), 1–53. <https://doi.org/10.1145/3398069>
- Thomsen, K., Iversen, L., Titlestad, T. L., & Winther, O. (2020). Systematic review of machine learning for diagnosis and prognosis in dermatology. *Journal of Dermatological Treatment*, 31(5), 496–510. <https://doi.org/10.1080/09546634.2019.1682500>
- Thyde, D. N., Mohebbi, A., Bengtsson, H., Jensen, M. L., & Mørup, M. (2021). Machine learning-based adherence detection of type 2 diabetes patients on once-daily basal insulin injections. *Journal of Diabetes Science and Technology*, 15(1), 98–108. <https://doi.org/10.1177/1932296820912411>
- Tolkachev, A., Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., & Ibragimov, B. (2021). Deep learning for diagnosis and segmentation of pneumothorax: The results on the Kaggle competition and validation against radiologists. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1660–1672. <https://doi.org/10.1109/JBHI.2020.3023476>
- Triantafyllidis, A. K., & Tsanas, A. (2019). Applications of machine learning in real-life digital health interventions: Review of the literature. *Journal of Medical Internet Research*, 21(4), e12286. <https://doi.org/10.2196/12286>
- Tuti, T., Paton, C., & Winters, N. (2020). *Learning to represent health-care providers knowledge of neonatal emergency care: Findings from a smartphone-based learning intervention targeting clinicians from LMICs* [Paper presentation]. ACM International Conference Proceeding Series (pp. 320–329). <https://doi.org/10.1145/3375462.3375479>
- Van Woensel, W., Roy, P. C., Abidi, S. S. R., & Abidi, S. R. (2020). Indoor location identification of patients for directing virtual care: An AI approach using machine learning and knowledge-based methods. *Artificial Intelligence in Medicine*, 108, 101931. <https://doi.org/10.1016/j.artmed.2020.101931>
- Venne, J., Busshoff, U., Poschadel, S., Menschel, R., Evangelatos, N., Vysyaraju, K., & Brand, A. (2020). International consortium for personalized medicine: An international survey about the future of personalized medicine. *Personalized Medicine*, 17(2), 89–100. <https://doi.org/10.2217/pme-2019-0093>
- Vicente, A. M., Ballensiefen, W., & Jönsson, J. I. (2020). How personalised medicine will transform healthcare by 2030: The ICPMed vision. *Journal of Translational Medicine*, 18(1), 180. <https://doi.org/10.1186/s12967-020-02316-w>
- Visalakshi, S., & Radha, V. (2015). *A literature review of feature selection techniques and applications: Review of feature selection in data mining* [Paper presentation]. 2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014. <https://doi.org/10.1109/ICCIC.2014.7238499>
- Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., & Weidt, S. (2016). Mobile sensing and support for people with depression: A pilot trial in the wild. *JMIR mHealth and uHealth*, 4(3), e111. <https://doi.org/10.2196/mhealth.5960>
- Wang, L., Fan, R., Zhang, C., Hong, L., Zhang, T., Chen, Y., Liu, K., Wang, Z., & Zhong, J. (2020). Applying machine learning models to predict medication nonadherence in Crohn's disease maintenance therapy. *Patient Preference and Adherence*, 14, 917–926. <https://doi.org/10.2147/PPA.S253732>
- Wang, S., Sporrel, K., van Hoof, H., Simons, M., de Boer, R. D. D., Ettema, D., Nibbeling, N., Deutekom, M., & Kröse, B. (2021). Reinforcement learning to send reminders at right moments in smartphone exercise application: A feasibility study. *International Journal of Environmental Research and Public Health*, 18(11), 6059. <https://doi.org/10.3390/ijerph18116059>
- Waring, J., Lindvall, C., & Umeton, R. (2020). Automated machine learning: Review of the state-of-the-art and opportunities for health-care. *Artificial Intelligence in Medicine*, 104, 101822. <https://doi.org/10.1016/j.artmed.2020.101822>
- Wittig, M., & Wittig, A. (2018). *Amazon web services in action* (2nd ed.). Simon and Schuster.
- Xu, P., Chen, D., Liu, X., & Loo, J. (2019). *Image-based dietary assessment system for chinese children* [Paper presentation]. Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018 (pp. 5471–5473). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/BigData.2018.8622579>
