Development of Novel Big Data Analytics Framework for Smart Clothing

MOMINUL AHSAN [ID], SIEW TEAY HON [ID], AND ALHUSSEIN ALBARBAR [ID]

Smart Infrastructure and Industry Research Group, Department of Engineering, Manchester Metropolitan University, Manchester M1 5GD, U.K.

Corresponding author: Mominul Ahsan (m.ahsan@mmu.ac.uk)

ABSTRACT Recent advances in micro electro-mechanical systems (MEMS) have produced wide variety of wearable sensors. Owing to their low cost, small size and interfacability, those MEMS based devices have become increasingly commonplace and part of daily life for many people. Large amount of data from heart and breath rates to electrocardiograph (ECG) signals, which contain a wealth of health-related information, can be measured. Hence, there is a timely need for novel interrogation and analysis methods for extracting health related features from such a Big Data. In this paper, the prospects from smart clothing such as wearable devices in generating Big Data are critically analyzed with a focus on applications related to healthcare, sports and fashion. The work also covers state-of-the-art data analytics methods and frameworks for health monitoring purposes. Subsequently, a novel data analytics framework that can provide accurate decision in both normal and emergency health situations is proposed. The proposed novel framework identifies and discusses sources of Big Data from the human body, data collection, communication, data storage, data analytics and decision making using artificial intelligence (AI) algorithms. The paper concludes by identifying challenges facing the integration of Big Data analytics with smart clothing. Recommendation for further development opportunities and directions for future work are also suggested.

INDEX TERMS Wearable technology, smart clothing, big data, data analytics, artificial intelligence (AI), machine learning (ML) and decision-making.

I. INTRODUCTION

A. SMART CLOTHING AS A WEARABLE DEVICE

Wearable technologies (WT) have made significant influence on people's way of modern life due to globalisation, Internet-of-Things (IoT) paradigm and advanced communication technologies. Wearable devices including smart watch, fitness band, smart ring etc. are gaining popularity to monitor health condition for the people of all age groups, athletes and sportsmen. WT found application in numerous areas such as healthcare, sports, fitness and wellbeing, glamour and fashion. Diverse vital signs are collected by the wearable devices to assess body fitness through vibrant data analytics techniques. Conventional WT are not fully capable of efficiently monitoring and tracking human health due to their rigidity and bulkiness. Poor design and integration techniques of wearable devices with the human body cause discomfort when used for long term on a daily basis [1]. To overcome this limitation, Smart clothing idea has evolved by integrating

wearable technologies within cloths such as coat, pant, shirts, leggings, socks etc. Smart clothing covering maximum part of the body creates a great opportunity to continuously generate huge amount of physiological data compared to the other wearable devices. The advent of flexible electronics and their seamless integration make smart clothing to bring in direct contact with human body [2].

B. BIG DATA

Big Data generated with a mix of structured and unstructured formats are enormously huge in size and multifarious. To collect, process, analyse and store Big Data accurately are the puzzling tasks for traditional software system, algorithms, and data repositories [3], [4]. The Big Data is increasing dramatically and has already been expanded to an unimaginable scale. This could be due to the development of internet, mobile internet, IoT and social media. It has four main characteristics named as 4Vs: volume, veracity, velocity and variety (Fig. 1). Data volume refers to collection of huge data in terabytes level. Such massive-scale data are required to manage, store and analyse with highly effective data
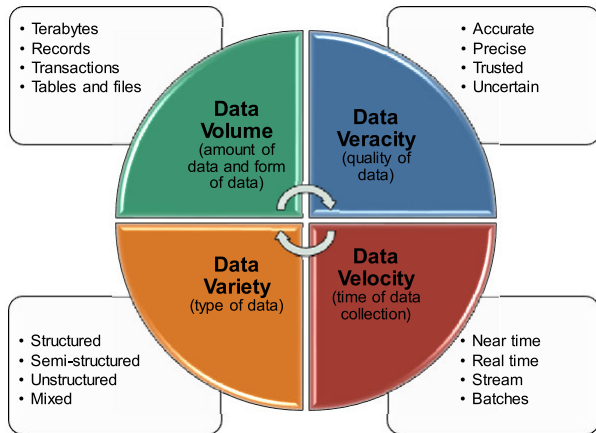
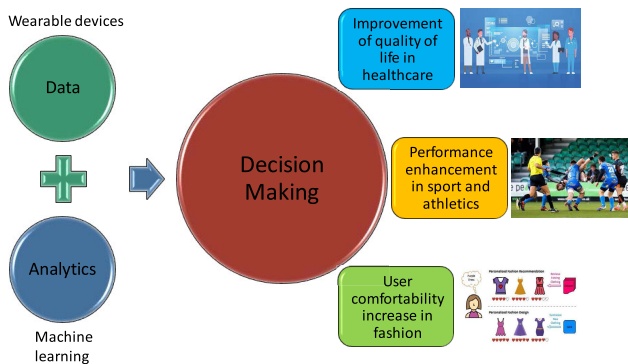**FIGURE 1.** Characteristics of Big Data (adapted from [8]).

FIGURE 2. Decision making with Big Data and data analytics.

processing techniques. The high volume of data generated from human body with the use of modern health IoT system drives the obligation of cost-effective data analysis and decision making. Thus, public health improvement and disease prediction can be performed efficiently by extracting valuable insights from the data [5]. Velocity represents the exponential growth of data generation and collection continuously such as real-time heart rate streaming data [6]. Variety of Big Data signifies the data in structured or unstructured forms. For example, structured data in healthcare contains classified terminologies. They include numerous diseases, symptoms, diagnosis report, pathology information, analysis and patient information such as admission details, medicine and payment information. On the contrary, unstructured data has no inherent structure such as clinical letters, biomedical writings, discharge information etc. Finally, veracity of Big Data represents the degree to which the data are consistent, accurate, precise and trusted [7]. For example, inconsistent blood pressure data could be generated when different equipment, people and measurement procedures are employed.

C. IMPACT OF BIG DATA

The application of smart clothing in healthcare, sports and athletics, and fashion industries for health monitoring (Fig.2) can generate data at a tremendous rate, but processing of the large data is becoming challenging at the same time. Therefore, new strategies need to be adopted to organise,

manage and process them more efficiently in order to derive meaningful information and to use in practical applications. Data analytics by machine learning (ML) techniques have diverse features and capability to extract meaningful hidden information that helps in making effective and timely decision.

1) HEALTHCARE: IMPROVEMENT OF QUALITY OF LIFE

The economics of healthcare systems have brought attention due to dynamics of global demographics. The investment on healthcare services is one of the important priorities worldwide. It is estimated that healthcare costs will increase from 20% to 30% of GDP in some countries by 2050 [9], [10]. The motivation of cost reduction is leading to change the attention of the approaches for developing predictive smart healthcare system.

Medical practitioners gain real insight about the progress of the treatments through observing massive amount of data generated from various phases of diagnosis and treatment plans. For example, the phases could include clinical services, reports of lab test, clinical notes, imaging data and patient's behaviour under different scenarios using wearable devices. Clinical Disease Repository (CDR) as a source of Big Data helps in improving public health surveillance by offering quicker response through effective analysis of disease patterns. In addition, other advantages include facilitating physicians to track the usage of drugs, monitoring the patient's health condition at any stage. of Big Data. Hospital operators can extremely rely upon the outcome of Big Data sources for managing the patient's experiences and to enhance resources. Besides, Big Data helps in resource allocation and deployment of optimization techniques to fill up manpower in clinical sectors. Pharmaceutical and clinical researchers can build predictive models using healthcare Big Data to realise effective drug designs. Pharmaceutical companies integrate and analyse many forms of health-related Big Data in order to build end-to-end product solutions coupling in-memory computing technologies during manufacturing [11]. Big Data (vital signs, age, gender, family history, income, nature of job) release new analytical possibilities for healthcare insurers by introducing novel health plans with minimal premium cost. Appropriate predictive modelling developed through Big Data associated with IoT based techniques can help in predicting reliable claims and rare outliers. This reduces cost of abuse through analysing unstructured data from claim history and customer behaviour data [12]. The diverse Big Data analytics and associated decision-making approach can improve the quality of life. The improvement is happened through continuous monitoring of vital signs and effective decision making for taking appropriate actions before occurring any health damage.

2) SPORTS AND ATHLETICS: PERFORMANCE ENHANCEMENT

Sports and athletics programs have become popular in the arena of entertainment due to increasing demand of digital

broadcasting applications [13]. With the development of wearable technologies, sports data is being generated in a huge scale. Data accessibility and data interpretation are the main challenges when dealing with those collected from athletics. In addition, precise mining of large data is complicated and costly for implementing analytical techniques [14]. Big Data in sports and athletics with appropriate analytical techniques using ML have great impact to mine the data accurately by extracting precise information for making effective decision on performance enhancement.

3) FASHION: USER COMFORTABILITY INCREMENT

Big Data has been a prominent term in fashion world since the last decade [15]. Market trend analysis, customer behaviour analysis and supply chain management are the key activities using Big Data in fashion industry. For example, excessive stock sometimes causes huge loss for fashion industry due to the changes of trend. Big Data analytics forecasts the trend of customer behaviour and emotions that can help in identifying the needs of customer and reduces unnecessary production of similar items. Collaborative filtering technique endorses customer demand based on the preferences of different age's groups. However, Big Data in fashion faces challenges due to lack of technical competence, opportunity to access fashion data etc. [16].

D. CONTRIBUTIONS OF THIS ARTICLE

Majority of the frameworks for health monitoring using wearable technologies focused on data collection and processing techniques but still there is a lack of application of data analytics for making decision on health status through developing user interface. The scientific value of the entire paper is about development of complete big data analytics and decision-making framework particularly for smart clothing, which has not been attempted in the literature. Critical performance assessment of ML algorithms from state of the art has been conducted and identified suitable algorithms for Big data analytics. Furthermore, the challenge section has been developed based on rigorous analysis and limitation of existing research to give direction to researchers for future development in this area.

E. PAPER ORGANISATION

This article introduces smart clothing as a wearable device with the sources and opportunities of Big data from human body and its impact on human lifestyle in Section I. A novel data analytics framework is presented in Section II including bio-signal sources, data acquisition, Big data analytics and decision making using smart clothing. Overall challenges of Big Data analytics and subsequent future development are reported in Section III. Finally, key conclusions are drawn in Section IV.

II. BIG DATA COLLECTION AND ANALYSIS FRAMEWORK

Big Data analytics encompasses quick data collection, processing ability, technologies and approaches for managing
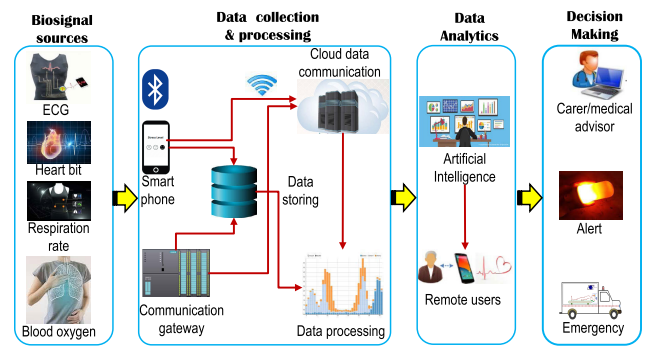

FIGURE 3. Proposed Big Data analytics framework using smart clothing.

the data for the purpose of analysis, modelling, prediction and hypothesis verification [17]. The infrastructure of Big Data analytics includes servers and storage systems, cloud service, networking kit, parallel and distributed file systems and data-mining software [18], [19]. Recent research efforts have presented data analytics frameworks for healthcare applications. Data analytics frameworks for healthcare are summarised in Table 1 including tools and techniques used, framework characteristics and limitations.

Recent advancement in smart clothing in the areas of healthcare, sports and athletics, and fashion creates an opportunity for new era of Big Data [29]. A novel data analytics approach for smart clothing has been developed as shown in Fig. 3 that has merged the opportunities of all existing frameworks. For smart clothing, data collection entails capturing raw bio signals from human body such as ECG, heart bit, blood oxygen, knee movement etc. The collected raw data is stored in local hardware-based system or cloud-based system and pre-processed. The processing is conducted through eliminating outlier, transforming, categorizing and turning it in a structured form such as CSV or text file. The processed data are then analysed applying advanced techniques such as ML. Finally, decision is made by the care supporter or medical advisor based on the outcome from the analysis. The whole section 2 has been dedicated for elaboration of various parts of the framework in the light of current literature.

