



香港城市大學  
City University of Hong Kong

專業 創新 胸懷全球  
Professional · Creative  
For The World

## CityU Scholars

### Disease diagnosis in smart healthcare Innovation, technologies and applications

Chui, Kwok Tai; Alhalabi, Wadee; Pang, Sally Shuk Han; de Pablos, Patricia Ordóñez; Liu, Ryan Wen; Zhao, Mingbo

**Published in:**  
Sustainability (Switzerland)

**Published:** 01/12/2017

**Document Version:**  
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

**License:**  
CC BY

**Publication record in CityU Scholars:**  
[Go to record](#)

**Published version (DOI):**  
[10.3390/su9122309](https://doi.org/10.3390/su9122309)

**Publication details:**  
Chui, K. T., Alhalabi, W., Pang, S. S. H., de Pablos, P. O., Liu, R. W., & Zhao, M. (2017). Disease diagnosis in smart healthcare: Innovation, technologies and applications. *Sustainability (Switzerland)*, 9(12), [2309]. <https://doi.org/10.3390/su9122309>

#### Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

#### General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

#### Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

#### Take down policy

Contact [lbscholars@cityu.edu.hk](mailto:lbscholars@cityu.edu.hk) if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Review

# Disease Diagnosis in Smart Healthcare: Innovation, Technologies and Applications

Kwok Tai Chui <sup>1,\*</sup> , Wadee Alhalabi <sup>2</sup> , Sally Shuk Han Pang <sup>3</sup>, Patricia Ordóñez de Pablos <sup>4</sup>, Ryan Wen Liu <sup>5</sup>  and Mingbo Zhao <sup>6</sup>

<sup>1</sup> Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China

<sup>2</sup> Virtual Reality Research Center, Effat University, Jeddah 21577, Saudi Arabia; walhalabi@effatUniversity.edu.sa

<sup>3</sup> School of Biological Sciences, Faculty of Science, The University of Hong Kong, Hong Kong, China; sallypsh@connect.hku.hk

<sup>4</sup> Department of Business Administration and Accountability, Faculty of Economics, The University of Oviedo, 33003 Oviedo, Spain; patriop@uniovi.es

<sup>5</sup> Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan 430063, China; wenliu@whut.edu.cn

<sup>6</sup> School of Information Science & Technology, Donghua University, Shanghai 200051, China; mzhao4@dhu.edu.cn

\* Correspondence: ktchui3-c@my.cityu.edu.hk; Tel.: +852-3442-6718

Received: 18 October 2017; Accepted: 8 December 2017; Published: 18 December 2017

**Abstract:** To promote sustainable development, the smart city implies a global vision that merges artificial intelligence, big data, decision making, information and communication technology (ICT), and the internet-of-things (IoT). The ageing issue is an aspect that researchers, companies and government should devote efforts in developing smart healthcare innovative technology and applications. In this paper, the topic of disease diagnosis in smart healthcare is reviewed. Typical emerging optimization algorithms and machine learning algorithms are summarized. Evolutionary optimization, stochastic optimization and combinatorial optimization are covered. Owing to the fact that there are plenty of applications in healthcare, four applications in the field of diseases diagnosis (which also list in the top 10 causes of global death in 2015), namely cardiovascular diseases, diabetes mellitus, Alzheimer’s disease and other forms of dementia, and tuberculosis, are considered. In addition, challenges in the deployment of disease diagnosis in healthcare have been discussed.

**Keywords:** automation; computational intelligence; data analysis; data mining; disease diagnosis; healthcare; smart living; smart city; social progress; sustainability

## 1. Introduction

There are different definitions and interpretations of a “smart city”. In the research area, the smart city is viewed as having characteristics such as: (i) adopting information and communication technology (ICT) to solve the daily life problems in governance, environment, economy, healthcare, etc.; (ii) improving the quality of life of human beings; (iii) using computational intelligence to address real-world issues by mathematical formulations and machine learning algorithms; and (iv) distributing the workload to computers, robots and machines [1].

Human health is wealth; there is nothing more valuable than good health. Researchers have devoted vast efforts in proposing new policies, algorithms, systems and architectures for healthcare. Healthcare is defined as the amelioration of health through the prevention, treatment and examination of physical damage, mental damage, illness, injury and disease.

Today's world is experiencing three challenges in healthcare: shortage of medical personnel, ageing and high total expenditure on healthcare. Reports from the World Health Organization (WHO) stated that the global need and the factual amount of health workforce were 60.4 million and 43 million, respectively, in 2013 [2]. These figures will increase to 81.8 million and 67.3 million, respectively, by 2030. Therefore, the shortage of medical personnel is unsolved and remains serious.

From 2000 to 2050, the percentage of world population over 60 years will double (from 11% to 22%) [3]. One of the major reasons is that the birth rate has remained low in past and it is estimated to keep low in the coming decades. The older the person, the higher is the chance of getting diseases, illnesses and requiring long-term caring and medical treatments. Consequently, more human resources and expenditures should be allocated to the age group of 60 or above.

The total expenditure on healthcare has occupied a significant portion as a percentage of the gross domestic product (GDP). The survey from WHO concluded that the corresponding figures in China, U.S.A., Canada, Brazil, Russian Federation, India, and Australia are 5.6%, 17.1%, 10.5%, 8.3%, 7.1%, 4.7%, and 9.4%, respectively [4]. Attributed to ageing, these figures will be increased in the coming decades.

It is good news that, with the rapid increase in computation power and availability of health data, healthcare applications can include artificial intelligence and thus become smart healthcare. This can certainly help solving some of the aforementioned challenges in healthcare. In addition, it will facilitate sustainable development [5–7]. Healthcare sustainability aims at simultaneously optimizing the financial and social impacts of the health service, without compromising the health of our patients and our ability to provide healthcare in the future. In fact, the inadequate amounts of medical personnel and increase portion of elderly increase government expenditure and lower the social productivity.

Optimization algorithms and machine learning algorithms are core tools that can benefit healthcare applications. Three kinds of optimization algorithms, evolutionary [8–17], stochastic [18–29] and combinatorial optimization [30–38] will be addressed. For machine learning algorithms, the discussion is based on un-supervised learning [39–49], supervised learning [50–70] and semi-supervised learning [71–81]. The technical content in terms of mathematical formulation of these algorithms will not be discussed in this review, but it will illustrate examples to reveal the potential opportunities between optimization/machine learning algorithms and healthcare application.

Due to the multitude of smart healthcare applications, only four applications in the field of diseases diagnosis, cardiovascular diseases [82–99], diabetes mellitus [100–112], Alzheimer's disease and other forms of dementias [113–126], and tuberculosis [127–140] are considered. These are listed on the top 10 causes of annual global death. The applications of various disease diagnosis in smart healthcare are related to automated decision making. If the performance of the classification algorithm is good in terms of overall accuracy and testing time, it may ultimately replace the role of medical doctors in disease diagnosis (and medical doctors can devote their time majorly in complicated surgery). The second case is that the classification algorithm will be utilized as rapid test (fair overall accuracy and rapid decision) for low-cost and large-scale screening.

This paper is organized as follows. Section 2 discusses the emerging optimization algorithms and machine learning algorithms in healthcare. Section 3 provides an up-to-date literature on seven key applications. Finally, a conclusion is drawn.

## 2. Emerging Optimization Algorithms and Machine Learning Algorithms

When it comes to mathematical or engineering problems, optimization algorithms play an important role because people desire to select the best solution (generally better solution) among plenty of possible solutions based on some criteria. The difficulty of the optimization problem depends on its nature and complexity. Solving convex optimization problems is generally less difficult than non-convex problems because the local minimum must be a global solution and thus first-order conditions are optimally sufficient conditions [141]. Multiobjective optimization problems (that often

have multiple conflicting objectives) are more complex than single objective optimization problem. A tradeoff solution or near optimal solution is sometimes being selected when the global optimal solution is undetermined or requires huge computation power.

The overview of the algorithms and their applications in smart healthcare is provided in the following subsections.

### 2.1. Optimization Algorithms

There are many types of optimization algorithms, for instance evolutionary optimization, stochastic optimization, and combinatorial optimization. It is worth mentioning that an algorithm can be a mixture of different types of optimization algorithms, depending on their nature and the formulation of the problem. Readers interested in other optimization techniques are encouraged to consult Bellman [142] for dynamic programming, Bertsekas [143] for gradient method, and Glover et al. [144] for metaheuristics.

#### 2.1.1. Evolutionary Optimization

Biological evolution is the inspiration of evolutionary optimization, also known as evolutionary computation. The first idea of evolutionary computation was proposed in the 1950s [8,9]. Biologically, there are many solutions for the population which is subjected to natural selection, breeding and mutation. The population evolves in a way that leads to an increase of its suitability to the environment (in the computational environment, the fitness function is used to measure this suitability). Computationally, a set of candidate solutions is initialized and updated iteratively. For new iterations (also known as generations), the less desired solutions will be removed according to the fitness values. The effectiveness of evolutionary algorithms can be examined by the convergence analysis.

The classical steps for evolutionary algorithms have been summarized in [10] and are as follows: (i) a population of candidate solutions is initialized, and each solution is encoded using alphabet of cardinality; (ii) each solution is evaluated with a fitness value; (iii) new solutions are generated by altering the existing solutions; (iv) each new solution is evaluated with a fitness value and inserted into the population; (v) replace the solutions of population with new solutions; and (vi) return to (iii) until reaching the maximum allowable time.

A problem solved by an evolutionary algorithm can be either single objective or multiobjective. Multiobjective optimization deals with optimization problems involving more than one objective function [12]. In this situation, the objective functions are conflicting. Thus, there exists many (could be infinite) Pareto optimal solutions. If none of the objective functions can be enhanced in value without scarifying other objective values, the solution is named non-dominated Pareto optimal solution.

For healthcare facility location-allocation in smart city (Hong Kong as an example) [13], four objectives were considered: (i) maximization of the population accessibility; (ii) minimization of the inequity of accessibility; (iii) minimization of the number of people beyond the threshold of travel distance; and (iv) minimization of the cost of setting up any new public healthcare facility. The problem is solved by a genetic algorithm (GA) modified to deal with multiple objectives. The cost of the mean solution is  $1.7050 \times 10^6$ , the population out of coverage is reduced from  $3.4413 \times 10^6$  to  $3.1866 \times 10^6$ , the total accessibility is increased from  $2.9633 \times 10^7$  to  $5.8485 \times 10^7$  and inequity of accessibility is reduced from  $3.5915 \times 10^6$  to  $3.1768 \times 10^6$ .

Another GA based multiobjective optimization problem was formulated in partitioning the healthcare system in Brazil [145]. Minimization of inter-microregion traveling, maximization of population homogeneity and medical procedures are defined as objectives. The improvements in the three objectives are 7.1% ( $2211 \times 10^8$  versus  $2380 \times 10^8$ ), 23.5% (76,101 versus 99,537) and 7.8% (1773 versus 1644), respectively.

Emergency services are critical for saving lives and, consequently, a modified artificial bee colony algorithm was applied to solve the multiobjective optimization problem of emergency medical

resources allocation [14]. It was important to optimization both the distances between patients and first-aid institutions and available resources requirement.

GA and differential evolution (DE) were employed to deal with the multiobjective optimization for the design of rectangular façade in common space of healthcare building [15]. The objectives were: (i) minimizing the cost of façade construction; and (ii) maximizing the daylight performance. Results indicated that GA achieved better performance in terms of convergence and hyper-volume, whereas DE provided wider range of optimal solutions.

A home healthcare scheduling multiobjective optimization problem was considered to improve the patients' satisfaction [16]. Again, GA was utilized to solve the multiobjective optimization with objectives travel cost, service cost and penalty cost. It was highlighted that home healthcare could reduce the chances for patients' admission to hospitals due to non-emergency healthcare issues and thus relieved the workload of medical personnel.