- Yang, J., Xiao, W., Lu, H., & Barnawi, A. (2020). Wireless high-frequency NLOS monitoring system for heart disease combined with hospital and home. *Future Generation Computer Systems*, 110, 772–780. <https://doi.org/10.1016/j.future.2019.11.001>
- Yang, W., Chuah, M. C., & Du, L. (2020). *M&M: Deep learning aided multi-facet mental health support tool for college students* [Paper presentation]. Proceedings of Deep Learning for Wellbeing Applications Leveraging Mobile Devices and Edge Computing (pp. 10–15). <https://doi.org/10.1145/3396868.3400898>
- Yates, D., & Islam, M. Z. (2019). Readiness of smartphones for data collection and data mining with an example application in mental health. In *Australasian Conference on Data Mining (AusDM)* (pp. 235–246). Springer. [https://doi.org/10.1007/978-981-15-1699-3\\_19](https://doi.org/10.1007/978-981-15-1699-3_19)
- Ying, Z., Cao, S., Xu, S., Liu, X., Lyu, L., Chen, C., & Wang, L. (2021). Privacy-preserving optimal insulin dosing decision. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021 (pp. 2640–2644). <https://doi.org/10.1109/icassp39728.2021.9414807>
- Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., ... Segal, E. (2015). Personalized nutrition by prediction of glycemic responses. *Cell*, 163(5), 1079–1094. <https://doi.org/10.1016/j.cell.2015.11.001>
- Zeng, J. (2020). Developing a machine learning tool for dynamic cancer treatment strategies. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(10), 13742–13743. <https://doi.org/10.1609/aaai.v34i10.7143>
- Zhang, J. (2010). Multi-source remote sensing data fusion: Status and trends. *International Journal of Image and Data Fusion*, 1(1), 5–24. <https://doi.org/10.1080/19479830903561035>
- Zhang, Y., Chen, K., Weng, Y., Chen, Z., Zhang, J., & Hubbard, R. (2022). An intelligent early warning system of analyzing Twitter data using machine learning on COVID-19 surveillance in the US. *Expert Systems with Applications*, 198, 116882. <https://doi.org/10.1016/j.eswa.2022.116882>
- Zhou, X., Li, Y., & Liang, W. (2020). CNN-RNN based intelligent recommendation for online medical pre-diagnosis support. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(3), 912–921. <https://doi.org/10.1109/TCBB.2020.2994780>
- Zhu, T., Li, K., Herrero, P., & Georgiou, P. (2021). Deep learning for diabetes: A systematic review. *IEEE Journal of Biomedical and Health Informatics*, 25(7), 2744–2757. <https://doi.org/10.1109/JBHI.2020.3040225>
- Zingg, A., Singh, T., & Myneni, S. (2021). Analysis of online peripartum depression communities: Application of multilabel text classification techniques to inform digitally-mediated prevention and management. *Front Digit Health*, 3, 653769. <https://doi.org/10.3389/fdgth.2021.653769>

## About the Authors

**Oladapo Oyeboade** is a Ph.D. candidate at Dalhousie University. His research interests include Human-Computer Interaction, Persuasive Technology, Adaptive Systems, Affective Computing, Artificial Intelligence, Digital Health, and Health Informatics. He has published over 30 peer-reviewed research articles. He is currently designing and developing technologies to tackle health-related (including mental health) issues.

**Jonathon Fowles** is the professor with the School of Kinesiology and Director of the Centre of Lifestyle Studies at Acadia University, Chair of the National Advisory Council for Exercise is Medicine Canada, and Fellow of the Canadian Society for Exercise Physiology. He believes technology can enhance health and wellness.

**Darren Steeves** is an Adjunct Professor at the School of Health and Human Performance, Dalhousie University. His passion is to help people through living his values of collaboration, innovation, curiosity, and integrity. He has 25-year background in health and performance, consulting with top executives, Olympic medalists, and world champion athletes.

**Rita Orji** is a Canada Research Chair in Persuasive Technology and a Computer Science Professor at Dalhousie University where she directs Persuasive Computing Lab. Her research is at the intersection of technology and human behaviour, focusing on investigating user-centered approaches to designing technologies that improve lives and promote desirable behaviours.