A. BIOSIGNALS AND BIG DATA IN SMART CLOTHING

Current multidisciplinary research in wearable technologies integrated with smart clothing, wireless and mobile information communication technologies (ICT) has influenced monitoring and generating health indicator data (Fig. 4). For example, electrical activity of muscle can create EMG signals which are detected using skin electrodes. Electrical activity of brain creates EEG where scalp-placed electrodes are positioned to collect that signal [30], [31]. Accelerometer sensor can detect posture or limb movements which are responsible for generating activity signals, mobility and fall. Furthermore, heartbeat inspiration and expiration generate heart sound and respiration rate signals respectively which are collected by phonograph, piezoelectric and piezo

TABLE 1. Examples of existing Big Data analytics frameworks for healthcare.

Data source	Tools/techniques	Data processing and Data analytics	Decision Making	Ref
Healthcare data (external and internal)	middleware and data warehousing operations	Data source and transformation layer, platform and analytics layer of Big Data	Emphasizes the theoretical aspects. Decision making was not elaborated in detail	[6]
Disease risk profiles for individuals	Collaborative filtering	Similar patient identification and prediction of diseases	Lack of details in data analytics and decision-making process	[20]
ECG, respiration and motion data	Hadoop platform, Map-Reduce programming model	Extract feature values from raw data	Analysis and decision making in detail	[21]
Images of Chest X-ray	Pattern analysis	Data capturing, storing, analyzing, pre-processing and decision support	Lack of details in decision making	[22]
Pharmacy, hospital, and lab	encryption and authentication techniques	Integration and interoperability, availability, and sharing of healthcare data	Lack of assessment of the proposed security policies.	[23]
Pathology report, patient record	Offline Hadoop cluster and online distributed search cluster	Processing user query, search node, data analysis and load balanced cluster	Analyses only the historical medical records of patients	[24]
Electronic Healthcare Data (EHR),	Web Ontology Language (OWL)	Data collection, data classification and visualization	Only presented semantic technologies in biomedical research	[25]
Patient profile	cyber-physical system	Data sensing and storage, processing	Lack of analytical capability for Big data	[10]
Electronic health records	Three tier sensitive system	Data collection, data analytics and end users	Decision making has not been presented in detail	[26]
Electronic copy of test reports	Intelligent agents, ML and text mining	Data source, data analytics visualization, prediction and evaluation	Health decision outcome has not been mentioned properly	[27]
Medical equipment, supplies, medication and patient information	RFID technologies	monitoring, tracking and privacy	The system has focused only privacy issue	[28]

resistive sensor. Woven metal electrodes are also used to detect Galvan skin response through conductivity of skin electrical. Temperature of body can also be measured through temperature probe or skin patch during sports and fitness exercise. Glucose in blood produces glucose signals that
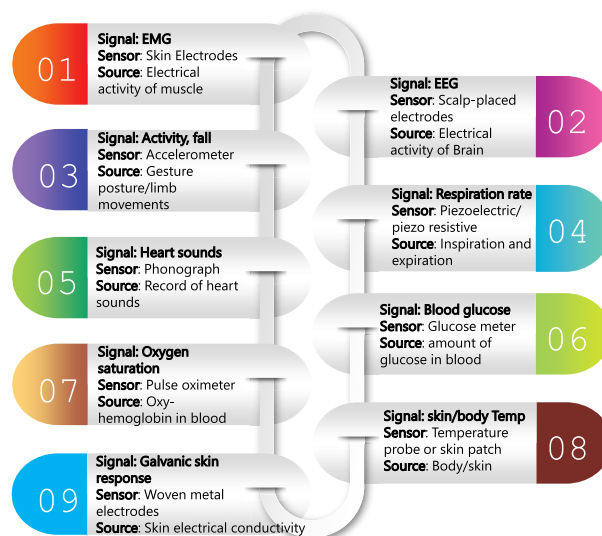
FIGURE 4. Sources of Big Data from human body by smart clothing (information taken from [31]).

can be detected and collected using Glucose meter sensor. In addition, Oxy hemoglobin is a part of blood where oxygen saturation signals are picked up through pulse oximeter sensors.

Researchers are working towards developing wearable biosensors having flexibility, stretchability and physical robustness in order to capture the bio-signals. In general, current sensors on smart clothing are categorised based on piezoresistivity, capacitance and piezoelectricity. Recently, the sensors materials are being developed for smart clothing based on carbon, metal and polymer. Graphite, graphene and carbon nanotube have recently become attractive choices as carbon-based sensor materials whereas nanowire (AgNW) and liquid metal (NiO, Ga & its alloys) are as metal based sensors materials. Poly (3, 4-ethylenedioxythiophene): poly(styrenesulfonate) (PEDOT:PSS), Polyvinylidene difluoride (PVDF) and Ionic liquid (IL) salt are used as popular materials for developing polymer-based sensors [9]. Further details about the sensor development for smart clothing are presented in a number of review papers [2], [9].

In general, wearable sensors have great mobility due to it's miniaturized size. Wearable sensors can be compared with the traditional wired sensors deployed in the environmental monitoring such as floor pressure sensors, temperature sensors, or cameras. Environment sensors are placed in a fixed location, therefore, mobility is a great issue for them. Compared to wearable sensors, environment sensors are costly due to it high instrumentation, installation and periodic maintenance cost. The data security of the wearable sensors has significantly improved due to technological development such as Integrated Circuit Metric (ICMetric) [32].

B. DATA COLLECTION AND PRE-PROCESSING
1) DATA COLLECTION, TRANSMISSION AND STORAGE
With the development of technology, smart clothing is becoming the part of the future Internet-of-Smart-Clothing [33].
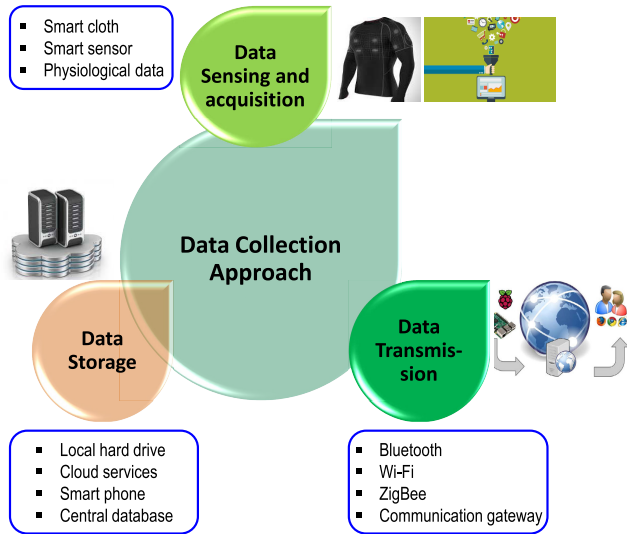
FIGURE 5. Physiological data collection, transmission and storage approach using smart clothing.

TABLE 2. Summary of health monitoring related Big Data [36].

Data type	Data name	Data description	Data acquisition	Ref.
Clinical Big Data	Electronic health record (EHR) and electronic medical record (EMR)	medical and health information	computer tomography (CT), radiography, magnetic resonance imaging (MRI), etc.	[37]
	Personal health record (PHR)	health-related data and information of patients	reports and laboratory test results, chronic diseases, family history, etc.	[38]
	Medical images	visual information of interior human body	radiography, MRI, nuclear medicine, elastography, tactile imaging	[39]
Big Data in health and behavior	Electrocardiogram	electrical graph recording heartbeat activity	electrocardiograph (ECG) signal	[39]
	Vitals	Blood pressure, temperature, pulse, respiratory rate, etc.	temperature, pulse, respiratory rate, and blood pressure	[38]
	Omics data	biology information data	RNA seq, metabolomics – Nuclear Magnetic Resonance (NMR) etc.	[38]
	Molecular biology experiment	Interactions between DNA, RNA, proteins etc.	polymerase chain reaction (PCR), probing	[36]
	Human body samples	cells, tissues, and organs	cells, tissues etc.	[36]

The advent of conductive fabrics and flexible printed electronics enables seamless integration between sensors and textiles [34], [35]. Conductive fabrics and advanced sensors communicate among different subsystems of a smart clothing. Fig. 5 presents a data collection and storage approach, which covers three steps: data sensing and acquisition, communication and storage.

Physiological data such as movement data are loaded from sensors for example, after sitting, standing, walking by embedding accelerometer or gyroscope sensors with smart clothing such as smart leggings. Multiple wearable sensors have the diverse capability of measuring physiological vital signs. For example, skin temperature data is connected to a health-tracking network for the acquisition of health condition or fitness data.

All data acquisition (DAQ) systems are developed with sensor, signal conditioning, and analog-to-digital converter (ADC). Data acquisition system can be developed further with combination of DAQ hardware, sensors and actuators, and a DAQ software of computer. Big Data has already been collected from the fields related to healthcare and fitness including medical diagnosis from imaging data, quantifying lifestyle data such as nutrition, physical activity, sleeping etc. Table 2 summarises examples of Big Data for health monitoring.

Data transmission process sends digital or analog data over a communication medium through computer and communication network. It enables transferring data and making communication of the devices as point-to-point, point-to-multipoint and multipoint-to-multipoint environment. The monitored physiological data are transmitted from the human body by various data transmission components. Data acquisition using sensors platform is employed with a short-range radio where ZigBee or low-power Bluetooth acts as transferring sensor data from users to remote center or cloud. Collected data are then transmitted and stored to a local service center using a smartphone's Wi-Fi or cellular data connection. Internet of Things (IoT)-based architecture provides opportunity to transmit and access each individual sensor's data through the internet connection [40]–[42].

Data transmission is conducted from source to local storage centre (computer/cloud system) through Central Processing Units (CPUs), microcontrollers, Field-Programmable Gate Arrays (FPGAs), or System-on-Chips (SOCs). A communications gateway can exchange information with smart clothing through cloud servers or a block chain [43]. Body Area Network (BAN) provides a short data communication range through distributing electronic sensors embedded into smart cloth. Besides, a Local Area Network (LAN) will collect data from the smart clothing and sending them to a cloud or remote server wirelessly. Mesh network communications can be established to communicate with each objects and machine, which are setup surround them. Furthermore, Wide Area Network (WAN) works like the internet that covers a wide range of communication [2].

Diverse technologies are incorporated with smart clothing as the medium of communication. They include 3G/4G/5G, ultrasounds, infrared, ZigBee, Long-Range Wide Area Network (LoRa WAN), Ultra-Wide Band (UWB), Wireless HART, SigFox, ANT+, Weightless-P, Wi-SUN or IEEE
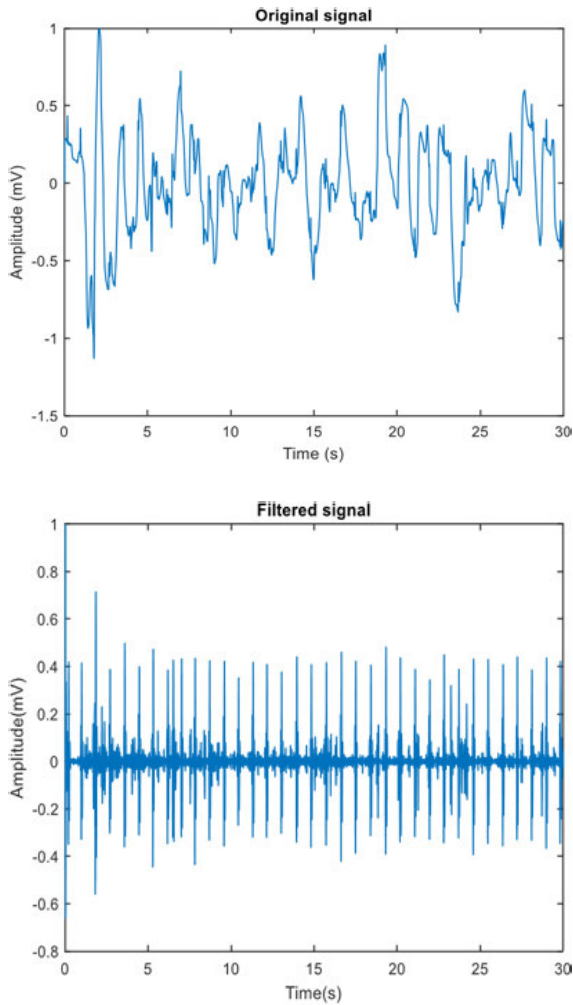
FIGURE 6. Current communications technologies used in smart clothing (information taken from [2]).

FIGURE 7. Data pre-processing steps by smart clothing.

802.11ah [2]. Fig. 6 provides examples for communications technologies that are embedded in smart clothing.