A multiobjective optimization was illustrated in the healthcare facility planning of location-allocation problem via particle swarm optimization [17]. The conflicting objectives are defined as minimization of travelled distance taken from service seekers outside the facility coverage area and maximization of the demand coverage.

### 2.1.2. Stochastic Optimization

There are many optimization problems that contain varying parameters, for instance, cost factor, demand parameter and noise factor. These kinds of parameters should be modeled in the form of random variables attributable to stochastic uncertainties. Trivial solutions include: (i) applying parameter estimation to replace the unknown parameters (nevertheless, the performance of the system is lowered); and (ii) online input correction of the system (this is generally time-consuming and has increased costs) [18]. Stochastic optimization via introduction of stochastic parameters into the optimization problem is a way out.

A summary of many techniques for stochastic optimization (for example, procedures of iterative solution, approximate problem, stochastic approximation, stochastic gradient, non-convex optimization, Laplace expansion of integrals, first order reliability method, regression and response surface method, and Taylor expansion) is presented in [19].

Attributed to random parameters variation, the optimal solutions are insensitive and thus a suitable deterministic substitute problem should be formulated. It is often found that the problem is related to probability of success, probability of failure, or expectation. This results in a non-convex optimization problem which is handled by approximation. There are two not uncommon categories of deterministic substitute problems: (i) expected total cost minimization [19–21]; and (ii) reliability-based optimization [22–24].

Recovery beds and operating rooms scheduling was formulated as a stochastic optimization problem [25]. The randomness was from capacity of resources and undetermined in the duration of surgery and recovery. For the objectives, authors have focused on the assignment and plan of surgery and recovery respectively. An orthopedic surgery department was chosen as simulation environment.

Radiotherapy is often used as a treatment to kill or control malignant cells offered to cancer patients. In [26], the management of facility in radiotherapy has been improved by introducing stochastic optimization on the waiting time before the first treatment and its duration. The uncertainty in this problem was defined as the arrival of patients and this follows a Poisson distribution.

Nurse scheduling problem has remained a great concern. A recent sample average approximation based stochastic optimization algorithm was proposed to find the optimal solution in minimizing the assignment cost [27]. There was randomness in the nurse preferences, medical demand and hospitalization.

An excellent management is always sought for in emergency department which specializes in acute care and emergency medicine. Sample average approximation was also employed to optimize the expected waiting time of acute care patients and shift scheduling [28]. The simulation

analysis has assumed that the patient arrivals into the emergency department obey inhomogeneous Poisson distribution.

Patient scheduling has been optimally designed by minimizing the completion time of medical treatments and waiting time of patients [29]. Since the service sequences and times are heterogeneous and stochastic, respectively, it increases the complexity of the optimization problem. The Multiobjective Tabu Search (MOTS) method was employed.

### 2.1.3. Combinatorial Optimization

Many real-world optimization problems are found to be exhaustive where the solution space is composed of multiple solutions. Examples of such type of problems that are solved using combinatorial optimization are minimum spanning tree problem [30] and travelling salesman problem [31]. The solvers aim at finding optimal solution from a finite set of solutions. The feasible solution should be reduced to discrete.

In general, combinatorial optimization problems can be solved by polynomial time algorithms which have the running time bounded by a polynomial expression in the size of the input. Examples include maximum matching in graph, arithmetic operations and selection sort sorting algorithm.

When it comes to public hospitals, the different surgeries taking place in an operating room can be classified into: emergency, urgent and elective [32]. The corresponding waiting times for surgery are immediate, up to a week or up to a few years. It is stressful and frustrated to wait for surgery, especially for those who are classified as elective surgery patients. A combinatorial optimization problem was formulated to solve the objectives of minimizing the number of patients waiting for elective surgery and its cost, and it was solved using genetic algorithm [33].

In private healthcare industries or hospitals, the providers can charge significant differences in price for the identical medical treatment. A report showed that, in United States, the price for an appendectomy was between USD 1529 and USD 182,955 [34]. It is not always true that the quality of service increases with the price and, therefore, in [35], a combinatorial optimization problem was formulated to minimize the patient's cost and maximize the providers' quality.

Another problem is represented by home hospice care scheduling and staffing. It was analyzed in [36] and it was concluded that there should be a balance between quality of hospice care and available manpower. Therefore, the problem was formulated as maximization of the number of patients serviced by medical personnel, while fulfilling the need of the patients.

Attention was drawn to the optimal plan for non-coplanar intensity modulated radiotherapy [37]. It could be characterized by three steps (i) iterative beam selection of solutions (i.e., non-coplanar beam directions); (ii) finding beam trajectory via combinatorial optimization; and (iii) obtaining the optimal non-coplanar intensity modulated radiotherapy plan.

Home care scheduling could be considered as a hierarchical based combinatorial optimization problem [38]. An optimal route with minimal travelled distance was required between staffs offered by medical service companies and patients at home. This is similar to the multiple traveling salesman problem.

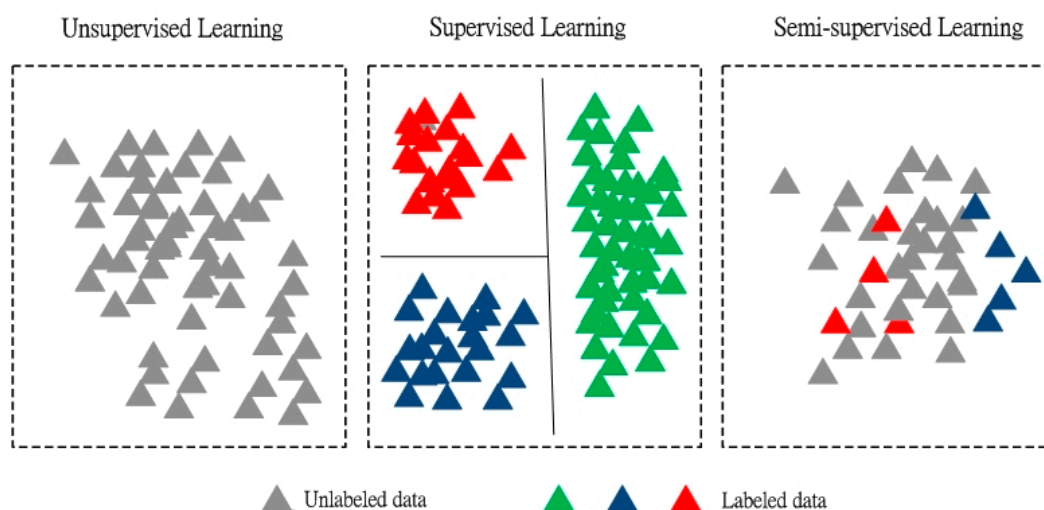
## 2.2. Machine Learning Algorithms

The term machine learning has a long history [146]. Its definition is as follows: the field of study that gives computers the ability to learn without being explicitly programmed. Many real-world problems can be modeled by machine learning algorithms. Some examples are: classification, regression, clustering, dimensionality reduction, structured prediction, face detection, decision making, speech recognition, signal de-noising, anomaly detection, deep learning and reinforcement learning.

Machine learning can be categorized into four types (i) un-supervised learning; (ii) supervised learning; (iii) semi-supervised learning; and (iv) reinforcement learning. Fundamentally, none of the data is labeled in type (i), the training/testing samples are labeled in type (ii) and there are many



unlabeled data and few labeled data in type (iii). The comparison of the first three aspects is illustrated in Figure 1.



**Figure 1.** Types of machine learning algorithm: Unsupervised, supervised and semi-supervised.

The challenge that people are confronted with is handling massive amounts of data generated by every entity in the world. Take the Internet as an example, over a billion of people are connected to the Internet every day. As a result, these people generate tremendous data. Appropriate utilization of data is valuable in terms of price and knowledge [147,148]. What we call data should make sense to proceed any further with intelligent data analysis. More important, the limited computational power of personal computer (even supercomputer) cannot solve all real-world problems. Data analytics should remove and neglect some of the data to enjoy a finite and reasonable computation time. In addition, efforts are majorly devoted to the more influencing applications to the society: healthcare, energy crisis, education, food security, overfishing, environmental pollution, migration crisis, urbanization, and water security.

### 2.2.1. Un-Supervised Learning

In reality, there exist many unlabeled data which can be intentional missing label or unintentional missing label. The former data are usually labeled initially; one may remove the label and formulate the problem as relation or correlation analysis between samples. The reasons for latter data are two-fold. First, the staff may forget to enter the label by mistake. Second, it is possible that the data owner company does not believe that the data are useful so that data collection managing is neglected.

Un-supervised learning deals with the problem of design of model to narrate the hidden pattern and relationship of unlabeled data using machine learning algorithms. Typical methods for unsupervised learning are un-supervised clustering and un-supervised anomaly detection.

Un-supervised clustering aims to organize similar unlabeled data into groups named clusters. Therefore, the data within the same cluster have similar characteristic and are dissimilar to data in other clusters. There are three basic techniques for clustering, proximity measure (similarity or dissimilarity measure), criterion function for clustering evaluation, and algorithms for clustering. It can be further divided into hierarchical clustering (divisive or agglomerative) [39], partitional clustering (centroid, model based, graph theoretic, or spectral) [40] and Bayesian clustering (decision based or non-parametric) [41].

It is curious if it is possible to influence the culture of staff of intensive care unit department to improve the patient satisfactory and safety. Unsupervised clustering was performed to analyze

the usual patterns among staff [42]. Results have concluded that we can recognize and group the influencers in developing intervention to alter the culture.

Hidden infectious diseases are frightful because medical experts are generally not familiar with their features. As a result, effective and reliable medical treatments are seldom available and further research studies are necessary. A recent study aimed at identifying hidden infectious diseases based on data retrieved from unlabeled social media messages [43].

Un-supervised anomaly detection is defined as the recognition of outliers, noise, deviations, exceptions or novelties which comply with the majority of the unlabeled data. In healthcare, medical errors [44], healthcare sensor faults [45], healthcare monitoring system [46], abnormal ECG signal [47], anomalous behavior signs [48] and clinical pathway [49] are examples of anomaly detection. The categories for anomalies are: (i) collective anomalies, a group of data points collectively facilitates the anomaly detection; (ii) contextual anomalies, it is context specific, normally in the form of time-series; and (iii) point anomalies, one instance is anomalous.

### 2.2.2. Supervised Learning

In machine learning, due to its application in classification and regression, a larger group of researchers focused on supervised learning. The datasets contain individual paired data which have input and output values. Generally, algorithms are selected to create inner relations based on the training data and to generalize unseen data.

There are five general steps for supervised learning model: (i) data collection of training and testing datasets; (ii) feature extraction; (iii) selection of machine learning algorithm; (iv) model construction using the selected algorithm; and (v) algorithm evaluation and comparison to other algorithms.

There exist packages for supervised learning in various programming languages such as MATLAB, Java, C++, and Python which provide build-in functions. Because there are numerous practical challenges, they cannot be simply treated as black-box (simply enter the input features). Some of the difficulties are: large dimensionality of feature vectors [50], bias/variance dilemma [51], input and output noise [52], large-scale training data [53], data heterogeneity [54], data redundancy [55] and non-linearity among features [56].

Applications include but are not limited to prediction of virus-host infectious association [57], human pose detection [58], virtual screening for hyperuricemia [59], speech recognition for remote healthcare [60] and biomedical information retrieval [61]. Some commonly used classification algorithms are nearest neighbors [62], discriminant analysis [63], decision trees [64], naïve Bayes classifier [65], neural network [66] and support vector machine [67]. For regression algorithms, typical examples are support vector regression [68], neural network [66], generalized linear models [69] and linear and non-linear regression [70].

Authors would like to give more elaboration on support vector machine and deep neural network which are believed to be more famous in supervised machine learning.

Many learning algorithms based on non-linear kernels require a kernel function  $K(x_i, x_j)$  well defined for all pairs of  $x_i$  and  $x_j$ . However, the computation times will be significantly increased when predicting the class for new data points. Thus, support vector machine (SVM) becomes more and more popular which offers sparse solutions, the predictions for new data points depend only on the kernel function evaluated at a subset of training data. Another characteristic of SVM is that it deals with convex optimization problem so that any local solution is a global optimal solution. To solve the convex optimization problem, the technique of Lagrange multipliers is adopted. Readers who are not familiar with this are highly encouraged to review the concept in [149].

The recent deep neural network learning application Google DeepMind's AlphaGo is in the limelight [150]. Machine learning, which is popular worldwide, is defined as computers able to learn from data to perform tasks such as prediction, clustering, classification, decision making and dimensionality reduction. Compared with the human beings, computers are superior in computation



power and can handle many complex problems. It is noted that the complete theoretical development of the algorithms is out of scope of this paper. Generally speaking, deep neural network is an artificial neural network with multiple hidden layers between input layers and output layers. It is trivial that it is more prone to overfitting and more time-consuming for training. However, it takes more advantages when handling complex problems with large dataset [151].

### 2.2.3. Semi-Supervised Learning

Semi-supervised learning is relatively young compared to supervised and un-supervised learning. Chances are that there are many unlabeled data but only few labeled data are available. This is the reason for the name semi-supervised as it lies between supervised learning (pairwise labeled inputs and outputs) and un-supervised learning (completely unlabeled data). Nevertheless, it is time consuming and costly to change unlabeled data into labeled data. For instance, web-based health-related questions are seldom being answered by users [71], providing that community-based question answering sites are important to address health information needs.

Here are the typical learning strategies of semi-supervised learning [72]. First, a supervised learning algorithm that sets some learning rules for labeled data is selected. Then, the rules are modified to include unlabeled data. The robustness of such approach depends on the reliability of labeling, i.e., it has poor robustness if it has severe labeling error.

When constructing semi-supervised learning models, at least one of the following assumptions is made: (i) Smoothness assumption is a concept that samples (in the forms of feature vector) which are close to each other have a higher opportunity have the same output label as they share similar characteristic in feature space [73]; (ii) The cluster assumption claims that the datasets have a tendency to form clusters where samples in the same clusters have higher chance to be the same output label [74]; and (iii) Similar to the first assumption, but referring to clustering, manifold assumption means that data lie on a low-dimensional manifold embedded in a higher-dimensional space [75]. It is noted that manifold is a topological space that locally resembles Euclidean space near each point.

Many existing approaches for semi-supervised learning have been discussed. Typical examples are graph-based model [76], self-training [77], co-training [78], generative model [79], low-density separation [80], and heuristic model [81].

### 2.2.4. Reinforcement Learning

Reinforcement learning is related to agents trying to maximize the total reward under interaction with uncertain and complex environment [152]. In the field of control and operation research, it is also named approximate dynamic programming [153]. Differed from standard supervised learning, reinforcement learning does not possess correct input/output pairs and sub-optimal actions.

Basically, two strategies are commonly used to solve reinforcement learning problems [154]. First, researchers aim at finding a good behavior in the environment via searching the space of behaviors. Second, estimation of the utility of taking actions in states of the world is considered with aid of dynamic programming and statistical techniques.

The lifelong function of dynamic treatment regimes was estimated using reinforcement learning [155]. Testing datasets were retrieved from about 6000 patients of Acute Myeloid Leukemia. The algorithm comprised of deep neural network and deep Q-learning which aimed at handling heterogeneous high dimensional actions and states in practical problems.

A multi-agent system was proposed for continuous patient monitoring for improvement in overall user (with poor motor skills or vision) experience [156]. To learn the behavior of user, the reinforcement learning tracked the errors in the interface use and the actions of users. It is noted that behavior consists of statistical interactions with the graphical user interface, whereas the action consists of user preferences and choices.

A reinforcement learning based neuroprosthesis controller was trained to evaluate target-oriented task performed using planar musculoskeletal human arm [157]. The results reflected that human rewards are effective measures to train the controller.

When it comes to mobile health applications, medical video streaming via adaptive rate control algorithm was studied [158]. Reinforcement learning was applied to fulfill the requirement of high quality of service. The end-to-end delay was lower than the minimum requirement of 350 ms and image quality was improved by 2.5 dB.

### 3. Smart Healthcare Applications

Since there are numerous applications in smart healthcare, this review only focuses on disease diagnosis. Cardiovascular diseases, diabetes mellitus, Alzheimer's disease and other forms of dementia and tuberculosis are on the list of top 10 causes of annual global death in 2015. Table 1 summarizes the mortality due to these diseases in 2000, 2005, 2010 and 2015 [159]. It can be seen that the first three types of diseases increase, whereas the tuberculosis decreases from 2000 to 2015. The percentage changes in 2000–2005, 2005–2010 and 2010–2015 in each type of disease are 6.7%, 8.03%, and 6.4%; 19.9%, 17.7%, and 17.8%; 29.9%, 36.1%, and 33.7%; and –5.58%, –10.6%, and –2.34%, respectively. These diseases led to more than 22 million deaths in 2015.

**Table 1.** Mortality due to cardiovascular diseases, diabetes mellitus, Alzheimer's disease and other forms of dementia, and tuberculosis in 2000, 2005, 2010 and 2015.

Disease	Deaths ('000) in Particular Year			
	2000	2005	2010	2015
Cardiovascular diseases	14,375	15,338	16,570	17,631
Diabetes mellitus	945	1133	1333	1570
Alzheimer's disease and other forms of dementia	649	843	1147	1534
Tuberculosis	1666	1573	1406	1373

The optimization algorithms and machine learning approaches that have been utilized for smart healthcare covering the diseases from Table 1 will be discussed in the following subsections.

#### 3.1. Cardiovascular Diseases

Cardiovascular diseases are a group of disorders involving heart and blood vessels. Some common conditions include atherosclerosis and may even result in stroke. Most of the risk factors of cardiovascular diseases are related to human lifestyle and the corresponding percentages of factors attributable to mortality are as follows: 40.6% for high blood pressure; 13.7% for smoking; 13.2% for poor diet; 11.9% for insufficient physical activity; 8.8% for abnormal glucose levels; and the rest for other factors [82]. Therefore, improving the eating habit and lifestyle is an effective measure for preventing cardiovascular disease. The main principle of healthy lifestyle includes low-salt and low-fat diet, regular exercising and quitting smoking [83].

Stroke is one kind of cardiovascular disease. A stroke will happen when there is an interruption of the blood supply to brain. The common causes will be burst blood vessels in brain or a blockage by a clot. This in turn cuts off the supply of oxygen and nutrients to brain cell, and causes damages to the brain tissue. Hypertension is the most notable risk factor of stroke [84] and thus having a good management of blood pressure will be a preventive measure of stroke.

The key priority for effective primary stroke prevention should be focusing on behavioral and lifestyle risk factors, including smoking, unhealthy diet, sedentary lifestyle, and improper use of alcohol [85]. Other than diet modification, engaging in physical activity for at least 30 min every day will help to prevent stroke [86].

The surface electrocardiogram (ECG) is the most widely adopted means to record electrical activity of the heart and for diagnosis of cardiovascular diseases [87]. Besides, ECG can be used for

risk stratification and selection of optimal management for different types of cardiovascular diseases. An alternative method, vectorcardiography is rarely used in clinical practice [88].

In [89], it was concluded that there are three types of features namely fiducial features (FF), non-fiducial features (NFF) and hybrid features (HF). FF is defined as any feature that is related to the characteristic points (P wave, QRS complex and T wave) of ECG signal, whereas NFF is the opposite of FF. For HF, it has both FF and NFF. Table 2 shows some of the methodologies that have been proposed for cardiovascular diseases classification [90–94] or ECG recognition [95–99].

**Table 2.** Summary of related works on cardiovascular diseases classification or heartbeat classification.

Work	Diseases	Methodology	TF	N <sub>s</sub>	Performance (%)		
					S <sub>e</sub>	S <sub>p</sub>	OA
[90]	HC + Ca + MI + Hy + Dy	LS; SVM	NFF	65	/	/	90.34
[91]	HC + Ca	RF; SVM	FF	221	87	92	94
[92]	HC + CHF	CART	FF	41	93.3	63.6	/
[93]	HC + CHF + VT + AF	FL; GA	NFF	300	/	/	93.34
[94]	HC + CA + MI	KNN	NFF	207	99.7	98.5	98.5
[95]	HB	SKF; SVM	HF	48	/	/	98.3
[96]	HB	NN	FF	17	/	/	95
[97]	HB	CNN	NFF	47	96.71	91.64	93.47
[98]	HB	NN; SVM	NFF	48	98.91	97.85	98.91
[99]	HB	FCM	FF	48	/	/	81.21

AF = atrial fibrillation; Ca = cardiomyopathy; CART = classification and regression tree; CHF = congestive heart failure; CNN = convolutional neural network; Dy = dysrhythmia; FCM = fuzzy c-means clustering; FL = fuzzy logic; GA = genetic algorithm; HB = heartbeat; HC = healthy control; Hy = Hypertrophy; KNN = K-nearest neighbor; LS = least-square; MI = myocardial infarction; NN = neural network; N<sub>s</sub> = sample size; OA = overall accuracy; RF = random forest; S<sub>e</sub> = sensitivity; SKF = switching Kalman filter; S<sub>p</sub> = specificity; SVM = support vector machine; TF = type of features; VT = ventricular tachyarrhythmia.

Eleven machine learning or optimization algorithms have been applied in [90–99]. Different algorithms may have their superiorities in particular application. The number of patients, N<sub>s</sub>, is very small compared to the number of cardiovascular diseases sufferers (more than 10 million). However, it is difficult for researcher to tackle the problem of small sample size because the datasets are generally retrieved from public domain (hospitals). It depends on the willingness of the policy of government that the data from patients can be fully disclosed to public. The performances are good (>90%) in most of the works, except [92,99].

### 3.2. Diabetes Mellitus

Diabetes mellitus, or simply called diabetes, is one of the serious epidemic diseases in the world. Over 400 million people are living with diabetes [100]. It is expected that by 2035, the total number of adults with diabetes would increase to 592 million [101]. Its subsequent macrovascular complications could be fatal. Enormous financial burden is putting on diabetes related area. Many patients have to rely on medication in their lifetimes to control diabetes.

There are two types of diabetes, insulin-dependent diabetes (or type 1 diabetes) and noninsulin-dependent diabetes (or type 2). Generally speaking, most diabetes cases belong to type 2 as it is often onset in adulthood and caused by unhealthy lifestyle, improper diet and obesity [102]. These factors then cause a combined defect of insulin secretion and insulin resistance resulting in different severity of diabetes.

Recently, there are many studies in diabetes mellitus via machine learning algorithms. The applications are summarized in Table 3. The applications on diabetes are type 2 diabetes diagnosis [103–105], prediction of fasting plasma glucose status [106], analysis of predictive power of hypertriglyceridemic waist phenotype [107], detection of hypoglycemic episodes in children [108], prediction of protein–protein interaction [109], prediction of vascular occlusion [110], prediction of development of liver cancer for diabetes sufferers [111] and detection of microalbuminuria [112] related to diabetes.