Data storage is a process of archiving data in different forms through different types of medium such as computer or other electronic devices. Big Data are temporarily stored in the memory chips and video card. Sometimes signal data are transmitted and stored in a cloudlet, while the opportunity of data storage is not available in cloud system. Desktop computer acts as cloudlet sometimes for storing the data through Wi-Fi network. Cloudlet transmits aggregated data further to the cloud while a temporary lack of connectivity or energy exists [44]. The data collected by smart clothing can also be stored in local cloud-based system or SD cards. For, IoT devices the data are stored in an external server. Smart cloths are connected to an internal LAN using wireless interface that enables interacting with it. However, a software-based back-end storage system is developed as data storage. Block chain acts as digital information distributor that can store and transfer data through secured way among entities [40], [9].

2) DATA PRE-PROCESSING

Before Big Data analysis, monitored data are stored either in local storage or cloud-based service center and pre-processed where Fig. 7 presents different data processing steps. The

cleaned data are then stored in the local database either in computer or cloud system for further processing such as aggregation of Big Data analysis. Bio-signals are usually noisy due to reading extremely weak signals during health monitoring. The noise of low and high frequency is reduced without changing the desired signal using Local Outlier Factor (LOF), Chi-Square statistical method, AJAX etc. [45], [46].

Data filtering process enables removing some frequencies from collected signals. Outliers can also be eliminated from the data, which exist outside of data clusters. Sometimes missing values are also removed from the data sets. During data fusion process varying degree of noise is removed on a defined criterion that can enhance model robustness. Various filtering process are used including, finite impulse response FIR), Kalman filtering collaborative filtering, low-pass filter, high-pass filter, band-pass filter, notch filter, etc. [47]. For example, Kalman filter-based methods are limited to linear system where application consequence may be nonlinear. It can take advantages of sigma points to approximate statistical properties accurately from random variable. Kalman Filter (KF)-based noise optimization scheme can protect privacy issue by improving utility of sanitized data. Furthermore, an accelerometer sensor can measure vibration signal in the wearable device. In that case a high-pass filter can eliminate low frequency noise from patient movement during data analytics [48].

Meaningful and accurate data can be extracted using data filtering by reducing noise from the monitored original physiological signals. A wearer (smart t-shirt) trials were successfully carried out to gather ECG signals as a part of our preliminary work. An adaptive filtering algorithm has been designed using Butterworth high pass filter to rectify the gathered signals. Fig. 8 shows an example of raw ECG signal and corresponding filtered signal. The results obtained from the experiments showed that ECG signals could be portably obtained and filtered using the developed adaptive algorithm to determine the heart rate in Beats Per Minute (BPM). A BPM detection difference of only ±5.09% from the original signal was successfully achieved. It was also concluded that additional pressure at the interface between the smart cloth and body did not affect the retrieval of the signal, as along as enough direct contact was established in the first place. This knowledge can be implemented within the next generation of smart cloths for health monitoring.

Data Segmentation is a process for dividing and grouping similar data together based on the monitored parameters to receive a deeper insight of the data. There are different types of data segmentation techniques on large data. K-Means Clustering can identify similar groups of respondents based on selected characteristics. Latent Class Cluster Analysis is another probability modeling technique that maximises the overall fit of a model to the data. In addition, Factor segmentation is used to analyse factor- or form groups of attributes that represents common theme in the data [49]. For large amount of monitored data segmentation

FIGURE 8. An example of original ECG signal collected using smart cloth and filtered signal by adaptive filtering algorithm.
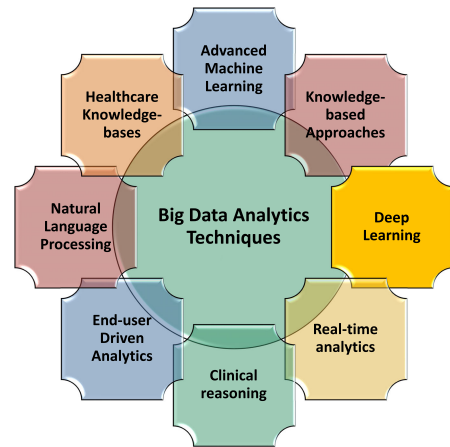

FIGURE 9. Big Data analytics techniques (information taken from [29]).
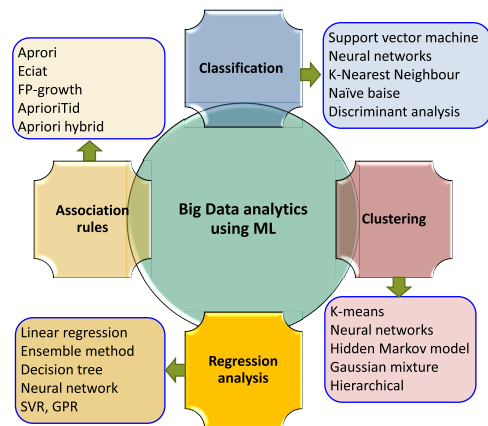
FIGURE 10. Machine learning techniques for Big Data analytics (information taken from [36]).

can separate them into distinct groups. In healthcare, ECG signals are analysed using heartbeat segmentation. Various methods have been applied for detecting the accuracy of QRS complex [50]. For example, the morphology of QRS complex along ECG is segmented using Hilbert transform. Making judgmental splits on values of explanatory variables is widely used as segmentation approach. Prescribed binning criteria method known as descriptive segmentation approach is used in healthcare. The method can channel patients into one of a set of pre-defined cohorts according to their attributes using an off-the-shelf set of binning rules. Classification and Regression Tree (CART) algorithm analysis is applying to segment claims data on healthcare. Time-domain data are segmented into fixed-duration time segments whereas dynamic time warping (DTW) based detection and evaluation systems are employed in time domain without any obvious segmentation and feature extraction.

C. BIG DATA ANALYTICS APPROACH

Many data mining techniques are adopted to gain insights from large volumes of data. It is observed from existing research works that classification, clustering, regression analysis, deep learning and other techniques are used to analysis the data [29]. A number of Big Data analytics techniques are presented in Fig. 9 where the focus has given particularly on ML methods in this article.

1) DATA ANALYSIS USING ML

Machine learning is a highly popular technique particularly for data analytics that helps in making decision. The techniques for data analytics using ML is presented in Fig. 10. Feature extraction is one of the important tasks during signal analysis as useful information are extracted at the stage, which is considered as main feature for the corresponding data. Fourier transform and wavelet transform are used for processing physiological signals such as ECG. Fourier transform can preserve frequency domain information while wavelet transform retains both frequency and time domain information at a time [51]. The real time information based on collected data is visualized using computer, mobile phones and tablets to understand the data behavior. The system raises alarm to medical staff and user if any risky health situation is observed.

For example, feature extraction process reduces dimensionality of data to make it manageable groups.

ML techniques (i.e. classification, clustering, regression and association rules) are useful for extracting features for identification, classification and computation. Feature vectors extract feature such as signals from all three axes of gravity and body acceleration components [52]. Frequency-derived feature is one of the dominating approaches that employs parameters such as averages or correlations among them and estimated over long time [53]. Statistical methods detect activity of human body calculating variance, standard deviation, mean value, correlation etc. using vital signs [54]–[56]. Electrocardiogram (ECG), images and videos health signals data are converted into EHR (Electronic health record) with the use of ML.

Advanced ML techniques are used to learn and relate information from various sources and understand hidden relationship among source parameters [29]. They are massively used in building predictive models due to its excellent training speed, great memory capability and good accuracy in predicting health condition. Neural network (NN), Naïve Bayes, k nearest neighbor (k-NN), Support Vector Machine (SVM), Decision Tree (DT), regression, clustering, Hidden Markov model (HMM), Gaussian Mixture Modelling (GMM) etc. are currently used for detecting and predicting human health condition.

K-NN classifiers can measure geometrical distances between feature vectors from different classes. KNN is a simple, efficient non-parametric method used for classification, regression, object recognition etc. [57]. In KNN, input consists of the K, closest training examples, in the feature space. Classification of targeted object or an unknown sample is performed by assigning to a test pattern the class label of its K nearest neighbors. A voting procedure of nearest neighbors helps classifying object. During classification, most of the votes of its neighbors are considered, with the object assigned to the most common class among its K nearest neighbors. For example, if K = 1, that means one nearest neighbour is there where the object is simply assigned to that class.

SVM can classify data creating linear decision boundary to separate all data points in binary classes. In addition, Naive Bayes is a probabilistic technique performing data classification based on uppermost probability of its belonging to individual classes. Model and feature based ANN learn, recall and generalise from the given data through adjustment of weights that can perform checking and testing models for predicting final outcome [58]. Reinforcement learning, a promising advanced ML method, works as trial-and-error in analysing data and optimising sequential treatments particularly for chronic disease [29].

Extensive current literature is surveyed and information is extracted based on the type of sensor, sensor location, activities detection, ML techniques and corresponding prediction accuracies for health monitoring. A summary of current ML techniques used for human posture detection system is presented in Table 3.

TABLE 3. Examples of activity detection and corresponding Big Data analytics using ML techniques for health monitoring.

Type of sensor	Sensor location	Activities detection	ML techniques	Health status prediction accuracy	Ref.
Accelerometer	Thigh, arm, ankle and abdomen	Sitting, standing, walking	SVM, HMM, ANN	92%-99%	[59]
Force and acceleration sensors	Chair	Different sitting positions	SVM, Neural Network, Random Forest	81% and 98%	[60]
Pressure sensor and accelerometer	heel, heads of metatarsal bones,	Sitting, standing, walking, running	SVM	95.2%	[61]
Accelerometer	Chest, thigh, ankle	Stairs, walking, standing up, sitting on ground etc.	HMM, K-NN, Naïve-Bayes, SVM	80%-95%	[62]
Tri-axial accelerometer	Waist	standing, jumping, sitting-down, walking, running, and falling performed	Wavelet-based principle component analysis, SVM	92.25%- and 94.87%	[34]
Pressure sensor, accelerometer.	Sole and heel	Sitting, standing, jogging, cycling	SVM, MLP	96%	[53]
Accelero-meters	Waist, thighs	Sitting, lying, standing and walking speed	Extended Kalman Filtering	80-92%	[63]
Acc., gyro, magneto-meters	Waist, arms, legs	Sitting, standing, and walking behavior	HMM	92%	[64]
2D Accelerometer	4 limbs and the right hip	Walking, sitting, running, bicycling, stretching, lying down and relaxing	Naive Bayesian k-NN, binary decision, ANN	84%-92%	[65]
2D Accelero-meter	Connected a pair of pants with accelerometer to a laptop	Sitting, standing, descending stairs, walking etc.	ANN	96%	[66]
Wristwatch accelerometer	Wrist	Long-term activities (shopping, showering, sinner, working, etc.	Hidden Markov models, conditional random fields	75%	[67]
Wrist-wearable sensors	Wrist	arm and body movements	naive Bayes, decision trees, SVM and ANN	85.8%	[68]
Wearable physiological sensors	Fingers	electro-dermal activity (EDA) and the pulse plethysmograph (PPG) signals.	SVM	82%	[69]
Electrocardiogram (ECG), thermistor, Skin Response Sensor,	Chest and hand	Mental Stress Monitoring	SVM and K-means clustering	91.26%	[70]
GPS and accelerometer sensor	Attached with wooden board in backpack	To identify travel behaviors	Random forest	91.9%.	[71]
Accelerometer and barometer sensors	waist	posture transitions and falls	SVM and random forest	90%	[72]

Current literature shows that most of the big data analysis techniques for health monitoring are dominated by ANN, SVM, KNN, Naïve Bayesian and kalman filter [73]. Based on the reported prediction accuracy of various ML techniques used for processing health monitoring Big data (Table 3), it has been identified that ANN, SVM and KNN have shown accuracies in the range of 96% to 99.47%. Diverse data processing capabilities of the algorithms could be responsible for this. For example, ANN has the ability to learn and develop non-linear and complex relationships between inputs and outputs in real-world situations. ANNs are trained with a set of input and target values and composed of multiple nodes. The nodes can take input data and perform simple operations on the data as those are connected by links and can interact with other nodes. The result is passed on to other neurons where each link is associated with a weight. The learning process are taken place by altering these weight values. Due to having fault tolerant capability ANN performs minor changes in input without changing in output during training. A gradient descent is applied to perform model training by minimizing a cost function between estimated and desired network outputs.

On the other hand, SVM is a discriminative classifier that performs data mapping to a high-dimensional feature space to become linearly separable for categorising data points. The algorithm finds an optimal hyperplane that assists in separating new data points into classes even for unstructured and semi structured data like text, images etc. The real strength of SVM is kernel trick which is a set of mathematical functions taking data as input and transform it into the desired form. The kernel trick is used to transform data for finding an optimal boundary between possible outputs. The kernel trick avoids explicit mapping to obtain linear learning algorithms which can learn a nonlinear function or decision boundary. Certain functions can be defined as inner product in another space which is referred to as a kernel or a kernel function.