**Table 3.** Summary of related works on machine learning applications in diabetes mellitus.

Work	Applications	Methodology	N <sub>s</sub>	Performance
[103]	Diagnosis of type 2 diabetes	RF; SVM	7913	Precision: 94.2%; Recall: 93.97%
[104]	Diagnosis of type 2 diabetes	DT; KNN; NBC LR; RF; SVM	300	average AUC = 98%
[105]	Diagnosis of type 2 diabetes	ACO; FL	768	Accuracy = 84.24%
[106]	Predicting of fasting plasma glucose status	LR; NBC	4870	AUC (Female): 0.74; (Male): 0.68
[107]	Analysis of predictive power of hypertriglyceridemic waist phenotype for type 2 diabetes	LR; NBC	11,937	waist-to-hip ratio + triglyceride (men): AUC = 0.653; rib-to-hip ratio + triglyceride (women): AUC = 0.73
[108]	Detection of hypoglycemic episodes for type 1 diabetes children	NN	16	Sensitivity = 78%; Specificity = 60%
[109]	Prediction of type 2 diabetes related proteins	SVM	1296	Accuracy = 78.2%
[110]	Prediction of peripheral vascular occlusion in type 2 diabetes	SVM; WPS	33	Accuracy = 100%
[111]	Predicting the development of liver cancer in type 2 diabetes	LR; NN	2060	Sensitivity = 75.7%; Specificity = 75.5%
[112]	Detection microalbuminuria in type 2 diabetes	FL; LR; PSO	200	Sensitivity = 95%; Specificity = 85%; Accuracy = 92%

ACO = ant colony optimization; AUC = area under the curve; DT = decision tree; FL = fuzzy logic; KNN = k-nearest neighbor; LR = logistic regression; NBC = Naive Bayes classifier; NN = neural network; PSO = particle swarm optimization; RF = random forest; SVM = support vector machine; WPS = wolf pack search.

Similar to the cases in cardiovascular diseases (Table 2), there are 12 types of algorithms that have been utilized for diabetes mellitus. The number of samples has an increase tendency, which leads to more trustworthy results in evaluating the algorithms. It is still challenging (less than 80% in performance) in the fields of prediction of fasting plasma glucose status, analysis of predictive power of hypertriglyceridemic waist phenotype, detection of hypoglycemic episodes in children, prediction of protein–protein interaction and prediction of development of liver cancer for diabetes sufferers.

### 3.3. Alzheimer's Disease and Other Forms of Dementia

Dementia is a general term describing the decline in mental ability including memory and thinking skills that affect the human daily living activities. Dementia is resulted from damages of brain cells making them unable to communicate with each other [113]. Alzheimer's disease is the most common form of dementia, making around 60–80% of all cases. Some Alzheimer's disease patients in final stages may have lost basic bodily functions, including swallowing and moving their limbs. They will need around-the-clock care and this puts a huge financial burden on the medical system and social welfare department of government.

The number of dementia cases increases with aging. This leads to a huge concern among healthcare professionals. Dementia, unlike other forms of chronic illness, has a higher prevalence in developed countries. It is a pandemic disease that affects people in different regions. Moreover, studies have found that people with metabolic syndromes like diabetes and obesity are at higher risk of Alzheimer's disease [114]. Therefore, Alzheimer's disease and dementia will be one of the most challenging non-communicable diseases to battle in this era.

Table 4 shows some studies in Alzheimer's disease and other forms of dementia via machine learning algorithms. The applications include diagnosis of Alzheimer's disease [115,116], diagnosis of dementias [117], and detection of Alzheimer's disease related regions [118], prediction of mild cognitive impairment patients for conversion to Alzheimer's disease [119,120], detection of dissociable

multivariate morphological patterns [121], diagnosis of both Alzheimer's disease and mild cognitive impairment [122] and identification of genes related to Alzheimer's disease [125,126].

**Table 4.** Summary of related works on machine learning applications in Alzheimer's disease and other forms of dementias.

Work	Applications	Methodology	N <sub>s</sub>	Performance
[115]	Diagnosis of Alzheimer's disease	NBC; RF; RLO; RS; SVM	27	Accuracy = 97.14%
[116]	Diagnosis of Alzheimer's disease	SVM	53	Accuracy = 96.23%
[117]	Diagnosis of dementias	LR; SVM	29	Accuracy = 93%
[118]	Detection of Alzheimer's disease related regions	SVM	126	Accuracy = 92.36%
[119]	Predicting mild cognitive impairment patients for conversion to Alzheimer's disease	LDS	164	AUC = 0.7661
[120]	Predicting mild cognitive impairment patients for conversion to Alzheimer's disease	GA; SVM	458	Sensitivity = 76.92%; Specificity = 73.23%; Accuracy = 75%
[121]	Identification of dissociable multivariate morphological patterns	LC	801	AUC = 0.93
[122]	Identification for Alzheimer's disease and mild cognitive impairment	EM; SVM	338	Alzheimer's disease: sensitivity = 84.86%, specificity = 91.69%, accuracy = 88.73%; Mild cognitive impairment: sensitivity = 79.07%, specificity = 82.7%, accuracy = 80.91%
[123]	Identification for Alzheimer's disease and mild cognitive impairment	DCNN	900	Alzheimer's disease: sensitivity = 98.89%, specificity = 97.78%, accuracy = 98.33%; Mild cognitive impairment: sensitivity = 92.23%, specificity = 91.11%, accuracy = 92.12%
[124]	Identification for Alzheimer's disease and mild cognitive impairment	DCNN	142	Alzheimer's disease: sensitivity = 85%, specificity = 82%, accuracy = 85%; Mild cognitive impairment: sensitivity = 84%, specificity = 81%, accuracy = 85%
[125]	Identification of genes related to Alzheimer's disease	DT; QAR	33	90 genes are related to Alzheimer's disease
[126]	Identification of genes related to Alzheimer's disease	ELM; RF; SVM	31	Sensitivity = 78.77%; Specificity = 83.1%; Accuracy = 74.67%

DCNN = deep convolutional neural network; DT = decision tree; ELM = extreme learning machine; EM = expectation maximization; GA = genetic algorithm; LC = lasso classification; LDS = low density separation; LR = logistic regression; NBC = Naive Bayes classifier; QAR = quantitative association rules; RF = random forest; RLO = random linear oracle; RS = random subspace; SVM = support vector machine.

Fourteen different algorithms were employed in [115–126]. The datasets of Alzheimer's disease and other forms of dementia have relatively small sample size. Three applications, prediction of mild cognitive impairment patients for conversion to Alzheimer's disease [119,120], prediction of mild cognitive impairment [122–124] and identification of genes related to Alzheimer's disease [125,126] are required to have further improvement.

3.4. Tuberculosis

Tuberculosis (TB) was a lethal infectious disease caused by bacteria mycobacterium tuberculosis that usually attacks the lung. The transmission way is mainly through the air, and thus TB is actively spread among the community. It was a public health concern in early 19th century. With the development of drugs and hygiene awareness, the incidence rate has declined slowly since the 1990s [127].

People in poor and undeveloped countries are more vulnerable to TB because of poor environmental conditions, food insecurity, and inconvenient access to healthcare services [127]. Early detection is important for TB control. Without proper diagnosis, infected people are the major source of infection in the community [128]. It is estimated that about 3 million people remain undiagnosed or are not notified [129].



Intensive research was carried out on prevention and fighting against TB [130]. TB control is crucial with effective surveillance system for instantaneous reporting. With the development of science and technology, it is hoped that TB would not be a major public health issue in next generation.

Smart healthcare applications on tuberculosis include identification of drug resistance-associated mutations [131], detection of tuberculosis [132–135], detection of multidrug resistance tuberculosis [136], prediction of treatment failure [137], identification between tuberculosis and human immunodeficiency virus (HIV) [138], predicting recent transmission of tuberculosis [139] and detection of smear-negative pulmonary tuberculosis [140]. These are summarized in Table 5. Similar findings can be observed in applications in tuberculosis. There are 12 types of algorithms in seven applications in tuberculosis and the sample sizes of datasets are limited. The performance of two of the works [133,139] can be improved in the future. Other algorithms can be applied to obtain a favorable performance (>90%).

**Table 5.** Summary of related works on machine learning applications in tuberculosis.

Work	Applications	Methodology	N <sub>s</sub>	Performance
[131]	Identification of drug resistance-associated mutations in tuberculosis	LMM; LR; MOSS	144	Accuracy of over 90% in selected 9 drugs
[132]	Detection of tuberculosis	KNN; SVM	917; 869; 850	KNN: AUC = 0.84; 0.78; 0.82 SVM: AUC = 0.88; 0.79; 0.85
[133]	Detection of tuberculosis	CNN	4701	Accuracy = 62.07%
[134]	Detection of tuberculosis	CNN	138;662	Accuracy = 82.6%; 92.6% AUC = 84.7%; 92.6%
[135]	Detection of tuberculosis	SVM	150	Accuracy = 96.68%
[136]	Detection of multidrug resistance tuberculosis	NN	280	Sensitivity = 95.1%; Specificity = 85%
[137]	Prediction of tuberculosis treatment failure	Xpert; Response5	153	Sensitivity = 83–100%; Specificity = 26–100%
[138]	Identification between tuberculosis and HIV	Filtering	54	Accuracy = 79–93%
[139]	Predicting recent transmission of tuberculosis	LR	1552	Sensitivity = 53%; Specificity = 67%
[140]	Detection of smear-negative pulmonary tuberculosis	MLP	136	Sensitivity = 100%; Specificity = 80%; Accuracy = 88%; AUC = 91.8%

CNN = convolutional neural network; KNN = k-nearest neighbor; LMM = linear mixed model; LR = logistic regression; MLP = multilayer perception; MOSS = mode oriented stochastic search; NN = neural network; SVM = support vector machine.

#### 4. Challenges in Smart Healthcare

For the transition of healthcare to smart healthcare, numerous challenges are normally encountered and research is still ongoing. In the following subsections, privacy, pilot studies and real project, communication between data scientist and medical personnel, no free lunch theorem and increase short-term to medium-term expenditure are discussed.

##### 4.1. Privacy

Medical data contain meaningful information for modeling and analysis, which ultimately can improve medical practice and research. The privacy must be protected from misuse and violation so that patients and medical institutions will agree to release and share the data. The increase in the data availability will improve the quality of healthcare [160]. A report has stated that about 60% of patients realize that the wide-ranging employment of electronic health record will lead to more personal information being stolen or lost and about half of patients think that the privacy of medical data is not protected [161].

Ideally, medical data should be accessible to authorized parties or public institutions only if security and privacy are guaranteed. More important, researchers need medical data for carrying out data analysis, statistical analysis and machine learning applications. Common privacy-preserving methods include disclosure control [162,163], output perturbation [164,165] and anonymization [166,167].

It is hoped that, in the future, the shared datasets for medical applications (especially disease diagnosis) will be increased in sample size. The key elements will be medical data privacy and the government policies.

#### *4.2. Pilot Studies and Real Projects*

Since, in the long run, they reduce the risk of non-professional project deployment, pilot projects are important and they help to select the appropriate risk mitigation strategies and application directions. There is an increasing tendency to carry out pilot studies [168].

Professionals gain experience and intuition to improve the protocol design, algorithms, and hardware design of smart healthcare. However, in reality, only a few companies and researchers may take part in real projects in healthcare. Most researchers fail to learn from the experience since it is not uncommon that there are no timely publications, reports and feedbacks provided by the pilot projects. Usually, the projects are target-oriented [169]. However, chances are that there exist alternate solutions that are cost-effective which may replace the current solution.

It is highly encouraged that before the technology platforms are adopted, various algorithms and technologies that are tested and selected for adoption should be consolidated into technical reports and papers, and shared to the public.

Economically, it is desired that governments can support basic research and innovation [170]. Regarding the private sectors, they take up the responsible for the commercial and applied research since they are keen on market needs and demands.

For the selection of the policy tool, the features of basic research are essential. If there is a large-scale and highly focused project, conduct of research or direct government support is reliable to maintain a good project management with numerous researchers and staffs [171].

The effectiveness in research, innovation and development will be increased if government serves as long-term investor. The time lag between basic research and real deployment commercially can be large. Here are two examples for the benefit of long-term investment. The Internet revolution in the 1990s was the results of long-term investment covering twenty years [172]. The second example is that the biotechnology commercialization was facilitated by the research findings in the 1950s [173].