In addition, k-Nearest Neighbour (k-NN) is a nonparametric distance based supervised classifier having high efficiency and simplicity for implementation [74]. Although the algorithm needs high computation load for classification, but less time requirement to build model is one of the great advantages. With ML techniques such as SVM, KNN, NN etc high-speed data processing performance and long-term data storage are significant issues for mobile applications. For example, healthcare organisations should move active data to a high-performance platform for processing large physiological data collected through smart clothing. Due to increasing amount of data, parallel-access architecture is essential. Furthermore, hyper scale data centres with purpose-built server architectures and public cloud can also be useful to solve data storage issue. Organisations should use long-term archives like object stores or the public cloud to store the processed data in well-indexed platforms. Computation complexities can be decreased by reducing the dimension of the collected data if required.

Data labelling, used for detecting and tagging data samples, can be costly in some applications, especially in the healthcare. In unsupervised learning, a machine provides assumptions about data through learning without it being labelled [75]. Machine can able to find patterns by itself from the data to find desired outcome. It can create grouping data through clustering or association. However, subjective problem is observed in clustering. For example, clustering data related to people according to their gender or interests are considered to solve the problem of gender classification by unsupervised learning using similarities in the data. On the other hand, association provides solutions to find interesting and hidden relationships in the data. Multiple features in physiological data sometimes cause problems during data pre-processing such as requiring professionals during data annotation, time consuming, difficulty in acquiring labelled data etc. Semi-supervised learning can train learners without external interaction where small numbers of labelled and unlabelled samples are required during training. It helps improving classification accuracy and performance of the learners to enhance the applicability of the ML model. In case of limited labelled samples, semi-supervised algorithm performs better to reflect the overall structure of data samples [76]. Active learning is a smarter way of prioritising large amount of data which needs to be labelled in order to train a supervised model [77]. Optimisation of the data points chosen for labelling and training a model is performed. In active learning, manual labelling is performed through making small subsample of data for subsequent training. Then the model is used to predict the class of remaining unlabelled data points.

Hyperparameters can directly control the behaviour of a training algorithm and have a significant impact on the performance while a model is being trained [78]. A hyperparameter is not tuned during learning phase, it finds tuple of hyper-parameters that generates optimal model which minimizes objective function. Manual optimization is time-consuming where expert knowledge of four optimization strategies are commonly used including grid search, random search, hill climbing, and Bayesian optimization. A simple grid search can yield optimal hyperparameters by applying all possible combinations of hyper parameters. A typical optimization procedure elucidates a set of hyperparameters and metric to be maximized or minimized for a problem. An automatic optimization procedure tracks an iterative procedure where the model is trained on a new set of hyperparameters at each iteration and evaluated on the test set. Finally, the optimal set is identified based on hyperparameter set corresponding to the best metric score.

Generalization represents the ability of a model to adapt accurately to new data drawn from the same distribution as the one used to create the model [79]. It is critical to generalize beyond training data of any learning algorithm that helps making accurate predictions. The performance of generalization is measured based on the performance on out-of-sample data of the models. Training on one task provides instant improvement to performance on the new task. Consecutively training participants on tasks sharing a common
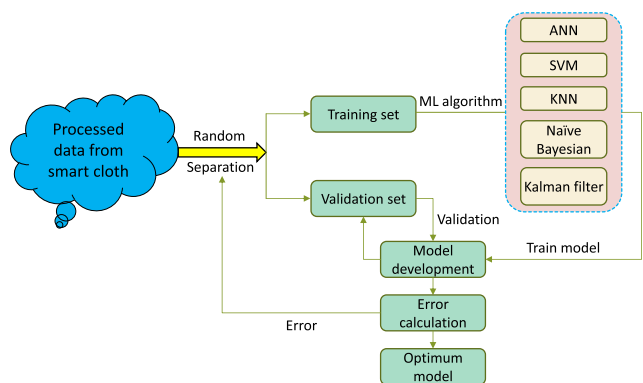
FIGURE 11. Data analytics framework for smart clothing.

high-level task structure can harvest faster learning of new tasks, where no immediate benefit can be obtained. Generalization is linked to the distribution of the training data. Test data distribution should follow training data distribution on new subject then the model performs well. For example, model needs to be changed or retrained in order to fit the new distribution of population if distribution is changed for a new subject. Sometimes it is difficult to obtain correct prediction once training data is not representative of the entire population. In practice, distribution of population changes over time where frequent re-training is required to keep the model up to date with the changes of population. Some systems are trained in one subject that may help model to do predictions that are more concrete, less variance and minimal bias ness. However, due to lack of massive data sets to train on one subject, model predictions should not be inclusive and of good quality.

A conceptual framework for the physiological Big Data analytics with the use of ML techniques has been developed and presented in Fig. 11. Conventional data analytics techniques are not fully flexible and efficient due to increasing data size and slow processing with noisy data. Big Data analytics can provide customized solutions for handling large volume of data [80]. Therefore, advanced data analytics techniques are prerequisite for making appropriate decision using smart clothing. Training and validation data sets are prepared from the processed raw physiological data. During training, the data are taken randomly to build ML models. ANN, SVM and KNN are identified as the efficient algorithms for Big Data analytics for processing data collected by smart clothing. Validation data sets are then applied in the developed models and compared with known outputs. If the error is high, then the models are rebuilt again using random selection of processed data until developing optimum models for achieving best outputs.

2) DATA ANALYSIS USING DEEP LEARNING
Over the last few decades, machine learning approaches have achieved reliable results across many applications with small or medium sized data sets. Currently, deep learning has been under dynamic development on Big data analytics as a more advanced technique. It performs processing large data

by using artificial neural networks to solve complex problems. Traditional machine learning algorithms learn from model functions and predict future action from historical data whereas deep learning contains more than one hidden layer to interpret data features and relations [81]. It is becoming popular due to its superior performance in learning feature representation from raw data. Currently, deep learning algorithms including a Deep Neural Network (DNN), Convolutional Neural Networks (CNNs), Recurrent neural networks (RNN), Restricted Boltzmann machines and Deep autoencoders, are being considered for the purpose of Big data analytics and decision making [81].

Deep learning algorithms are largely self-directed on data analysis, can create new features and offers end-to-end problem solution which can help providing appropriate data analytics and decision making through the physiological signals collected by smart clothing [82]. For example, CNNs performs better in extracting health related information compared to machine learning methods like SVM. Besides, RNN is used for analysing clinical temporal sequence data. The main advantages including learning features and classifiers automatically with high accuracy and performing automatic feature extraction without the need for human intervention. A deep architecture is developed to direct complex hypotheses through adding more hidden layers to a neural network as the hidden layers capture nonlinear relationships. For example, deep learning explores clinical data in the direction of precision medicine through combining with the large data and Graphical Processing Units (GPU). The main disadvantages of deep learning are highly time-consuming and requiring high performances hardware. Deep learning has diverse ability to provide solutions for complex analysis. For example, Deep neural networks contain various levels of complexity. Therefore, the networks usually have capability of making complex decision on various cases. However, it is costly to train and build the model in terms of time.

Turner and Hayes [83] have proposed a method with combining non-invasive wearable sensors and deep learning for classifying artificially induced gait alterations. The method can diagnosis gait abnormalities based on a symptom which can identify movement disorders. Long-term and short-term memory networks with deep learning architecture was applied on collected pressure data for 12 patients. The method has achieved 82.0% accuracy in classifying gait function of patients accurately that can help in making decision the patient status.

Gumaei *et. al.* [84] have proposed sensor-based framework for recognising human activity using a hybrid deep learning method. The method has been developed combining simple recurrent units (SRUs) and gated recurrent units (GRUs) of neural networks using multimodal body sensing data. The experimental results achieved with 0.99 precision for the datasets used which helps user to take decision for monitoring health condition. However, the proposed system was unable to analyse bigger and complex data sets and monitor human behaviour in real time.

Hong *et. al.* [85] have proposed a method combining hybrid recommendation algorithm and deep neural networks (DNN) that can provide disease prediction based on medical history while unclear description of symptoms observed. The proposed method has used real-world datasets that can predict the potential Phenotype through handling high and low order disease features. Although the experimental results in this approach has shown significant accuracy in predicting disease. However, further research is required to add additional methods to improve the prediction accuracy and criteria that can help making decision.

Alhussein and Muhammad *et. al.* [86] have proposed a method applying deep learning to efficiently analyse healthcare data. The authors have conducted case study for predicting type 2 diabetic patients to establish a relation between laboratory and medical assessment variables. Prediction performances of myocardial infarction were measured on patients admitted in hospital employing linear discriminant analysis (LDA), support vector machines (SVM) and recurrent neural network (RNN). RNN can be useful in predicting disease which can help in making decision accurately as it has provided best accuracy among them.

Sierra-Sosa *et al.* [87] have proposed a voice pathology detection system on mobile healthcare framework using a convolutional neural network (CNN) through a transfer learning technique. The voices are captured using smart phone and processed. Experimental results revealed that the accuracy of voice pathology detection was achieved by 97.5%. Based on the results cloud manager sends the decision and samples of voice signal to local carer or doctor. The doctor analyses the signal, check the decision and provides feedback to cloud for patients.

3) OTHER APPROACHES
Apart from ML, few other approaches including Knowledge-based approaches [88], Deep learning [89], Real-time analytics [90], Clinical reasoning [91], End-user Driven data analytics [92], Natural language processing [93] and Healthcare knowledge-bases [29] have shown to be effective in healthcare to explore complex medical knowledge through structuring, integrating and managing the physiological data in health monitoring purposes.

4) COMMERCIAL PLATFORMS FOR BIG DATA ANALYTICS
It is a challenge to select appropriate platform based on data characteristics for accommodating required insights to make accurate decision. Many commercial software platforms are currently used for handling large volume of health-related data [94]. For example, Hadoop is useful for analysing image, patient information in healthcare analytics. The software tool uses data as batch format. However, it is not fully capable of recording the real time ECG reading. Apache Spark [95] is used for analysing large-scale magnetic resonance imaging (MRI) data. The platform is capable to analysis both stream and batch data. MapReduce [96] processes large-scale medicine data whereas Apache Storm [97]

FIGURE 12. Big Data analytics software platforms (adapted from [94]).

performs health data streaming to make quick decision for controlling cardio logical injuries in an emergency situation. In addition, Apache Kafka [98] is a data processing software used for storing and retrieving bioinformatics data. NVidia CUDA [99] is a parallel computing platform analyses health related batch and stream medical image data. Additional list of software platforms used in health related visualisation and analytics, compute, storage and data warehouse are presented in Fig. 12.

D. DECISION MAKING
Smart decision making supports accurate diagnosis and effective decision that provides proper healthcare plan to avoid any sudden health risk of people [100]. Subsequently, activity monitoring, fitness measurements, sign of early health damage can be tracked effectively on real time. This section describes few current examples of decision-making systems on health care application.

1) EXISTING FRAMEWORK BASED DECISION MAKING
Currently, researches on framework based decision-making systems have been developed particularly for healthcare applications. For example, Mora *et al.* [101] has presented an IoT based distributed framework using wearable technology (i.e. smart watch) for monitoring human biomedical signals for making decision on reducing sudden death and possible fatal injuries. Real-time heart telemetry system was used to collect cardiac parameters. The proposed system can broadcast abnormal behaviours observed for athletes compared with normal cardiac activities (Fig. 13). The cardiologist can take preventive decisions based on the cardiac disorders within few seconds and provides service to the athlete by analysing cardiac parameters, number of minutes played, temperature, humidity and distance during playing football. All the heart rates are recorded for all players during treadmill exercise. Distributed computation approach and Body Area Network (BAN) were employed for making decision. The system makes decision and displays green signal on screen for normal condition of heart through monitoring ECG signal. For any unusual ECG signal, red signal is displayed that implies emergency cardiac situation and medication are to be arranged immediately.
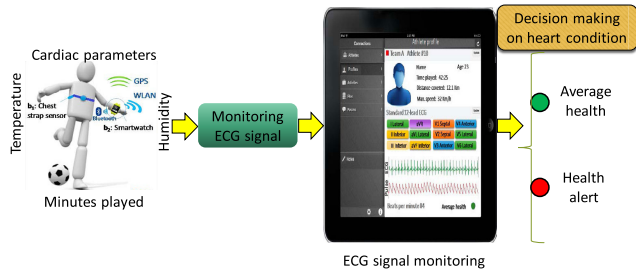
FIGURE 13. Real time health telemetry system for football players (information taken from [79]).

TABLE 4. Decision making score for early warning (information taken from [23]).