#### *4.3. Communication between Data Scientists and Medical Personnel*

Data scientists generally lack sufficient medical knowledge and they require medical experts. Similar situation happens for medical personnel. When data scientists would like to enter the medical institutions, medical personnel and management teams often refuse [174]. From the data scientists' perspective, medical data are useful for statistical analysis, prediction, classification and knowledge discovery. From medical personnel perspective, they are used to diagnose patients via their medical knowledge and rarely rely on the decision generated by machines.

Medical workers may argue that smart healthcare via machine learning and optimization algorithms has a conflict of interest with them; however, this is not the case [175]. First, most of the countries are suffering from a shortage of medical personnel and the medical shortage will become more accentuated. Second, smart healthcare is mainly targeting routine works so that medical workers can spend more time on professional consultation and surgery activities. Third, medical workers will earn a higher social status and better job satisfaction because patients are more satisfied with the medical services.

Both data scientists and medical personnel should make a step forward to see what the best collaboration method is. The ultimate goal is to improve the quality of life.

#### 4.4. No-Free Lunch Theorem

However, it can be seen that there are numerous approaches in different machine learning application in smart healthcare. There is a “No free lunch” theorem in machine learning which says that no unique algorithm works best for every application [176]. As a result, when people design algorithm for smart healthcare application, the analysis becomes tedious as one should go through many algorithms and select the one with best performance. An important knowhow is to only try appropriate algorithms for the problem. For instance, clustering algorithms seem inappropriate in solving classification problems.

Deep learning has gained popularity in recent years, attributable to its superiority in solving complex problem. However, it is reminded that deep learning is not always the best or necessary for all problems. It depends on the complexity of the problem, available amount of data, computational power and training time [177].

#### 4.5. Increase Short-Term to Medium-Term Expenditure

Authors would like to emphasize that owing to the fact that machine learning in smart healthcare is still young, government and institutions should spend more money as short-term to medium-term expenditure. Many applications are only in initial stage where data collection, feature extraction, and methodology are key criteria for successful deployment.

Ultimately, the smart healthcare applications will benefit human beings by increase of human life expectancy, early disease examination, ambient assisted living, patient monitoring, etc. As a result, we are in a sense making profit in the long-term using machine learning.

### 5. Conclusions

This review has provided an up-to-date literature on emerging machine learning algorithms, optimization algorithms and applications in smart healthcare. Important challenges, including privacy, pilot studies and real project and communication between data analytics and medical staffs, have been discussed. These challenges are essential for the breakthrough of smart healthcare; otherwise, the wide adoption of machine learning and optimization algorithms in real deployment is difficult to succeed. To conclude, it is believed that human beings will enjoy more and more smart healthcare applications in the coming decades. It is worth ensuring the health services are financially viable and environmentally sustainable to safeguard the future health of our patients and the health services availability in the future.

**Author Contributions:** All of the authors have contributed to the drafting of manuscript. The idea and structure of this manuscript was proposed by Kwok Tai Chui. Wadee Alhalabi advised and guided on structuring and presenting the information clearly. All authors have collected the list of research publications and written the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

### References

1. Kondepudi, S.N.; Ramanarayanan, V.; Jain, A.; Singh, G.N.; Nitin Agarwal, N.K.; Kumar, R.; Singh, R.; Bergmark, P.; Hashitani, T.; Gemma, P.; et al. *Smart Sustainable Cities: An Analysis of Definitions*; International Telecommunication Union: Geneva, Switzerland, 2014.
2. Scheffler, R.; Cometto, G.; Tulenko, K.; Bruckner, T.; Liu, J.; Keuffel, E.L.; Preker, A.; Stilwell, B.; Brasileiro, J.; Campbell, J. *Health Workforce Requirements for Universal Health Coverage and the Sustainable Development Goals*; World Health Organization: Geneva, Switzerland, 2016.
3. Beard, J.; Ferguson, L.; Marmot, M.; Nash, P.; Phillips, D.; Staudinge, U.; Dua, T.; Saxena, S.; Ogawa, H.; Petersen, P.E.; et al. *World Report on Ageing and Health 2015*; World Health Organization: Geneva, Switzerland, 2015.

4. *Total Expenditure on Health as a Percentage of Gross Domestic product (US\$)*; Global Health Observatory (GHO) Data; World Health Organization: Geneva, Switzerland, 2017. Available online: [http://www.who.int/gho/health\\_financing/total\\_expenditure/en/](http://www.who.int/gho/health_financing/total_expenditure/en/) (accessed on 15 September 2017).
5. Du, G.; Sun, C. Location planning problem of service centers for sustainable home healthcare: Evidence from the empirical analysis of Shanghai. *Sustainability* **2015**, *7*, 15812–15832. [[CrossRef](#)]
6. Castro, M.D.F.; Mateus, R.; Serôdio, F.; Bragança, L. Development of benchmarks for operating costs and resources consumption to be used in healthcare building sustainability assessment methods. *Sustainability* **2015**, *7*, 13222–13248. [[CrossRef](#)]
7. Momete, D.C. Building a sustainable healthcare model: A cross-country analysis. *Sustainability* **2016**, *8*, 836. [[CrossRef](#)]
8. Friedman, G.J. Selective Feedback Computers for Engineering Synthesis and Nervous System Analogy. Master's Thesis, University of California, Los Angeles, CA, USA, 1956.
9. Friedberg, R.M. A Learning Machine: Part I. *IBM J. Res. Dev.* **1958**, *2*, 2–13. [[CrossRef](#)]
10. Fogel, D.B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*; Wiley-IEEE Press: Hoboken, NJ, USA, 2006.
11. Bäck, T. Selective Pressure in Evolutionary Algorithms: A Characterization of Selection Methods. In Proceedings of the First IEEE Conference on Evolutionary Computation, Orlando, FL, USA, 27–29 June 1994; IEEE Press: Piscataway, NJ, USA, 1994.
12. Deb, K. Multi-objective optimization. In *Search Methodologies*; Springer: Berlin, Germany, 2014; pp. 403–449.
13. Zhang, W.; Cao, K.; Liu, S.; Huang, B. A multi-objective optimization approach for health-care facility location-allocation problems in highly developed cities such as Hong Kong. *Comput. Environ. Urban Syst.* **2016**, *59*, 220–230. [[CrossRef](#)]
14. Wen, T.; Zhang, Z.; Qiu, M.; Wu, Q.; Li, C. A Multi-Objective Optimization Method for Emergency Medical Resources Allocation. *J. Med. Imaging Health Inform.* **2017**, *7*, 393–399. [[CrossRef](#)]
15. Karaman, S.; Ekici, B.; Cubukcuoglu, C.; Koyunbaba, B.K.; Kahraman, I. Design of rectangular façade modules through computational intelligence. In Proceedings of the 2017 IEEE Congress Evolutionary Computation (CEC), San Sebastián, Spain, 5–8 June 2017; IEEE Press: Piscataway, NJ, USA, 2017.
16. Du, G.; Liang, X.; Sun, C. Scheduling Optimization of Home Health Care Service Considering Patients' Priorities and Time Windows. *Sustainability* **2017**, *9*, 253. [[CrossRef](#)]
17. Elkady, S.K.; Abdelsalam, H.M. A modified multi-objective particle swarm optimisation algorithm for healthcare facility planning. *Int. J. Bus. Syst. Res.* **2016**, *10*, 1–22. [[CrossRef](#)]
18. Marti, K. *Stochastic Optimization Methods*; Springer: Berlin, Germany, 2008.
19. Martin, B.; Correia, M.; Cruz, J. A certified Branch & Bound approach for reliability-based optimization problems. *J. Glob. Optim.* **2017**, 1–24. [[CrossRef](#)]
20. Duan, L.; Li, G.; Cheng, A.; Sun, G.; Song, K. Multi-objective system reliability-based optimization method for design of a fully parametric concept car body. *Eng. Optim.* **2017**, *49*, 1247–1263. [[CrossRef](#)]
21. Zhang, L.Z. A reliability-based optimization of membrane-type total heat exchangers under uncertain design parameters. *Energy* **2016**, *101*, 390–401. [[CrossRef](#)]
22. Rostami, M.A.; Kavousi-Fard, A.; Niknam, T. Expected cost minimization of smart grids with plug-in hybrid electric vehicles using optimal distribution feeder reconfiguration. *IEEE Trans. Ind. Inform.* **2015**, *11*, 388–397. [[CrossRef](#)]
23. Noor, E.; Flamholz, A.; Bar-Even, A.; Davidi, D.; Milo, R.; Liebermeister, W. The protein cost of metabolic fluxes: Prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.* **2016**, *12*, e1005167. [[CrossRef](#)] [[PubMed](#)]
24. Yong, K.L.; Nguyen, H.V.; Cajucom-Uy, H.Y.; Foo, V.; Tan, D.; Finkelstein, E.A.; Mehta, J.S. Cost minimization analysis of precut cornea grafts in Descemet stripping automated endothelial keratoplasty. *Medicine* **2016**, *95*. [[CrossRef](#)] [[PubMed](#)]
25. Saadouli, H.; Jerbi, B.; Dammak, A.; Masmoudi, L.; Bouaziz, A. A stochastic optimization and simulation approach for scheduling operating rooms and recovery beds in an orthopedic surgery department. *Comput. Ind. Eng.* **2015**, *80*, 72–79. [[CrossRef](#)]
26. Legrain, A.; Fortin, M.A.; Lahrichi, N.; Rousseau, L.M. Online stochastic optimization of radiotherapy patient scheduling. *Health Care Manag. Sci.* **2015**, *18*, 110–123. [[CrossRef](#)] [[PubMed](#)]