Physiological parameters	Score 0	Score 1	Score 2	Score 3	Decision making on controlling heart rate
Heart rate (beats/sec)	51-90				Normal
		91-110			Keep attention
			111-130		Under monitoring
				≥131	Critical condition

Another investigation in [102] presents a generalised multi-sensor data fusion approach for assessing and making decision (Health-RAD) on health risk using a Wireless Body Sensor Network (WBSN). The approach has shown health severity level using vital signs (respiration rate, oxygen saturation, temperature and scores) where a risk variable contains the values between 0 and 1. For example, a score of 0 indicates that respiration rate is between 12 bpm and 20 bpm. The decision for critical condition of health is made for the value of high-risk variable (outside the range 0-3) and needed attention for that scenario. A heart rate scoring system on making decision with the use of fusion approach is presented in Table 4. The score of vital signs are calculated with the use of past and current value signs. Then the score is used for assessing health status based on its progress during a period of time with the employment of fuzzy inference system (FIS) and early warning score systems (EWS). The proposed approach was used to improve energy consumption (86% less than other approaches), risk assessment of vital signs and determination of the individual's health risk level.

A green cognitive body sensor network (BSN) is presented in [103] for monitoring and making decision of user's emotion status by analysing ECG signals with the combination of biosensors, smart clothing and artificial intelligence. The experiments for emotions data analysis were conducted using electrocardiograph (ECG) and photo plethysmography (PPG) signals. Deep learning algorithm was employed to train and build model with different emotional states based on user's ECG. The prediction for new ECG data set shows that
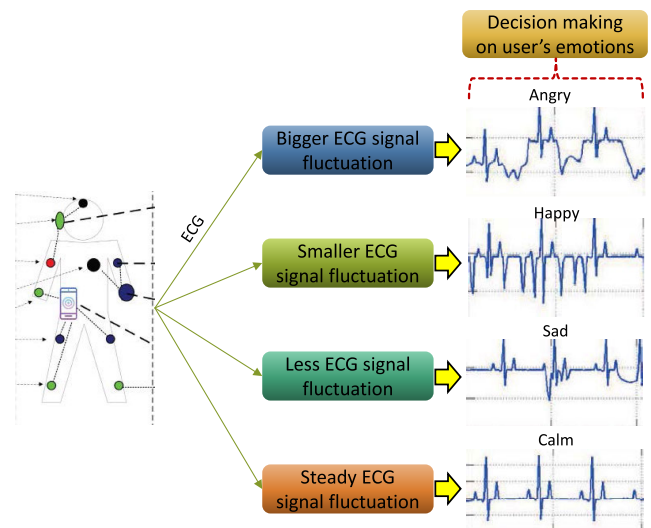
FIGURE 14. A conceptual framework for decision making in patient's emotions using body sensor network through smart clothing (information taken from [103]).

TABLE 5. Physical activities and health decisions are made with health score (information taken from [104]).

Hours	Heart rate	Score	Activity	Decision making on heartbeat control
0-10	70 beats/Min	0	Sleeping	Normal
10-11	90 beats/Min	1	Light sleeping	Normal
11-12	110 beats/Min	2	Sedentary	Keep monitoring

ECG signal waveform having bigger signal fluctuation are observed during user's angry mood whereas a happy mode is observed through smaller fluctuation (Fig. 14). In addition, small and steady signals indicated sad and calm mood.

Azimi *et al.* [104] have proposed an IoT based health decision-making approach with the use of heart rate data of 20 pregnant women for 7 months. A weighted arithmetic mean was exploited, and rule-based indicator was used for continuous monitoring of heart rate with changing physical activity to make decision on health status. The ranges of decision were made based on scores measured from heartbeat between the score 0 to 3 (Table 5).

Thaung *et al.* [105] have proposed a decision support system on health monitoring based on different activities including sitting, standing, walking, climbing and running using multibody sensor for hospital patients. The system was consisted of heartbeat sensor, temperature sensor and pulse oxygen saturation level (SPO2) sensor. The result obtained from their research work has been presented in the Table 6. It was observed that temperature and oxygen saturation level were stayed in same normal ranges for all five cases and which are 96°F to 99°F and 95% to 100% respectively. The heart rate has increased above 125 bpm during climbing and

TABLE 6. Health decisions are made considering temperature, heart rate and oxygen saturation level (information taken from [104]).

Position	Temperature sensor,	heart rate sensor,	oxygen saturation level sensor	Decision making
Sitting	95°F - 98°F	95 bpm - 88bpm	95% - 100%.	Normal
Standing	Same	Same	Same	Normal
Walking	Same	Same	Same	Normal
Climbing	Same	>125 bpm	Same	Emergency
Running	Same	>125 bpm	Same	Emergency

running. Then the proposed system can detect as emergency situation for the two activities.

Zarkogianni *et al.* [106] have presented biosensors-based glucose and lifestyle monitoring system and developed clinical decision support system (CDSS) to help self-management of underlying health conditions and to support healthcare professionals in making quick decisions. As a case study to decision support in caring diabetes, the proposed approach was verified by integrated sensor data and intelligent data analytics methods.

2) PROPOSED AI BASED DECISION MAKING FRAMEWORK
Intelligent decision-making is still at early stage in the context of integrating multisource Big Data including electronic medical records, medical images, health archives etc. Majority of the research work have proposed different Big Data analytics for health monitoring using smart clothing though the outputs are not quite linked to making effective decision thorough easy to use user interface. Therefore, integration of machine learning and user interface development for data analytics and decision making is required to take the full advantages of large volume of data collected by the smart clothing compared to the other traditional wearable technologies. A conceptual framework is presented in Fig. 15 to enhance better decision making through monitoring health using smart clothing particularly in the area of healthcare and fashion. ML techniques are used to train and build predictive models using the sensed data collected from human body with smart clothing over a longer period. Decision points for primary and emergency care are identified based on a set of threshold values (normal and emergency). A user interface is developed that will continuously update and display the measured health status of individuals. For non-emergency condition, primary care instructions are provided remotely through smart phones, tablets etc. On the other hand, for critical condition, emergency services will bring the monitored person immediately for hospitalisation and treatment plan.

Thus, appropriate action is taken avoiding any severe health damage. Sensing based decision making by smart clothing will significantly improve the quality of life and

minimise service cost. Early decision making on health condition, huge amount of treatment cost or most importantly a patient's life is saved through applying the proposed framework.

The competitive environment in sports and athletics causes physical and mental stresses that can negatively affect the performance of the sportsmen [107]. Additionally, safeguarding them from severe injuries during training and actual competition is an important issue. Therefore, to keep them fit and to improve their performance, it is essential to provide feedback based on real-time continuous monitoring of their physiological parameters such as body temperature, blood pressure, heart rate etc. and correlate with their performances.

Similar to health care and fashion applications, a conceptual framework has been developed (Fig. 16) to enhance better decision-making using ML for improving the performance of athletes and sportsmen through continuous monitoring of their physiological parameters using smart cloth, storing in cloud server and measuring relevant performance. The framework provides a collaborative platform among coaches, athletes and sportsman through knowledge sharing. ML techniques are employed to develop predictive models with the measured physiological data to help in decision making for performance improvement via a user interface. The monitored parameters are constantly updated and compared with thresholds performance by ML and feedback messages are provided for further improvement of the performance [108]. Minor performance improvement suggestion such as weight checking can be made if 100 m running time of an athlete during training session is slightly increased. In the case of significant performance drop, feedback message with full medical check-up can be suggested. The coach can set appropriate training plan for performance improvement. In addition, the coach can also provide suggestions to boost mental performance through concentration, confidence building, emotion control and commitment.

III. CHALLENGES AND FUTURE DEVELOPMENT
In order to take full advantage of the profound patterns contained in the massive data for making quick decision, data storage, mining, analysis, and privacy are essential [29], [36]. Current research presents that Big Data analytics for smart clothing is not fully matured yet. Although Big Data analytics in the area of smart clothing has potentials, still many challenges need to be resolved. A summary of current challenges of Big Data analytics is presented in Fig. 17 and the following subsections present further details of the challenges along with future development opportunities.

A. MINING OF BIG DATA
Majority of real-life data comes in a variety of unstructured form collected from sensor readings such as ECG measurements in intensive care, text data, imaging data, or omics-data, which cause difficulties in integration and analysis. Specialised algorithm designed for a particular application can organise the data as structured form. Low
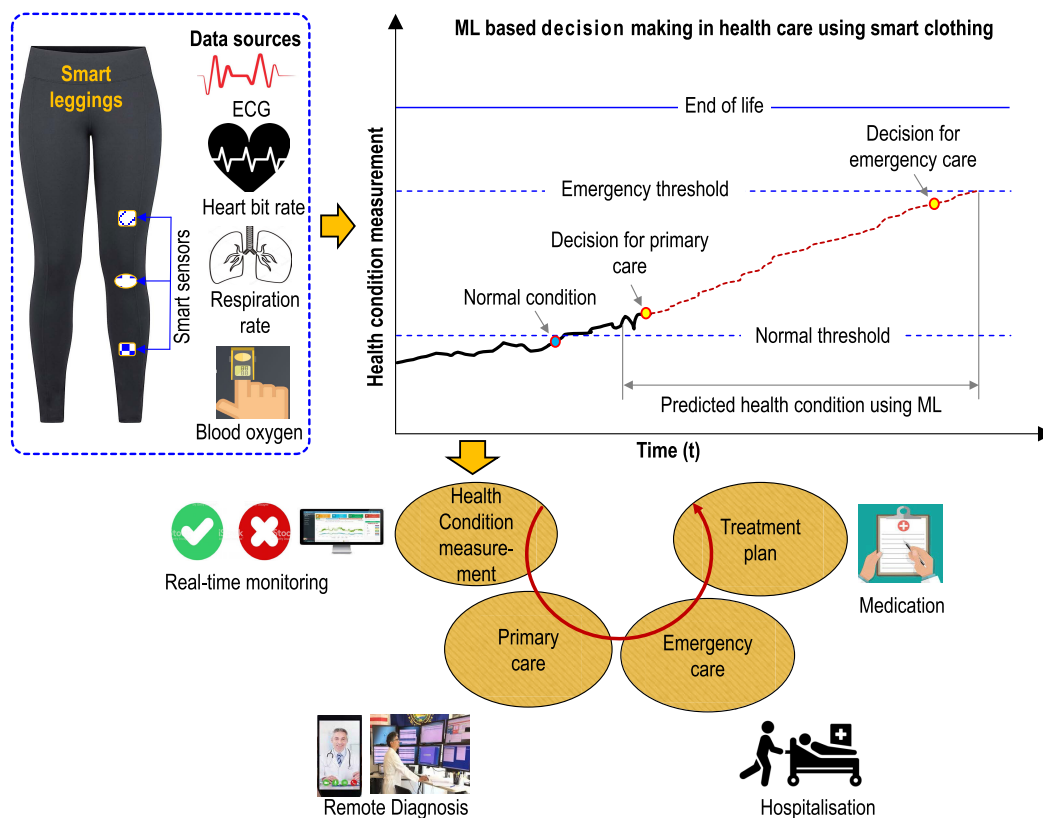
FIGURE 15. A conceptual framework for decision making in healthcare and fashion using smart clothing.

processing speed of Big Data is another important challenge whereas high bandwidth networking system can be employed to import and export large amount of data to the cloud [111]. Besides, large amount of unstructured data causing inaccuracy in mining the Big Data could be a common issue responsible for longer processing time. The development of advanced data mining tools with the opportunity of ML and statistical analysis can provide efficient data mining.

B. DATA STORING

High cost associated with storing of massive data size particularly medical data such as diagnostics images and pathological analysis could pose a huge challenge. Medical data are usually derived from millions of people with their various disease information. It is important that patients' data are to be retained for more than 50 years. Due to keeping the huge quantity of data, it is difficult to manipulate the data including storing, extracting, and downloading [23]. Secure cloud services can be improved by increasing data storing capacity and downloading speed to minimize the data storing cost.

C. DATA SHARING

Health care data are recently gathered in terabyte (TB) level and even sometimes, it reaches petabyte (PB) level, which

is beyond the capabilities of current computing devices and network file sharing programs. Therefore, development of new sharing mechanism will help to accommodate all data in real time [112].

Modern technologies are insufficient to meet the requirements for the integrative health care Big Data applications. Unique standards, constant description style, and presentation approach are important challenges in Big Data analytics. Diverse levels of structured, semi-structured, and unstructured data integration are difficult and different software are used to manage each of them. Thus, data comparison, analysis, transfer and sharing are challenging due to data incompatibility. Integrating various types of data may reduce the cost of managing unstructured data separately [6].