27. Bagheri, M.; Devin, A.G.; Izanloo, A. An application of stochastic programming method for nurse scheduling problem in real word hospital. *Comput. Ind. Eng.* **2016**, *96*, 192–200. [[CrossRef](#)]
28. Omar, E.R.; Garaix, T.; Augusto, V.; Xie, X. A stochastic optimization model for shift scheduling in emergency departments. *Health Care Manag. Sci.* **2015**, *18*, 289–302. [[CrossRef](#)]
29. Saremi, A.; Jula, P.; ElMekkawy, T.; Wang, G.G. Bi-criteria appointment scheduling of patients with heterogeneous service sequences. *Expert Syst. Appl.* **2015**, *42*, 4029–4041. [[CrossRef](#)]
30. Graham, R.L.; Hell, P. On the history of the minimum spanning tree problem. *Ann. Hist. Comput.* **1985**, *7*, 43–57. [[CrossRef](#)]
31. Dorigo, M.; Gambardella, L.M. Ant colonies for the travelling salesman problem. *Biosystems* **1997**, *43*, 73–81. [[CrossRef](#)]
32. Cardoen, B.; Demeulemeester, E.; Beliën, J. Operating room planning and scheduling: A literature review. *Eur. J. Oper. Res.* **2010**, *201*, 921–932. [[CrossRef](#)]
33. Malik, M.M.; Khan, M.; Abdallah, S. Aggregate capacity planning for elective surgeries: A bi-objective optimization approach to balance patients waiting with healthcare costs. *Oper. Res. Health Care* **2015**, *7*, 3–13. [[CrossRef](#)]
34. Hsia, R.Y.; Kothari, A.H.; Srebotnjak, T.; Maselli, J. Health care as a “market good”? Appendicitis as a case study. *Arch. Intern. Med.* **2012**, *172*, 818–819. [[CrossRef](#)] [[PubMed](#)]
35. Denoyel, V.; Alfandari, L.; Thiele, A. Optimizing healthcare network design under Reference Pricing and parameter uncertainty. *Eur. J. Oper. Res.* **2017**, *263*, 996–1006. [[CrossRef](#)]
36. Heching, A.; Hooker, J.N. Scheduling home hospice care with logic-based Benders decomposition. In Proceedings of the International Conference on AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, Banff, AB, Canada, 29 May–1 June 2016; Springer International Publishing: Cham, Switzerland, 2016.
37. Papp, D.; Bortfeld, T.; Unkelbach, J. A modular approach to intensity-modulated arc therapy optimization with noncoplanar trajectories. *Phys. Med. Biol.* **2015**, *60*. [[CrossRef](#)] [[PubMed](#)]
38. Jemai, J.; Chaieb, M.; Mellouli, K. The home care scheduling problem: A modeling and solving issue. In Proceedings of the 2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), Hammamet, Tunisia, 28–30 April 2013; IEEE Press: Piscataway, NJ, USA, 2013.
39. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
40. Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2973–2987. [[CrossRef](#)]
41. Wallstrom, G.L.; Hogan, W.R. Unsupervised clustering of over-the-counter healthcare products into product categories. *J. Biomed. Inform.* **2007**, *40*, 642–648. [[CrossRef](#)] [[PubMed](#)]
42. Fong, A.; Clark, L.; Cheng, T.; Franklin, E.; Fernandez, N.; Ratwani, R.; Parker, S.H. Identifying influential individuals on intensive care units: Using cluster analysis to explore culture. *J. Nurs. Manag.* **2017**, *25*, 384–391. [[CrossRef](#)] [[PubMed](#)]
43. Lim, S.; Tucker, C.S.; Kumara, S. An unsupervised machine learning model for discovering latent infectious diseases using social media data. *J. Biomed. Inform.* **2017**, *66*, 82–94. [[CrossRef](#)] [[PubMed](#)]
44. Sipes, T.; Jiang, S.; Moore, K.; Li, N.; Karimabadi, H.; Barr, J.R. Anomaly Detection in Healthcare: Detecting Erroneous Treatment Plans in Time Series Radiotherapy Data. *Int. J. Semant. Comput.* **2014**, *8*, 257–278. [[CrossRef](#)]
45. Haque, S.A.; Rahman, M.; Aziz, S.M. Sensor anomaly detection in wireless sensor networks for healthcare. *Sensors* **2015**, *15*, 8764–8786. [[CrossRef](#)] [[PubMed](#)]
46. Kadri, F.; Harrou, F.; Chaabane, S.; Sun, Y.; Tahon, C. Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems. *Neurocomputing* **2016**, *173*, 2102–2114. [[CrossRef](#)]
47. Khan, F.A.; Haldar, N.; Ali, A.; Iftikhar, M.; Zia, T.; Zomaya, A. A Continuous Change Detection Mechanism to Identify Anomalies in ECG Signals for WBAN-based Healthcare Environments. *IEEE Access* **2017**, *5*, 13531–13544. [[CrossRef](#)]
48. Ordóñez, F.J.; de Toledo, P.; Sanchis, A. Sensor-based Bayesian detection of anomalous living patterns in a home setting. *Pers. Ubiquitous Comput.* **2015**, *19*, 259–270. [[CrossRef](#)]
49. Huang, Z.; Dong, W.; Ji, L.; Yin, L.; Duan, H. On local anomaly detection and analysis for clinical pathways. *Artif. Intell. Med.* **2015**, *65*, 167–177. [[CrossRef](#)] [[PubMed](#)]



50. Fukumizu, K.; Bach, F.R.; Jordan, M.I. Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces. *J. Mach. Learn. Res.* **2004**, *5*, 73–99.
51. Geman, S.; Bienenstock, E.; Doursat, R. Neural networks and the bias/variance dilemma. *Neural Netw.* **2008**, *4*, 1–58. [[CrossRef](#)]
52. Nettleton, D.F.; Orriols-Puig, A.; Fornells, A. A study of the effect of different types of noise on the precision of supervised learning techniques. *Artif. Intell. Rev.* **2010**, *33*, 275–306. [[CrossRef](#)]
53. Akata, Z.; Perronnin, F.; Harchaoui, Z.; Schmid, C. Good practice in large-scale learning for image classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 507–520. [[CrossRef](#)] [[PubMed](#)]
54. Lewis, D.D.; Catlett, J. Heterogeneous uncertainty sampling for supervised learning. In Proceedings of the Eleventh International Conference on Machine Learning, New Brunswick, NJ, USA, 10–13 July 1994; Morgan Kaufmann Publishers: San Francisco, CA, USA, 1994.
55. Unler, A.; Murat, A.; Chinnam, R.B. MR 2 PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification. *Inf. Sci.* **2011**, *181*, 4625–4641. [[CrossRef](#)]
56. Raducanu, B.; Dornaika, F. A supervised non-linear dimensionality reduction approach for manifold learning. *Pattern Recognit.* **2012**, *45*, 2432–2444. [[CrossRef](#)]
57. Zhang, M.; Yang, L.; Ren, J.; Ahlgren, N.A.; Fuhrman, J.A.; Sun, F. Prediction of virus-host infectious association by supervised learning methods. *BMC Bioinform.* **2017**, *18*. [[CrossRef](#)] [[PubMed](#)]
58. Cai, Y.; Tan, X.; Tan, X. Selective weakly supervised human detection under arbitrary poses. *Pattern Recognit.* **2017**, *65*, 223–237. [[CrossRef](#)]
59. Ichikawa, D.; Saito, T.; Ujita, W.; Oyama, H. How can machine-learning methods assist in virtual screening for hyperuricemia? A healthcare machine learning approach. *J. Biomed. Inform.* **2016**, *64*, 20–24. [[CrossRef](#)] [[PubMed](#)]
60. Muhammad, G. Automatic speech recognition using interlaced derivative pattern for cloud based healthcare system. *Clust. Comput.* **2015**, *18*, 795–802. [[CrossRef](#)]
61. Wang, Y.; Wu, S.; Li, D.; Mehrabi, S.; Liu, H. A Part-Of-Speech term weighting scheme for biomedical information retrieval. *J. Biomed. Inform.* **2016**, *63*, 379–389. [[CrossRef](#)] [[PubMed](#)]
62. Aydın, E.A.; Kaya Keleş, M. Breast cancer detection using K-nearest neighbors data mining method obtained from the bow-tie antenna dataset. *Int. J. RF Microw. Comput. Aided Eng.* **2017**, *27*, e21098. [[CrossRef](#)]
63. Li, H.; Luo, M.; Luo, J.; Zheng, J.; Zeng, R.; Du, Q.; Ouyang, N. A discriminant analysis prediction model of non-syndromic cleft lip with or without cleft palate based on risk factors. *BMC Pregnancy Childbirth* **2016**, *16*, 368. [[CrossRef](#)] [[PubMed](#)]
64. Li, J.; Fong, S.; Mohammed, S.; Fiaidhi, J.; Chen, Q.; Tan, Z. Solving the under-fitting problem for decision tree algorithms by incremental swarm optimization in rare-event healthcare classification. *J. Med. Imaging Health Inform.* **2016**, *6*, 1102–1110. [[CrossRef](#)]
65. Miranda, E.; Irwansyah, E.; Amelga, A.Y.; Maribondang, M.M.; Salim, M. Detection of cardiovascular disease risk's level for adults using naive Bayes classifier. *Healthcare Inform. Res.* **2016**, *22*, 196–205. [[CrossRef](#)] [[PubMed](#)]
66. Choi, E.; Schuetz, A.; Stewart, W.F.; Sun, J. Using recurrent neural network models for early detection of heart failure onset. *J. Am. Med. Inform. Assoc.* **2016**, *24*, 361–370. [[CrossRef](#)] [[PubMed](#)]
67. Hsu, W.C.; Lin, L.F.; Chou, C.W.; Hsiao, Y.T.; Liu, Y.H. EEG classification of imaginary lower limb stepping movements based on fuzzy support vector machine with Kernel-induced membership function. *Int. J. Fuzzy Syst.* **2017**, *19*, 566–579. [[CrossRef](#)]
68. Gu, B.; Sheng, V.S.; Tay, K.Y.; Romano, W.; Li, S. Incremental support vector learning for ordinal regression. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 1403–1416. [[CrossRef](#)] [[PubMed](#)]
69. Eckardt, M.; Brettschneider, C.; Bussche, H.; König, H.H. Analysis of health care costs in elderly patients with multiple chronic conditions using a finite mixture of generalized linear models. *Health Econ.* **2017**, *26*, 582–599. [[CrossRef](#)] [[PubMed](#)]
70. Hahne, J.M.; Biessmann, F.; Jiang, N.; Rehbaum, H.; Farina, D.; Meinecke, F.C.; Parra, L.C. Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 269–279. [[CrossRef](#)] [[PubMed](#)]

71. Wongchaisuwat, P.; Klabjan, D.; Jonnalagadda, S.R. A Semi-Supervised Learning Approach to Enhance Health Care Community-Based Question Answering: A Case Study in Alcoholism. *JMIR Med. Inform.* **2016**, *4*. [[CrossRef](#)] [[PubMed](#)]
72. Albalade, A.; Minker, W. *Semi-Supervised and Unsupervised Machine Learning: Novel Strategies*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
73. Shi, D.; Zurada, J.; Guan, J. A Neuro-fuzzy System with Semi-supervised Learning for Bad Debt Recovery in the Healthcare Industry. In Proceedings of the 2015 48th Hawaii International Conference on System Sciences (HICSS), Kauai, HI, USA, 5–8 January 2015; IEEE Press: Piscataway, NJ, USA, 2015.
74. Reitmaier, T.; Sick, B. The responsibility weighted Mahalanobis kernel for semi-supervised training of support vector machines for classification. *Inform. Sci.* **2015**, *323*, 179–198. [[CrossRef](#)]
75. Wang, Y.; Chen, S.; Xue, H.; Fu, Z. Semi-supervised classification learning by discrimination-aware manifold regularization. *Neurocomputing* **2015**, *147*, 299–306. [[CrossRef](#)]
76. Nie, L.; Zhao, Y.L.; Akbari, M.; Shen, J.; Chua, T.S. Bridging the vocabulary gap between health seekers and healthcare knowledge. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 396–409. [[CrossRef](#)]
77. Cvetković, B.; Kaluža, B.; Gams, M.; Luštrek, M. Adapting activity recognition to a person with Multi-Classifer Adaptive Training. *J. Ambient Intell. Smart Environ.* **2015**, *7*, 171–185. [[CrossRef](#)]
78. Jin, L.; Xue, Y.; Li, Q.; Feng, L. Integrating human mobility and social media for adolescent psychological stress detection. In Proceedings of the International Conference on Database Systems for Advanced Applications, Dallas, TX, USA, 16–19 April 2016; Springer: Cham, Switzerland, 2016.
79. Ashfaq, R.A.R.; Wang, X.Z.; Huang, J.Z.; Abbas, H.; He, Y.L. Fuzziness based semi-supervised learning approach for intrusion detection system. *Inf. Sci.* **2017**, *378*, 484–497. [[CrossRef](#)]
80. Zhang, X.; Guan, N.; Jia, Z.; Qiu, X.; Luo, Z. Semi-supervised projective non-negative matrix factorization for cancer classification. *PLoS ONE* **2015**, *10*, e0138814. [[CrossRef](#)] [[PubMed](#)]
81. Yan, Y.; Chen, L.; Tjhi, W.C. Fuzzy semi-supervised co-clustering for text documents. *Fuzzy Sets Syst.* **2013**, *215*, 74–89. [[CrossRef](#)]
82. Go, A.S.; Mozaffarian, D.; Roger, V.L.; Benjamin, E.J.; Berry, J.D.; Borden, W.B.; Franco, S. Executive summary: Heart disease and stroke statistics—2013 update: A report from the American Heart Association. *Circulation* **2013**, *127*, 143–152. [[CrossRef](#)] [[PubMed](#)]
83. Thompson, P.D.; Buchner, D.; Piña, I.L.; Balady, G.J.; Williams, M.A.; Marcus, B.H.; Fletcher, G.F. Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease. *Circulation* **2003**, *107*, 3109–3116. [[CrossRef](#)] [[PubMed](#)]
84. Ohira, T.; Iso, H. Cardiovascular disease epidemiology in Asia. *Circ. J.* **2013**, *77*, 1646–1652. [[CrossRef](#)] [[PubMed](#)]
85. Feigin, V.L.; Norrving, B.; George, M.G.; Foltz, J.L.; Roth, G.A.; Mensah, G.A. Prevention of stroke: A strategic global imperative. *Nat. Rev. Neurol.* **2016**, *12*, 501–512. [[CrossRef](#)] [[PubMed](#)]
86. Wannamethee, S.G.; Shaper, A.G. Physical activity and cardiovascular disease. In *Seminars in Vascular Medicine*; Thieme Medical Publishers: New York, NY, USA, 2002.
87. Luna, A.B.D. *Clinical Electrocardiography: A Textbook*; Wiley-Blackwell: Oxford, UK, 2012.
88. Macfarlane, P.W.; Edenbrandt, L.; Pahlm, O. *12-Lead Vectorcardiography*; Butterworth Heinemann: Oxford, UK, 1995.
89. Odinaka, I.; Lai, P.H.; Kaplan, A.D.; O’Sullivan, J.A.; Sirevaag, E.J.; Rohrbaugh, J.W. ECG biometric recognition: A comparative analysis. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1812–1824. [[CrossRef](#)]
90. Tripathy, R.K.; Sharma, L.N.; Dandapat, S. A new way of quantifying diagnostic information from multilead electrocardiogram for cardiac disease classification. *Healthcare Technol. Lett.* **2014**, *1*, 98–103. [[CrossRef](#)] [[PubMed](#)]
91. Rahman, Q.A.; Tereshchenko, L.G.; Kongkatong, M.; Abraham, T.; Abraham, M.R.; Shatkay, H. Utilizing ECG-based heartbeat classification for hypertrophic cardiomyopathy identification. *IEEE Trans. Nanobiosci.* **2015**, *14*, 505–512. [[CrossRef](#)] [[PubMed](#)]
92. Melillo, P.; De Luca, N.; Bracale, M.; Pecchia, L. Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. *IEEE J. Biomed. Health Inform.* **2013**, *17*, 727–733. [[CrossRef](#)] [[PubMed](#)]