Big Data from healthcare covers a wide range including clinics, regional medical centres, sports and athletics, thus corresponding data resources are distributed in different data pools such as patient's records, settlement and cost data etc. Therefore, lack of connection among the data sets can be observed. In addition, data-sharing mechanism is inadequate due to the information barriers among hospitals and other research institutes [112]. Neither humans nor algorithms can always guarantee to deliver an optimal solution with noisy data to take important decisions within a short time. Smart decision support with the use of medical data, pathology,
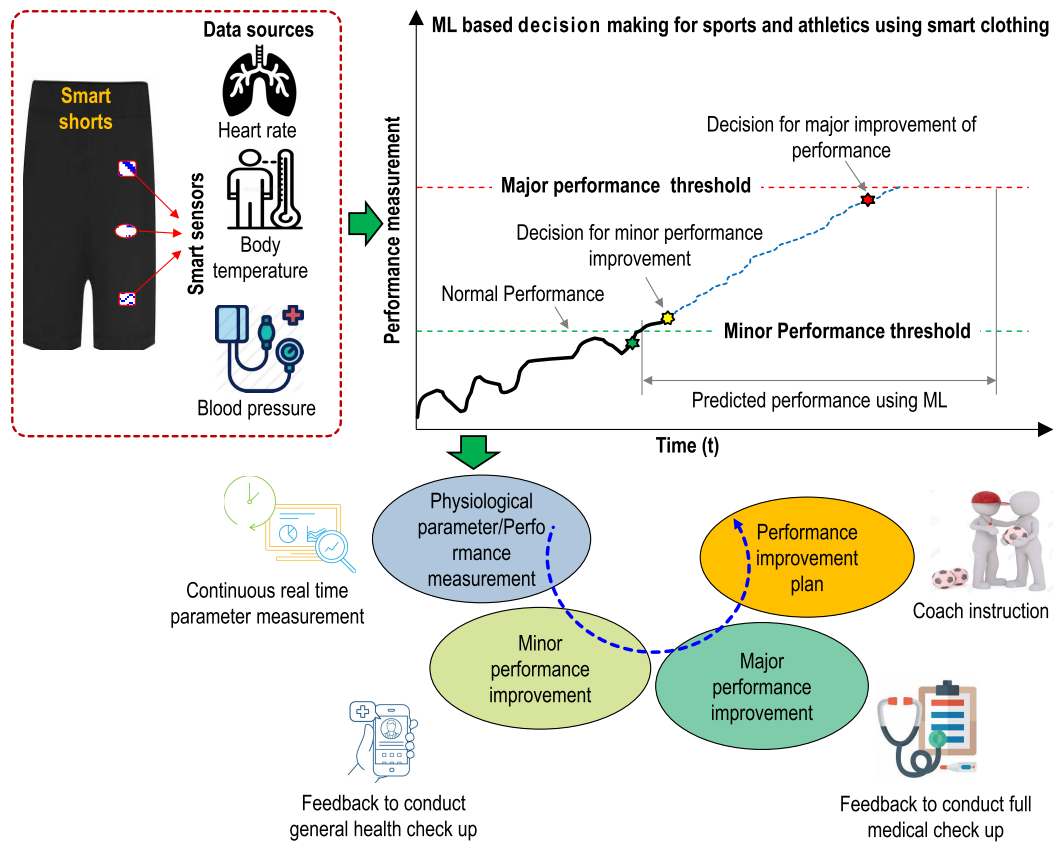
FIGURE 16. A conceptual framework for decision making in sports and athletics using smart clothing.

FIGURE 17. Challenges of Big Data analytics using smart clothing (information taken from [109], [110]).

intensive care monitoring for healthcare can solve health related issues and lead healthier lives with the use of smart-phones, smart clothing and sensor technologies [36].

Current data integration methods with software system have not yet been fully embedded. For example, health-monitoring data are yet to be integrated into clinical

diagnosis, treatment, and same for clinical data, which has not been integrated into public health services [19]. End-users in medicine such as doctors, researchers and bio-informaticians are not highly trained to exploit full potential of the software tools. Hence, an optimal and user-friendly analytical approach not requiring any specialist skills can be developed in order to cross check trustworthiness of the results. Self-service analytics can be developed in future allowing conducting the analytics process automatically [36].

D. DATA PRIVACY
Unnecessary personal information of the patients is collected along with actual Bio signals. In many cases where Big Data analytics could cause leakage of personal information and important data could be lost during copying and preserving. De-identification and digital identity encryption can be employed with the technology to avoid such risk [36]. Although large database uses anonymous personal encrypted data, personal information can be re-determined by re-identifying for the use of pseudonymized personal confidential data. In addition, de-anonymization is risky where anonymous data and other sources of data are compared for re-identifying the anonymous data sources [113]. Information hacking is one of the leading privacy breaches. Security risks are increasing due to lack of understanding in using

technology by the health care community [114]. Big data in healthcare raises huge concerns related to security and Privacy issues of patients. There is still risk of keeping patient records safely although their information is stored through applying different levels of security in the data centres. It is noticeable that lack of security can affect patient's privacy due to the increasing use of mobile devices. Researches are being undertaken to protect the data breaches during transmission, storage and usages. For example,

a. Wiping personal details after ending patient's session from a device
b. Applying two-steps authentication to protect the stored data
c. Setting Secure Sockets Layer (SSL) or Transport Layer Security (TLS) between user's app and other systems
d. Establishing End-to-End data encryption and decryption during transmission
e. Not to encourage interference with third party application including commercial advertisements.

Additionally, Big data analytics can cause various ethical issues. A failure to handle data ethically can lead loss of trust towards organisations. During Big data collection, it is required to audit based on legal requirements. The data should be obtained taking consent from the persons that should not be exposed for use by any organisations containing any hints to their identity. Third party persons or organisations should share the data with restrictions and high confidentially. Customers should be well aware about the transparency of their personal information (i.e. date of birth, sex etc.) in case of how the data are being managed through third party analytical system.

A sophisticated framework can enable the analysis of such data by incorporating adequate privacy-preserving analytical tools considering ethical issues. Strict security procedures such as employing specialized anti-virus software, data encryption, and multi-factor authentication along with smart decision-making framework could help in preventing data breaching.

E. DATA VISUALISATION
Uncleaned data interrupts visualizing accurate interpretation and assessment of health condition. Most cases health related data cannot be involved for the incapability of traditional techniques. Modern visualised techniques are yet to apply achieving actual insight of the analysed data for making accurate health decision. Color-coding visualization technique has diverse characteristics that can be employed for developing immediate response during data analytics. For examples, red, yellow, and green colour help understanding universally to mean stop, caution, and go respectively. In addition, poor representation of the analysed data such as complex flowcharts, overlapping text, and low-quality graphics cause frustration to recipients [115]. Advanced statistical tools such as heat maps, bar and pie charts, scatterplots, and histograms may help illustrating concepts and information accurately.

F. DATA UPDATING SYSTEM
Healthcare data is dynamic and traditional techniques are not capable to update information in real time. The vital signs from human body are required to update every few seconds while monitoring their health. Understanding the volatility of data in terms of what degree and how often the data changes are the important challenges to consistently monitor the targeted vital signs. Manual updating of health information is another drawback as all information of monitored parameters may not be included in the system and thus decision-making cannot be achieved accurately [115]. Health service providers are required to identify datasets for manual and automatic updating. To overcome the problem, the process can be conducted in a way without having downtime for end-users and damaging the quality or integrity of the dataset. In addition, unnecessary duplicate records can be eliminated while attempting an update of the data as it creates complications to make decision accurately for health providers.

IV. CONCLUSION
Existing research provided explicitly with details framework on making decision for monitoring health condition using wearable technology. Smart clothing-based health monitoring existing frameworks are incomplete with missing interpretations. This article presents a comprehensive framework with the sources of health indicator data, data collection, analytics and decision making for monitoring health condition using smart clothing in healthcare, sports and athletics, and fashion applications. It has been found that a number of bio-signals such as EMG, EEG, posture movements, blood glucose signals etc. are used as the source of Big Data to be collected continuously from the human body with the use of smart clothing. Lifestyle data such as nutrition, physical activity, sleeping also create great opportunities as the source of Big Data. Wi-Fi or Bluetooth is the appropriate medium identified for collecting bio signals from human body. Big Data analytics approaches such as ML, is considered as an appropriate approach helps in making effective decision for the proposed framework.

The significance of this article lies in proposing a new approach of Big Data analytics for data obtained by smart clothing. The approach suggests a decision-making system to enhance the quality of human life. The combination of Big Data, cloud-based services along with the application of ML techniques open up the opportunity for radically improving healthcare system and reduce costs by extracting meaningful information from Big Data.

Further research is required to address the challenges of Big Data analytics such as ineffective data mining, lack of data storage, limited data sharing capability and lack of privacy. A prototyping set-up can be developed in a laboratory environment to test the proposed framework. A software interface can also be developed to display the health status of users to get a real-time human health monitoring system. Deep learning for example DNN, CNN can be included for

analysing physiological Big and complex data. Verification for stability of the algorithm can be performed by comparing additional classification methods in future. Furthermore, reliability of the sensors can also be taken into consideration in future for collecting accurate biosignals.

In addition, identification of new bio-signals through advanced sensing, speedy data communication, real time data analytics and visualization are the major requirements to take full advantage of the knowledge hidden in Big Data.

REFERENCES

[1] C. Gonçalves, A. Ferreira da Silva, J. Gomes, and R. Simoes, "Wearable E-textile technologies: A review on sensors, actuators and control elements," *Inventions*, vol. 3, no. 1, p. 14, Mar. 2018, doi: 10.3390/inventions3010014.

[2] T. Fernández-Caramés and P. Fraga-Lamas, "Towards the Internet-of-smart-clothing: A review on IoT wearables and garments for creating intelligent connected E-textiles," *Electronics*, vol. 7, no. 12, p. 405, Dec. 2018.

[3] A. Asante-Korang and J. P. Jacobs, "Big data and paediatric cardiovascular disease in the era of transparency in healthcare," *Cardiol. Young*, vol. 26, no. 8, pp. 1597–1602, Dec. 2016.

[4] S. S.-L. Tan, G. Gao, and S. Koch, "Big data and analytics in healthcare," *Methods Inf. Med.*, vol. 54, no. 6, pp. 546–547, 2015.

[5] M. Herland, T. M. Khoshgoftaar, and R. Wald, "A review of data mining using big data in health informatics," *J. Big Data*, vol. 1, no. 1, p. 2, 2014.

[6] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: Promise and potential," *Health Inf. Sci. Syst.*, vol. 2, no. 1, pp. 1–10, Dec. 2014.

[7] S. Kumar and M. Singh, "Big data analytics for healthcare industry: Impact, applications, and tools," *Big Data Mining Anal.*, vol. 2, no. 1, pp. 48–57, Mar. 2019.

[8] C. Rodriguez, A. Barrow, S. Dangore, U. Pathak, and J. Talledo, "Applying data analytics to big data obtained from wearable devices," Student-Fac. Res. Day, CSIS, Pace Univ., New York, NY, USA, Tech. Rep. B6, 2018, pp. 1–6. [Online]. Available: http://csis.pace.edu/~ctappert/srd2018/2018PDF/b6.pdf

[9] Y. Liu, H. Wang, W. Zhao, M. Zhang, H. Qin, and Y. Xie, "Flexible, stretchable sensors for wearable health monitoring: Sensing mechanisms, materials, fabrication strategies and features," *Sensors*, vol. 18, no. 2, pp. 1–35, 2018.

[10] S. Sakr and A. Elgammal, "Towards a comprehensive data analytics framework for smart healthcare services," *Big Data Res.*, vol. 4, pp. 44–58, Jun. 2016.

[11] P.-Y. Wu, C.-W. Cheng, C. D. Kaddi, J. Venugopalan, R. Hoffman, and M. D. Wang, "Omic and electronic health record big data analytics for precision medicine," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 263–273, Feb. 2017.

[12] D. V. Dimitrov, "Medical Internet of Things and big data in healthcare," *Healthcare Informat. Res.*, vol. 22, no. 3, p. 156–163, 2016.

[13] H. Ma and X. Pang, "Research and analysis of sport medical data processing algorithms based on deep learning and Internet of Things," *IEEE Access*, vol. 7, pp. 118839–118849, 2019.

[14] H.-C. Shih, "A survey of content-aware video analysis for sports," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 5, pp. 1212–1231, May 2018.

[15] S. Jain, J. Bruniaux, X. Zeng, and P. Bruniaux, "Big data in fashion industry," in *Proc. IOP Conf Ser. Mater. Sci. Eng.*, vol. 254, 2017, pp. 1–6.