93. Vafaie, M.H.; Ataei, M.; Koofigar, H.R. Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. *Biomed. Signal Process. Control* **2014**, *14*, 291–296. [[CrossRef](#)]
94. Acharya, U.R.; Fujita, H.; Adam, M.; Lih, O.S.; Sudarshan, V.K.; Hong, T.J.; San, T.R. Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study. *Inf. Sci.* **2017**, *377*, 17–29. [[CrossRef](#)]
95. Oster, J.; Behar, J.; Sayadi, O.; Nemati, S.; Johnson, A.E.; Clifford, G.D. Semisupervised ECG ventricular beat classification with novelty detection based on switching Kalman filters. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 2125–2134. [[CrossRef](#)] [[PubMed](#)]
96. Li, P.; Wang, Y.; He, J.; Wang, L.; Tian, Y.; Zhou, T.S.; Li, J.S. High-Performance Personalized Heartbeat Classification Model for Long-Term ECG Signal. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 78–86. [[CrossRef](#)] [[PubMed](#)]
97. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adam, M.; Gertych, A.; San, T.R. A deep convolutional neural network model to classify heartbeats. *Comput. Biol. Med.* **2017**, *89*, 389–396. [[CrossRef](#)] [[PubMed](#)]
98. Elhaj, F.A.; Salim, N.; Harris, A.R.; Swee, T.T.; Ahmed, T. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Comput. Methods Programs Biomed.* **2016**, *127*, 52–63. [[CrossRef](#)] [[PubMed](#)]
99. Haldar, N.A.H.; Khan, F.A.; Ali, A.; Abbas, H. Arrhythmia classification using Mahalanobis distance based improved Fuzzy C-Means clustering for mobile health monitoring systems. *Neurocomputing* **2017**, *220*, 221–235. [[CrossRef](#)]
100. *Global Report on Diabetes*; World Health Organization: Geneva, Switzerland, 2016; Available online: [http://apps.who.int/iris/bitstream/10665/204871/1/9789241565257\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/204871/1/9789241565257_eng.pdf) (accessed on 15 September 2017).
101. Guariguata, L.; Whiting, D.R.; Hambleton, I.; Beagley, J.; Linnenkamp, U.; Shaw, J.E. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res. Clin. Pract.* **2014**, *103*, 137–149. [[CrossRef](#)] [[PubMed](#)]
102. Rodger, W. Non-insulin-dependent (type II) diabetes mellitus. *Can. Med. Assoc. J.* **1991**, *145*, 1571–1581.
103. Han, L.; Luo, S.; Yu, J.; Pan, L.; Chen, S. Rule extraction from support vector machines using ensemble learning approach: An application for diagnosis of diabetes. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 728–734. [[CrossRef](#)] [[PubMed](#)]
104. Zheng, T.; Xie, W.; Xu, L.; He, X.; Zhang, Y.; You, M.; Chen, Y. A machine learning-based framework to identify type 2 diabetes through electronic health records. *Int. J. Med. Inform.* **2017**, *97*, 120–127. [[CrossRef](#)] [[PubMed](#)]
105. Ganji, M.F.; Abadeh, M.S. A fuzzy classification system based on ant colony optimization for diabetes disease diagnosis. *Expert Syst. Appl.* **2011**, *38*, 14650–14659. [[CrossRef](#)]
106. Lee, B.J.; Ku, B.; Nam, J.; Pham, D.D.; Kim, J.Y. Prediction of fasting plasma glucose status using anthropometric measures for diagnosing type 2 diabetes. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 555–561. [[CrossRef](#)] [[PubMed](#)]
107. Lee, B.J.; Kim, J.Y. Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on machine learning. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 39–46. [[CrossRef](#)] [[PubMed](#)]
108. Ling, S.H.; San, P.P.; Nguyen, H.T. Non-invasive hypoglycemia monitoring system using extreme learning machine for Type 1 diabetes. *ISA Trans.* **2016**, *64*, 440–446. [[CrossRef](#)] [[PubMed](#)]
109. Vyas, R.; Bapat, S.; Jain, E.; Karthikeyan, M.; Tambe, S.; Kulkarni, B.D. Building and analysis of protein-protein interactions related to diabetes mellitus using support vector machine, biomedical text mining and network analysis. *Comput. Biol. Chem.* **2016**, *65*, 37–44. [[CrossRef](#)] [[PubMed](#)]
110. Li, C.M.; Du, Y.C.; Wu, J.X.; Lin, C.H.; Ho, Y.R.; Lin, Y.J.; Chen, T. Synchronizing chaotification with support vector machine and wolf pack search algorithm for estimation of peripheral vascular occlusion in diabetes mellitus. *Biomed. Signal Process. Control* **2014**, *9*, 45–55. [[CrossRef](#)]
111. Rau, H.H.; Hsu, C.Y.; Lin, Y.A.; Atique, S.; Fuad, A.; Wei, L.M.; Hsu, M.H. Development of a web-based liver cancer prediction model for type II diabetes patients by using an artificial neural network. *Comput. Methods Programs Biomed.* **2016**, *125*, 58–65. [[CrossRef](#)] [[PubMed](#)]
112. Marateb, H.R.; Mansourian, M.; Faghihimani, E.; Amini, M.; Farina, D. A hybrid intelligent system for diagnosing microalbuminuria in type 2 diabetes patients without having to measure urinary albumin. *Comput. Biol. Med.* **2014**, *45*, 34–42. [[CrossRef](#)] [[PubMed](#)]

113. Gaugler, J.; James, B.; Johnson, T.; Weuve, J. 2017 *Alzheimer's Disease Facts and Figures*; Alzheimer's Association: Chicago, IL, USA, 2017; Available online: [https://www.alz.org/documents\\_custom/2017-facts-and-figures.pdf](https://www.alz.org/documents_custom/2017-facts-and-figures.pdf) (accessed on 15 September 2017).
114. Rizzi, L.; Rosset, I.; Roriz-Cruz, M. Global epidemiology of dementia: Alzheimer's and vascular types. *BioMed Res. Int.* **2014**, *2014*, 908915. [[CrossRef](#)] [[PubMed](#)]
115. Armañanzas, R.; Iglesias, M.; Morales, D.A.; Alonso-Nanclares, L. Voxel-Based Diagnosis of Alzheimer's Disease Using Classifier Ensembles. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 778–784. [[CrossRef](#)] [[PubMed](#)]
116. Chen, Y.; Sha, M.; Zhao, X.; Ma, J.; Ni, H.; Gao, W.; Ming, D. Automated detection of pathologic white matter alterations in Alzheimer's disease using combined diffusivity and kurtosis method. *Psychiatry Res. Neuroimaging* **2017**, *264*, 35–45. [[CrossRef](#)] [[PubMed](#)]
117. Tanaka, H.; Adachi, H.; Ukita, N.; Ikeda, M.; Kazui, H.; Kudo, T.; Nakamura, S. Detecting Dementia through Interactive Computer Avatars. *IEEE J. Transl. Eng. Health Med.* **2017**, *5*, 1–11. [[CrossRef](#)] [[PubMed](#)]
118. Zhang, Y.; Dong, Z.; Phillips, P.; Wang, S.; Ji, G.; Yang, J.; Yuan, T.F. Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. *Front. Comput. Neurosci.* **2015**, *9*. [[CrossRef](#)] [[PubMed](#)]
119. Moradi, E.; Pepe, A.; Gaser, C.; Huttunen, H.; Tohka, J.; Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* **2015**, *104*, 398–412. [[CrossRef](#)] [[PubMed](#)]
120. Beheshti, I.; Demirel, H.; Matsuda, H. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Comput. Biol. Med.* **2017**, *83*, 109–119. [[CrossRef](#)] [[PubMed](#)]
121. Doan, N.T.; Engvig, A.; Zaske, K.; Persson, K.; Lund, M.J.; Kaufmann, T.; Barca, M.L. Distinguishing early and late brain aging from the Alzheimer's disease spectrum: Consistent morphological patterns across independent samples. *NeuroImage* **2017**, *158*, 282–295. [[CrossRef](#)] [[PubMed](#)]
122. Shi, B.; Chen, Y.; Zhang, P.; Smith, C.D.; Liu, J. Nonlinear feature transformation and deep fusion for Alzheimer's Disease staging analysis. *Pattern Recognit.* **2017**, *63*, 487–498. [[CrossRef](#)]
123. Billones, C.D.; Demetria, O.J.L.D.; Hostallero, D.E.D.; Naval, P.C. DemNet: A convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment. In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016; IEEE Press: Piscataway, NJ, USA, 2016.
124. Morabito, F.C.; Campolo, M.; Ieracitano, C.; Ebadi, J.M.; Bonano, L.; Bramanti, A.; Desalvo, S.; Mammone, N.; Bramanti, P. Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings. In Proceedings of the 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow (RTSI), Bologna, Italy, 7–9 September 2016; IEEE Press: Piscataway, NJ, USA, 2016.
125. Martínez-Ballesteros, M.; García-Heredia, J.M.; Nepomuceno-Chamorro, I.A.; Riquelme-Santos, J.C. Machine learning techniques to discover genes with potential prognosis role in Alzheimer's disease using different biological sources. *Inf. Fusion* **2017**, *36*, 114–129. [[CrossRef](#)]
126. Miao, Y.; Jiang, H.; Liu, H.; Yao, Y.D. An Alzheimers disease related genes identification method based on multiple classifier integration. *Comput. Methods Programs Biomed.* **2017**, *150*, 107–115. [[CrossRef](#)] [[PubMed](#)]
127. *Global Tuberculosis Report 2013*; World Health Organization: Geneva, Switzerland, 2013; Available online: [http://apps.who.int/iris/bitstream/10665/91355/1/9789241564656\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/91355/1/9789241564656_eng.pdf) (accessed on 15 September 2017).
128. Lacerda, S.N.B.; de Abreu Temoteo, R.C.; de Figueiredo, T.M.R.M.; de Luna, F.D.T.; de Sousa, M.A.N.; de Abreu, L.C.; Fonseca, F.L.A. Individual and social vulnerabilities upon acquiring tuberculosis: A literature systematic review. *Int. Arch. Med.* **2014**, *7*, 35. [[CrossRef](#)] [[PubMed](#)]
129. Hwang, L.Y.; Grimes, C.Z.; Beasley, R.P.; Graviss, E.A. Latent tuberculosis infections in hard-to-reach drug using population-detection, prevention and control. *Tuberculosis* **2009**, *89*, S41–S45. [[CrossRef](#)]
130. Sulis, G.; Roggi, A.; Matteelli, A.; Raviglione, M.C. Tuberculosis: Epidemiology and control. *Mediterr. J. Hematol. Infect. Dis.* **2014**, *6*. [[CrossRef](#)] [[PubMed](#)]
131. Sergeev, R.S.; Kavaliou, I.; Sataneuski, U.; Gabrielian, A.; Rosenthal, A.; Tartakovsky, M.; Tuzikov, A. Genome-wide Analysis of MDR and XDR Tuberculosis from Belarus: Machine-learning Approach. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2017**. [[CrossRef](#)] [[PubMed](#)]