[16] E. S. Silva, H. Hassani, and D. Ø. Madsen, "Big data in fashion: Transforming the retail sector," *J. Bus. Strategy*, vol. 41, no. 4, pp. 21–27, Jul. 2019, doi: 10.1108/JBS-04-2019-0062.

[17] H. M. Krumholz, "Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system," *Health Affairs*, vol. 33, no. 7, pp. 1163–1170, Jul. 2014.

[18] J. E. Anderson and D. C. Chang, "Using electronic health records for surgical quality improvement in the era of Big Data," *Jama Surg.*, vol. 150, no. 1, pp. 24–29, 2015.

[19] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 171–209, Apr. 2014.

[20] N. V. Chawla and D. A. Davis, "Bringing big data to personalized healthcare: A patient-centered framework," *J. Gen. Internal Med.*, vol. 28, no. S3, pp. 660–665, Sep. 2013.

[21] T. W. Kim, K. H. Park, S. H. Yi, and H. C. Kim, "A big data framework for u-Healthcare systems utilizing vital signs," in *Proc. Int. Symp. Comput., Consum. Control*, Taichung, Taiwan, Jun. 2014, pp. 494–497.

[22] R. Fang, S. Pouyanfar, Y. Yang, S.-C. Chen, and S. S. Iyengar, "Computational health informatics in the big data age: A survey," *ACM Comput. Surv.*, vol. 49, no. 1, pp. 1–36, Jul. 2016.

[23] A. E. Youssef, "A framework for secure healthcare systems based on big data analytics in mobile cloud computing environments," *Int. J. Ambient Syst. Appl.*, vol. 2, no. 2, pp. 1–11, Jun. 2014.

[24] W. Lin, W. Dou, Z. Zhou, and C. Liu, "A cloud-based framework for Homediagnosis service over big medical data," *J. Syst. Softw.*, vol. 102, pp. 192–206, Apr. 2015.

[25] M. D. C. Legaz-García, C. Martínez-Costa, M. Menárguez-Tortosa, and J. T. Fernández-Breis, "A semantic Web based framework for the interoperability and exploitation of clinical models and EHR data," *Knowl.-Based Syst.*, vol. 105, pp. 175–189, Aug. 2016.

[26] B. K. Sarkar, "Big data for secure healthcare system: A conceptual design," *Complex Intell. Syst.*, vol. 3, no. 2, pp. 133–151, Jun. 2017.

[27] M. I. Pramanik, R. Y. K. Lau, H. Demirkan, and M. A. K. Azad, "Smart health: Big data enabled health paradigm within smart cities," *Expert Syst. Appl.*, vol. 87, pp. 370–383, Nov. 2017.

[28] F. Rahman, M. Z. A. Bhuiyan, and S. I. Ahamed, "A privacy preserving framework for RFID based healthcare systems," *Future Gener. Comput. Syst.*, vol. 72, pp. 339–352, Jul. 2017.

[29] *Big Data Technologies in Health Care-Needs, Opportunities and Challenges*, TF7 Healthcare Subgroup, Big Data Value Assoc. (BDVA), Brussels, Belgium, 2016, pp. 1–34.

[30] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: Management, analysis and future prospects," *J. Big Data*, vol. 6, no. 1, pp. 1–25, Dec. 2019.

[31] M. Chan, D. Estève, J.-Y. Fourniols, C. Escriba, and E. Campo, "Smart wearable systems: Current status and future challenges," *Artif. Intell. Med.*, vol. 56, no. 3, pp. 137–156, Nov. 2012.

[32] H. Tahir, R. Tahir, and K. McDonald-Maier, "On the security of consumer wearable devices in the Internet of Things," *PLoS ONE*, vol. 13, no. 4, Apr. 2018, Art. no. e0195487, doi: 10.1371/journal.pone.0195487.

[33] M. Chen, Y. Ma, J. Song, C.-F. Lai, and B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," *Mobile Netw. Appl.*, vol. 21, no. 5, pp. 825–845, 2016.

[34] Y. Wu, R. Chen, J. Wang, X. Sun, and M. F. H. She, "Intelligent clothing for automated recognition of human physical activities in free-living environment," *J. Textile Inst.*, vol. 103, no. 8, pp. 806–816, Aug. 2012.

[35] C. Ma, W. Li, R. Gravina, and G. Fortino, "Posture detection based on smart cushion for wheelchair users," *Sensors*, vol. 17, no. 4, pp. 1–18, 2017.

[36] H. Liang, L. Mengqi, W. Ruixue, L. Peixin, L. Wei, and L. Long, "Big data in health care: Applications and challenges," *Data Inf. Manage.*, vol. 2, no. 3, pp. 175–197, 2019.

[37] C. de la Fuente, A. Urrutia, and E. Chávez, "Using the random forest algorithm for searching behavior patterns in electronic health records," *IEEE Latin Amer. Trans.*, vol. 17, no. 5, pp. 875–881, May 2019.

[38] J. S. Rumsfeld, K. E. Joynt, and T. M. Maddox, "Big data analytics to improve cardiovascular care: Promise and challenges," *Nature Rev. Cardiol.*, vol. 13, no. 6, pp. 350–359, Jun. 2016.

[39] V. Kovalev and A. Kalinovsky, "Big medical data: Image mining, retrieval and analytics," in *Proc. Big Data Predictive Anal.*, Minsk, Belarus, Jun. 2015, pp. 33–46.

[40] M. Hassanalieragh, A. Page, T. Soyata, G. Sharma, M. Aktas, G. Mateos, B. Kantarci, and S. Andreescu, "Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: Opportunities and challenges," in *Proc. IEEE Int. Conf. Services Comput.*, Jun. 2015, pp. 285–292.

[41] W. Zhao, C. Wang, and Y. Nakahira, "Medical application on Internet of Things," in *Proc. IET Int. Conf. Commun. Technol. Appl. (ICCTA)*, Oct. 2011, pp. 660–665.

[42] F. Hu, D. Xie, and S. Shen, "On the application of the Internet of Things in the field of medical and health care," in *Proc. IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things IEEE Cyber, Phys. Social Comput. (GreenCom), (iThings/CPSCom)*, Aug. 2013, pp. 2053–2058.

[43] T. M. Fernández-Caramés and P. Fraga-Lamas, "A review on the use of blockchain for the Internet of Things," *IEEE Access*, vol. 6, pp. 32979–33001, 2018.

[44] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon, and W. Heinzelman, "Cloud-vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Cappadocia, Turkey, Jul. 2012, pp. 59–66.

[45] S. Stoyanov, M. Ahsan, C. Bailey, T. Wotherspoon, and C. Hunt, "Predictive analytics methodology for smart qualification testing of electronic components," *J. Intell. Manuf.*, vol. 30, no. 3, pp. 1497–1514, Mar. 2019.

[46] M. Ahsan, S. Stoyanov, C. Bailey, and A. Albarbar, "Developing computational intelligence for smart qualification testing of electronic products," *IEEE Access*, vol. 8, pp. 16922–16933, 2020.

[47] S. Luo and P. Johnston, "A review of electrocardiogram filtering," *J. Electrocardiol.*, vol. 43, no. 6, pp. 486–496, Nov./Dec. 2010.

[48] J. Wang, R. Zhu, and S. Liu, "A differentially private unscented Kalman filter for streaming data in IoT," *IEEE Access*, vol. 6, pp. 6487–6495, 2018, doi: 10.1109/ACCESS.2018.2797159.

[49] Q. Zheng, J. Chen, P. Huang, and R. Hu, "Urban scene semantic segmentation with insufficient labeled data," *China Commun.*, vol. 16, no. 11, pp. 212–221, Nov. 2019.

[50] R. M. Wood, B. J. Murch, and R. C. Betteridge, "A comparison of population segmentation methods," *Oper. Res. Health Care*, vol. 22, pp. 1–12, Sep. 2019.

[51] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, "A wavelet optimization approach for ECG signal classification," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 342–349, Jul. 2012.

[52] I. Mandal, S. L. Happy, D. P. Behera, and A. Routray, "A framework for human activity recognition based on accelerometer data," in *Proc. 5th Int. Conf.-Confluence Next Gener. Inf. Technol. Summit (Confluence)*, Sep. 2014, pp. 600–603.

[53] W. Tang and E. S. Sazonov, "Highly accurate recognition of human postures and activities through classification with rejection," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 1, pp. 309–315, Jan. 2014.

[54] J. Dai, X. Bai, Z. Yang, Z. Shen, and D. Xuan, "PerFallD: A pervasive fall detection system using mobile phones," in *Proc. 8th IEEE Int. Conf. Pervas. Comput. Commun. Workshops*, Mar. 2010, pp. 292–297.

[55] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 871–879, Mar. 2009.

[56] L. Cheng, Y. Guan, K. Zhu, and Y. Li, "Recognition of human activities using machine learning methods with wearable sensors," in *Proc. IEEE 7th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Las Vegas, NV, USA, Jan. 2017, pp. 1–7.

[57] B. Nukala, T. Nakano, A. Rodriguez, J. Tsay, J. Lopez, T. Nguyen, S. Zupancic, and D. Lie, "Real-time classification of patients with balance disorders vs. normal subjects using a low-cost small wireless wearable gait sensor," *Biosensors*, vol. 6, no. 4, p. 58, Nov. 2016.

[58] P. Randhawa, V. Shanthagiri, and A. Kumar, "A review on applied machine learning in wearable technology and its applications," in *Proc. Int. Conf. Intell. Sustain. Syst. (ICISS)*, Palladam, India, Dec. 2017, pp. 347–354.

[59] P.-H. Chiang and S. Dey, "Offline and online learning techniques for personalized blood pressure prediction and health behavior recommendations," *IEEE Access*, vol. 7, pp. 130854–130864, 2019.

[60] R. Zemp, M. Tanadini, S. Plüss, K. Schnüriger, N. B. Singh, W. R. Taylor, and S. Lorenzetti, "Application of machine learning approaches for classifying sitting posture based on force and acceleration sensors," *Biomed. Res. Int.*, vol. 2016, pp. 1–9, Oct. 2016.

[61] E. S. Sazonov, G. Fulk, J. Hill, Y. Schutz, and R. Browning, "Monitoring of posture allocations and activities by a shoe-based wearable sensor," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 4, pp. 983–990, Apr. 2011.

[62] F. Chamroukhi, S. Mohammed, D. Trabelsi, L. Oukhellou, and Y. Amirat, "Joint segmentation of multivariate time series with hidden process regression for human activity recognition," *Neurocomputing*, vol. 120, pp. 633–644, Nov. 2013.

[63] L. Atallah, B. Lo, R. King, and G.-Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," *IEEE Trans. Biomed. Circuits Syst.*, vol. 5, no. 4, pp. 320–329, Aug. 2011.

[64] A. Mannini and A. M. Sabatini, "Machine learning methods for classifying human physical activity from on-body accelerometers," *Sensors*, vol. 10, no. 2, pp. 1154–1175, Feb. 2010.

[65] L. Bao and S. S. Intille, "Intille activity recognition from user-annotated acceleration data," in *Proc. Int. Conf. Pervasive Comput.*, Viena, Austria, Apr. 2004, pp. 1–17.

[66] V. K. Laerhoven and O. Cakmakci, "What shall we teach our pants? Digest of papers," in *Proc. 4th Int. Symp. Wearable Comput.*, Atlanta, GA, USA, Oct. 2000, pp. 77–83.

[67] E. R. Garcia-Ceja, R. F. Brena, J. C. Carrasco-Jimenez, and L. Garrido, "Long-term activity recognition from wristwatch accelerometer data," *Sensors*, vol. 14, no. 12, pp. 22500–22524, 2014.

[68] Y. Lu, S. Zhang, Z. Zhang, W. Xiao, and S. Yu, "A framework for learning analytics using commodity wearable devices," *Sensors*, vol. 17, no. 6, p. 1382, Jun. 2017, doi: 10.3390/s17061382.

[69] V. Sandulescu, S. Andrews, D. Ellis, N. Bellotto, and O. M. Mozos, "Stress detection using wearable physiological sensors," in *Artificial Computation in Biology and Medicine*. Cham, Switzerland: Springer, 2015, pp. 526–532.

[70] T. Salafi and J. C. Y. Kah, "design of unobtrusive wearable mental stress monitoring device using physiological sensor," in *Proc. 7th WACBE World Congr. Bioeng.* Cham, Switzerland: Springer, 2015, pp. 11–14.

[71] K. Ellis, S. Godbole, S. Marshall, G. Lanckriet, J. Staudenmayer, and J. Kerr, "Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms," *Frontiers Public Health*, vol. 2, p. 36, Apr. 2014.