132. Melendez, J.; van Ginneken, B.; Maduskar, P.; Philipsen, R.H.; Reither, K.; Breuninger, M.; Sánchez, C.I. A novel multiple-instance learning-based approach to computer-aided detection of tuberculosis on chest X-rays. *IEEE Trans. Med. Imaging* **2015**, *34*, 179–192. [[CrossRef](#)] [[PubMed](#)]
133. Alcantara, M.F.; Cao, Y.; Liu, C.; Liu, B.; Brunette, M.; Curioso, W.H. Improving tuberculosis diagnostics using deep learning and mobile health technologies among resource-poor communities in Perú. *Smart Health* **2017**, *1*, 66–76. [[CrossRef](#)]
134. Lopes, U.K.; Valiati, J.F. Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. *Comput. Biol. Med.* **2017**, *89*, 135–143. [[CrossRef](#)] [[PubMed](#)]
135. Yahiaoui, A.; Er, O.; Yumusak, N. A new method of automatic recognition for tuberculosis disease diagnosis using support vector machines. *Biomed. Res.* **2017**, *28*, 4208–4212.
136. Évora, L.H.R.A.; Seixas, J.M.; Kritski, A.L. Neural network models for supporting drug and multidrug resistant tuberculosis screening diagnosis. *Neurocomputing* **2017**, *265*, 116–126. [[CrossRef](#)]
137. Thompson, E.G.; Du, Y.; Malherbe, S.T.; Shankar, S.; Braun, J.; Valvo, J.; Shenai, S. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* **2017**, *107*, 48–58. [[CrossRef](#)] [[PubMed](#)]
138. Sambarey, A.; Devaprasad, A.; Mohan, A.; Ahmed, A.; Nayak, S.; Swaminathan, S.; Vyakarnam, A. Unbiased Identification of Blood-based Biomarkers for Pulmonary Tuberculosis by Modeling and Mining Molecular Interaction Networks. *EBioMedicine* **2017**, *15*, 112–126. [[CrossRef](#)] [[PubMed](#)]
139. Mamiya, H.; Schwartzman, K.; Verma, A.; Jauvin, C.; Behr, M.; Buckeridge, D. Towards probabilistic decision support in public health practice: Predicting recent transmission of tuberculosis from patient attributes. *J. Biomed. Inform.* **2015**, *53*, 237–242. [[CrossRef](#)] [[PubMed](#)]
140. JoãoFilho, B.D.O.; de Seixas, J.M.; Galliez, R.; de Bragança Pereira, B.; de Q Mello, F.C.; dos Santos, A.M.; Kritski, A.L. A screening system for smear-negative pulmonary tuberculosis using artificial neural networks. *Int. J. Infect. Dis.* **2016**, *49*, 33–39. [[CrossRef](#)]
141. Rockafellar, R.T. Lagrange multipliers and optimality. *SIAM Rev.* **1993**, *35*, 183–238. [[CrossRef](#)]
142. Bellman, R. *Dynamic Programming*; Dover Publication: New York, NY, USA, 2013.
143. Bertsekas, D.P. *Nonlinear Programming*; Athena Scientific: Belmont, CA, USA, 1999.
144. Glover, F.W.; Kochenberger, G.A. *Handbook of Metaheuristics*; Kluwer Academic Publishers: New York, NY, USA, 2006.
145. Steiner, M.T.A.; Datta, D.; Neto, R.J.S.; Scarpin, C.T.; Figueira, J.R. Multi-objective optimization in partitioning the healthcare system of Parana State in Brazil. *Omega* **2015**, *52*, 53–64. [[CrossRef](#)]
146. Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **1959**, *3*, 210–229. [[CrossRef](#)]
147. Lanza, J.; Sotres, P.; Sánchez, L.; Galache, J.A.; Santana, J.R.; Gutiérrez, V.; Muñoz, L. Managing Large Amounts of Data Generated by a Smart City Internet of Things Deployment. *Int. J. Semant. Web Inf. Syst.* **2016**, *12*, 22–42. [[CrossRef](#)]
148. Assaf, A.; Senart, A.; Troncy, R. Towards an Objective Assessment Framework for Linked Data Quality: Enriching Dataset Profiles with Quality Indicators. *Int. J. Semant. Web Inf. Syst.* **2016**, *12*, 111–133. [[CrossRef](#)]
149. Bertsekas, D.P.; Rheinboldt, W. *Constrained Optimization and Lagrange Multiplier Methods*; Athena Scientific: Belmont, MA, USA, 1982.
150. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvan, V.; Lanctot, M.; et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, 484–489. [[CrossRef](#)] [[PubMed](#)]
151. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)] [[PubMed](#)]
152. Gatti, C. *Design of Experiments for Reinforcement Learning*; Springer: New York, NY, USA, 2014.
153. Si, J. *Handbook of Learning and Approximate Dynamic Programming*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
154. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285.
155. Liu, Y.; Logan, B.; Liu, N.; Xu, Z.; Tang, J.; Wang, Y. Deep Reinforcement Learning for Dynamic Treatment Regimes on Medical Registry Data. In Proceedings of the 2017 IEEE International Conference on Healthcare Informatics (ICHI), Park City, UT, USA, 23–26 August 2017; IEEE Press: Piscataway, NJ, USA, 2017.
156. Shakshuki, E.M.; Reid, M.; Sheltami, T.R. Dynamic healthcare interface for patients. *Procedia Comput. Sci.* **2015**, *63*, 356–365. [[CrossRef](#)]



157. Jagodnik, K.; Thomas, P.; Bogert, A.V.D.; Branicky, M.; Kirsch, R. Training an Actor-Critic Reinforcement Learning Controller for Arm Movement Using Human-Generated Rewards. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1892–1905. [CrossRef] [PubMed]
158. Istepanian, R.S.; Philip, N.Y.; Martini, M.G. Medical QoS provision based on reinforcement learning in ultrasound streaming over 3.5G wireless systems. *IEEE J. Sel. Areas Commun.* **2009**, *27*. [CrossRef]
159. *Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000–2015*; World Health Organization: Geneva, Switzerland, 2016; Available online: [http://www.who.int/healthinfo/global\\_burden\\_disease/estimates/en/index1.html](http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html) (accessed on 15 September 2017).
160. Makoul, G.; Curry, R.H.; Tang, P.C. The use of electronic medical records: Communication patterns in outpatient encounters. *J. Am. Med. Inform. Assoc.* **2001**, *8*, 610–615. [CrossRef] [PubMed]
161. Westin, A.; Krane, D.; Capps, K.; Peterson, T.; Kliner, S. *Making It Meaningful: How Consumers Value and Trust Health It Survey*; National Partnership for Women & Families: Washington, DC, USA, 2012; Available online: [http://go.nationalpartnership.org/site/DocServer/HIT\\_Making\\_IT\\_Meaningful\\_National\\_Partnership\\_February\\_2.pdf](http://go.nationalpartnership.org/site/DocServer/HIT_Making_IT_Meaningful_National_Partnership_February_2.pdf) (accessed on 15 September 2017).
162. Sweeney, L. Computational disclosure control for medical microdata: The Datafly system. In *Record Linkage Techniques 1997: Proceedings of an International Workshop and Exposition*; The National Academies Press: Washington, DC, USA, 1997; pp. 442–453.
163. Chang, C.C.; Li, Y.C.; Huang, W.H. TFRP: An efficient microaggregation algorithm for statistical disclosure control. *J. Syst. Softw.* **2007**, *80*, 1866–1878. [CrossRef]
164. Yang, J.J.; Li, J.Q.; Niu, Y. A hybrid solution for privacy preserving medical data sharing in the cloud environment. *Future Gener. Comput. Syst.* **2015**, *43*, 74–86. [CrossRef]
165. Liu, K.; Kargupta, H.; Ryan, J. Random projection-based multiplicative data perturbation for privacy preserving distributed data mining. *IEEE Trans. Knowl. Data Eng.* **2006**, *18*, 92–106. [CrossRef]
166. Fung, B.C.; Wang, K.; Philip, S.Y. Anonymizing classification data for privacy preservation. *IEEE Trans. Knowl. Data Eng.* **2007**, *19*, 711–725. [CrossRef]
167. Loukides, G.; Gkoulalas-Divanis, A.; Malin, B. Anonymization of electronic medical records for validating genome-wide association studies. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 7898–7903. [CrossRef] [PubMed]
168. Turner, J.R. The role of pilot studies in reducing risk on projects and programmes. *Int. J. Proj. Manag.* **2005**, *23*, 1–6. [CrossRef]
169. Zingg, W.; Holmes, A.; Dettenkofer, M.; Goetting, T.; Secci, F.; Clack, L.; Allegranzi, B.; Magiorakos, A.; Pittet, D. Hospital organization, management, and structure for prevention of health-care-associated infection: A systematic review and expert consensus. *Lancet Infect. Dis.* **2015**, *15*, 212–224. [CrossRef]
170. Lichtenberg, F.R. The effect of government funding on private industrial research and development: A re-assessment. *J. Ind. Econ.* **1987**, 97–104. [CrossRef]
171. VomBrocke, J.; Lippe, S. Managing collaborative research projects: A synthesis of project management literature and directives for future research. *Int. J. Proj. Manag.* **2015**, *33*, 1022–1039. [CrossRef]
172. Forkner-Dunn, J. Internet-based patient self-care: The next generation of health care delivery. *J. Med. Internet Res.* **2005**, *5*. [CrossRef] [PubMed]
173. James, C. *Global Status of Commercialized Biotech/GM Crops*; The International Service for the Acquisition of Agri-biotech Applications (ISAAA): Metro Manila, Philippines, 2016. Available online: <http://africenter.isaaa.org/wp-content/uploads/2017/06/ISAAA-Briefs-No-52.pdf> (accessed on 15 September 2017).
174. Huesch, M.D.; Mosher, T.J. Using It or Losing It? The Case for Data Scientists inside Health Care. 2017. Available online: <http://catalyst.nejm.org/case-data-scientists-inside-health-care/> (accessed on 1 December 2017).
175. Oaldem-Rayner, L. Artificial Intelligence Won't Replace Doctors Soon But It Can Help with Diagnosis. 2017. Available online: <http://www.abc.net.au/news/2017-09-19/ai-wont-replace-doctors-soon-but-it-can-help-diagnosis/8960530> (accessed on 1 December 2017).
176. Wolpert, D.H. *The Supervised Learning No-Free Lunch*; Springer: London, UK, 2002.
177. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