[72] D. Rodríguez-Martín, A. Samà, C. Pérez-López, A. Català, and J. Cabestany, "Posture transition analysis with barometers: Contribution to accelerometer-based algorithms," *Neural Comput. Appl.*, vol. 32, no. 2, pp. 335–349, Jan. 2020, doi: 10.1007/s00521-018-3759-8.

[73] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *IEEE Access*, vol. 5, pp. 8869–8879, 2017.

[74] M. M. Islam, M. A. Razzaque, M. M. Hassan, W. N. Ismail, and B. Song, "Mobile cloud-based big healthcare data processing in smart cities," *IEEE Access*, vol. 5, pp. 11887–11899, 2017, doi: 10.1109/ACCESS.2017.2707439.

[75] L. Syed, S. Jabeen, S. Manimala, and A. Alsaeedi, "Smart healthcare framework for ambient assisted living using IoMT and big data analytics techniques," *Future Gener. Comput. Syst.*, vol. 101, pp. 136–151, Dec. 2019.

[76] J. Zhang and D. Chang, "Semi-supervised patient similarity clustering algorithm based on electronic medical records," *IEEE Access*, vol. 7, pp. 90705–90714, 2019, doi: 10.1109/ACCESS.2019.2923333.

[77] K. C. Santosh, "AI-driven tools for coronavirus outbreak: Need of active learning and cross-population train/test models on multitudinal/multimodal data," *J. Med. Syst.*, vol. 44, no. 5, p. 93, May 2020.

[78] F. Dernoncourt, S. Nemati, E. B. Kassis, and M. M. Ghassemi, "Hyperparameter Selection," in *Secondary Analysis of Electronic Health Records*. Cham, Switzerland: Springer, 2016.

[79] F. Kattner, A. Cochrane, C. R. Cox, T. E. Gorman, and C. S. Green, "Perceptual learning generalization from sequential perceptual training as a change in learning rate," *Current Biol.*, vol. 27, no. 6, pp. 840–846, Mar. 2017, doi: 10.1016/j.cub.2017.01.046.

[80] C. Srilakshmi, "Wearable technology orientation using big data analytics for improving quality of human life," *Int. Res. J. Eng. Technol.*, vol. 4, pp. 1250–1254, Jan. 2017.

[81] Y. Yu, M. Li, L. Liu, Y. Li, and J. Wang, "Clinical big data and deep learning: Applications, challenges, and future outlooks," *Big Data Mining Anal.*, vol. 2, no. 4, pp. 288–305, Dec. 2019, doi: 10.26599/BDMA.2019.9020007.

[82] D. Ravì, C. Wong, B. Lo, and G.-Z. Yang, "A deep learning approach to on-node sensor data analytics for mobile or wearable devices," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 56–64, Jan. 2017, doi: 10.1109/JBHI.2016.2633287.

[83] A. Turner and S. Hayes, "The classification of minor gait alterations using wearable sensors and deep learning," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 3136–3145, Nov. 2019, doi: 10.1109/TBME.2019.2900863.

[84] A. Gumaei, M. M. Hassan, A. Alelaiwi, and H. Alsalman, "A hybrid deep learning model for human activity recognition using multimodal body sensing data," *IEEE Access*, vol. 7, pp. 99152–99160, 2019, doi: 10.1109/ACCESS.2019.2927134.

[85] W. Hong, Z. Xiong, N. Zheng, and Y. Weng, "A medical-history-based potential disease prediction algorithm," *IEEE Access*, vol. 7, pp. 131094–131101, 2019, doi: 10.1109/ACCESS.2019.2940644.

[86] M. Alhussein and G. Muhammad, "Voice pathology detection using deep learning on mobile healthcare framework," *IEEE Access*, vol. 6, pp. 41034–41041, 2018, doi: 10.1109/ACCESS.2018.2856238.

[87] D. Sierra-Sosa, B. Garcia-Zapirain, C. Castillo, I. Oleagordia, R. Nuno-Solinis, M. Urtaran-Laresgoiti, and A. Elmaghraby, "Scalable healthcare assessment for diabetic patients using deep learning on multiple GPUs," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5682–5689, Oct. 2019, doi: 10.1109/TII.2019.2919168.

[88] E. Dogdu, "Semantic Web in eHealth," in *Proc. 47th Annu. Southeast Regional Conf.*, Clemson, SC, USA, 2009, pp. 1–5.

[89] P. Nguyen, T. Tran, N. Wickramasinghe, and S. Venkatesh, "Deepr: A convolutional net for medical records," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 22–30, Jan. 2017.

[90] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2019. Accessed: Nov. 15, 2019. [Online]. Available: http://www.deeplearningbook.org

[91] S. El-Sappagh, F. Ali, A. Ali, A. Hendawi, F. A. Badria, and D. Y. Suh, "Clinical decision support system for liver fibrosis prediction in hepatitis patients: A case comparison of two soft computing techniques," *IEEE Access*, vol. 6, pp. 52911–52929, 2018.

[92] J. Zhou, S. Z. Arshad, X. Wang, Z. Li, D. Feng, and F. Chen, "End-user development for interactive data analytics: Uncertainty, correlation and user confidence," *IEEE Trans. Affect. Comput.*, vol. 9, no. 3, pp. 383–395, Jul. 2018.

[93] A. Holzinger, J. Schant, M. Schroettner, C. Seifert, and K. Verspoor, "Biomedical text mining: State-of-the-art, open problems and future challenges," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. Berlin, Germany: Springer-Verlag, 2014, pp. 271–300.

[94] G. Harerimana, B. Jang, J. W. Kim, and H. K. Park, "Health big data analytics: A technology survey," *IEEE Access*, vol. 6, pp. 65661–65678, 2018, doi: 10.1109/ACCESS.2018.2878254.

[95] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin, and A. Ghodsi, "Apache spark: A unified engine for big data processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, 2016.

[96] D. Borthakur, "The Hadoop distributed file system: Architecture and design," *Hadoop Project Website*, vol. 11, p. 21, Aug. 2007.

[97] M. T. Jones, "Process real-time big data with Twitter storm," IBM Tech. Library, New York, NY, USA, Tech. Rep., Nov. 2012. [Online]. Available: https://www.ibm.com/developerworks/library/os-twitterstorm/os-twitterstorm-pdf.pdf

[98] J. Kreps, N. Narkhede, and J. Rao, "Kafka: A distributed messaging system for log processing," in *Proc. NetDB*, Athens, Greece, Jun. 2011, pp. 1–7.

[99] E. Lindholm, J. Nickolls, S. Oberman, and J. Montrym, "NVIDIA tesla: A unified graphics and computing architecture," *IEEE Micro*, vol. 28, no. 2, pp. 39–55, Mar. 2008.

[100] Q. Cai, H. Wang, Z. Li, and X. Liu, "A survey on multimodal data-driven smart healthcare systems: Approaches and applications," *IEEE Access*, vol. 7, pp. 133583–133599, 2019.

[101] H. Mora, D. Gil. R. M. Terol, J. Azorín, and J. Szymanski, "An IoT-Based Computational Framework for Healthcare Monitoring in Mobile Environments," *Sensors*, vol. 17, no. 10, pp. 1–25, 2017.

[102] C. Habib, A. Makhoul, R. Darazi, and R. Couturier, "Health risk assessment and decision-making for patient monitoring and decision-support using wireless body sensor networks," *Inf. Fusion*, vol. 47, pp. 10–22, May 2019.

[103] Y. Miao, G. Wu, C. Liu, M. S. Hossain, and G. Muhammad, "Green cognitive body sensor network: Architecture, energy harvesting, and smart clothing-based applications," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8371–8378, Oct. 2019.

[104] I. Azimi, T. Pahikkala, A. M. Rahmani, H. Niela-Vilén, A. Axelin, and P. Liljeberg, "Missing data resilient decision-making for healthcare IoT through personalization: A case study on maternal health," *Future Gener. Comput. Syst.*, vol. 96, pp. 297–308, Jul. 2019.

[105] S. M. Thaung, H. M. Tun, K. K. K. Win, and M. M. Than, "Exploratory data analysis based on remote health care monitoring system by using IoT," *Communications*, vol. 8, no. 1, pp. 1–8, 2020.

[106] K. Zarkogianni, E. Litsa, K. Mitsis, P.-Y. Wu, C. D. Kaddi, C.-W. Cheng, M. D. Wang, and K. S. Nikita, "A review of emerging technologies for the management of diabetes mellitus," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 12, pp. 2735–2749, Dec. 2015.

[107] J. H. M. Bergmann, S. Anastasova-Ivanova, I. Spulber, V. Gulati, P. Georgiou, and A. McGregor, "An attachable clothing sensor system for measuring knee joint angles," *IEEE Sensors J.*, vol. 13, no. 10, pp. 4090–4097, Oct. 2013.

[108] G. Elumalai and R. Ramakrishnan, "A novel approach to monitor and maintain database about physiological parameters of (Javelin) athletes using Internet of Things (IoT)," *Wireless Pers. Commun.*, vol. 111, no. 1, pp. 343–355, Mar. 2020, doi: 10.1007/s11277-019-06862-5.

[109] B. Mounia and C. Habiba, "Big data privacy in healthcare moroccan context," *Procedia Comput. Sci.*, vol. 63, pp. 575–580, Jan. 2015.

[110] P. Fraga-Lamas, "Enabling technologies and cyber-physical systems for mission-critical scenarios," Ph.D. dissertation, Dept. Comput. Sci. Inf. Technol., Univ. A Coruña, A Coruña, Spain, 2017.

[111] K. Jee and G. H. Kim, "Potentiality of big data in the medical sector: Focus on how to reshape the healthcare system," *Healthcare Informat. Res.*, vol. 19, no. 2, pp. 79–85, 2013.

[112] C. S. Kruse, R. Goswamy, Y. Raval, and S. Marawi, "Challenges and opportunities of big data in health care: A systematic review," *JMIR Med. Informat.*, vol. 4, no. 4, p. e38, Nov. 2016.

[113] E. Yom-Tov, *CrowdSourced Health: How What You Do on the Internet Will Improve Medicine*. Cambridge, MA, USA: MIT Press, 2016.

[114] L. Syed, S. Jabeen, S. Manimala, and H. A. Elsayed, "Data science algorithms and techniques for smart healthcare using IoT and big data analytics," in *Smart Techniques for a Smarter Planet* (Studies in Fuzziness and Soft Computing), vol. 374, M. Mishra, B. Mishra, Y. Patel, and R. Misra, Eds. Cham, Switzerland: Springer, 2019.

[115] *Top 10 Challenges of Big Data Analytics in Healthcare*. Accessed: Dec. 10, 2019. [Online]. Available: https://healthitanalytics.com/news/top-10-challenges-of-big-data-analytics-in-healthcare

MOMINUL AHSAN received the bachelor's degree from the Department of Computer Science and Engineering, State University of Bangladesh, Dhaka, Bangladesh, in 2008, the M.Eng. degree in research from the Faculty of Engineering and Computing, Dublin City University, Dublin, Ireland, in 2014, and the Ph.D. degree from the School of Computing and Mathematical Sciences, University of Greenwich, London, U.K., in 2019. He is currently a Postdoctoral Researcher with the Department of Engineering, Manchester Metropolitan University. His research interests include prognostics, machine learning, data analytics, reliability, power electronics, wireless sensor data communication, and wearable technology.

Dr. Ahsan is currently a member of the Institution of Engineering and Technology (MIET), U.K., an Associate Fellow of the Higher Education Academy (AFHEA), U.K., and an Associate Member of the Bangladesh Computer Society. He was a recipient of the M.Eng. Stipend at Dublin City University, in 2010; the Ph.D. Scholarship with the University of Greenwich, in 2014; and the Excellent Poster Award at the IEEE International Spring Seminar on Electronic Technology, Sofia, Bulgaria, in 2017.

SIEW TEAY HON received the B.Sc. degree in control engineering and the Ph.D. degree in intelligent wireless rotating machinery condition monitoring systems from Manchester Metropolitan University, in 2013 and 2015, respectively. He is currently a Postdoctoral Researcher with the Department of Engineering, Manchester Metropolitan University. He has participated in over three projects that delivered cost effective solutions for industry. His main research interests include developing hardware and software platforms, and algorithms for data analysis and cyber security purposes.

ALHUSSEIN ALBARBAR is currently a Reader with the Department of Engineering, Manchester Metropolitan University. He has well over 27 years of industrial working experience and as an Academic Active Researcher. He led and participated in over $7M of major projects and supervised over 21 research degrees including 15 doctoral studies. He has published three books, five-book chapters, over 100 technical papers in refereed journals and international conference proceedings. His current research interests include Industry 4.0 applications, renewable power systems, smart sensing, intelligent control, and monitoring algorithms used for electromechanical power plants.