Data Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine Learning Applications in Type 1 Diabetes

Ashenafi Zebene Woldaregay a, Eirik Årsand b, David Albers c, Lena Mamykina c, Taxiarchis Botsis a, Gunnar Hartvigsen a

a Department of Computer Science, University of Tromsø – The Arctic University of Norway, Tromsø, Norway
b Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway
c Department of Biomedical Informatics, Columbia University, N.Y., USA

Abstract

Background

Diabetes mellitus is a chronic metabolic disorder that results in abnormal blood glucose regulations. Blood glucose level (BG) is preferably maintained close to normality through self-management practices, which involves actively tracking BG levels and taking proper actions including adjusting diet and insulin medications. It is obvious that a tight BG control could sometimes result in out of range levels, i.e., hyper- or hypoglycemia episodes. Factors such as carbohydrate intake, stress and infections could result in hyperglycemia episodes. BG anomalies could be defined as any undesirable reading either due to a precisely known (normal cause variation) or unknown reason (special cause variation) to the patient. Even if the advancement in self-management applications and diabetes monitoring technologies have made things easier, the challenge of BG anomalies remains to be managed by the patient her-/himself. There are increasingly more technological developments in the direction of personalized decision systems and BG event alarms to provide an alert and decision support to the patient for these challenges. Techniques like detection of glycemic variability, hypoglycemia, and hyperglycemia in particular and blood glucose anomalies in general are central to the development of these diabetes technologies. The ubiquitous nature and widespread use of mobile health applications (mHealth apps), sensors and wearables and other point of care (POC) devices for self-monitoring and management purposes have made possible the generation of automated and continuous diabetes related data, which brought an opportunity for the introduction of machine learning applications for an intelligent and improved systems, capable of solving complex tasks within a dynamic knowledge and environment.

Motivation

Recently, machine learning applications have been widely introduced within the diabetes research in general and BG anomalies detection in particular. However, irrespective of their expanding and increasing popularity, there is lack of updated reviews that materialize the current trends in modelling options and strategies in BG anomalies detection within the context of personalized decision support systems and BG alarm events applications in type 1 diabetes.
Objective

The objective of this review is to identify, assess and analyze the state of the art machine learning strategies and its hybrid systems focusing on blood glucose anomalies classification and detection including glycemic variability, hyperglycemia, and hypoglycemia in type 1 diabetes, which are important constituents for optimal diabetes self-management. The review covers machine learning approaches pertinent to personalized decision support systems and BG alarm events applications.

Method

A rigorous literature was conducted between September 1 and October 1, 2017, through various online databases including Google scholar, PubMed, ScienceDirect and others. Peer reviewed journals and articles were considered. Relevant articles were first identified by reviewing the title, keywords, and abstracts as a preliminary filter with our selection criteria, and then reviewed the full text articles that fulfilled the inclusion criteria. Information from the selected literature was extracted based on some predefined categories, which were based on previous research and further elaborated through brainstorming.

Results

The initial hit was vetted using the title, abstract, and keywords, and retrieved a total of 467 papers (DBLP Computer Science (17), Diabetes Technology and Therapeutics (23), Google scholar (150), IEEE (211), Journal of Diabetes Science and Technology (21), PubMed Medlin (19), ScienceDirect (26)) (see Figure 3 below). After removing duplicates from the list, 389 records remained. Then, we did an independent assessment of the articles and screening based on the inclusion and exclusion criteria, which eliminated another 210 papers, leaving 179 relevant papers. After a full-text assessment, 45 articles were left (hyperglycemia = 5, glycemic variability = 3, hypoglycemia = 37), which were critically analyzed. The inter-rater agreement was measured using Cohen Kappa test, and disagreements were resolved through discussion.

Conclusion

Despite the complexity of blood glucose dynamics, there are many attempts to capture hypoglycemia, hyperglycemia incidences and the extent of an individual’s glycemic variability using different approaches. Recently, due to the ubiquitous nature of self-management mHealth apps, sensors and wearables has paved the way for the continuous accumulation of self-collected health data, which in turn contributed for the widespread research of machine learning applications. The state of the art indicates that various classes of machine learning have been developed and tested in different BG pattern classification and anomalies detection tasks. These class includes feed forward artificial neural network, hybrid systems, support vector machine (SVM), decision tree, genetic algorithm, adaptive neural fuzzy inference system (ANFIS), nonlinear autoregressive network with exogenous inputs (NARX), and nonlinear autoregressive network (NAR), Gaussian process regression, deep belief network, and Bayesian neural network (BNN). These techniques have explored various kinds of input parameters such as blood glucose (BG), heart rate, QT interval, insulin, diet, physical activity, galvanic response, and skin impedance. Most of the identified studies used a theoretical threshold either suggest by physician or various concerned bodies like the American diabetes association. However, the problem with this kind of approaches is that the specified threshold
may vary from patient to patient and also some patients might feel no symptoms. Therefore, such models should consider the difference among patients and also track its temporal change overtime. Moreover, the studies should also give more emphasis on the time lag (TL) and the various types of inputs used. Generally, we foresee these developments might encourage researchers to further develop and test these systems on a large-scale basis.

1. Introduction

Diabetes mellitus is a chronic metabolic disorder that results in an abnormal blood glucose (BG) regulation. BG level is maintained close to normality through self-management practices, which involves actively tracking BG levels and taking proper actions including diet and insulin medications. The estimated number of people with diabetes aged between 20 and 79 was 415 million (uncertainty interval: 340–536 million) in 2015 and is expected to reach 642 million (uncertainty interval: 521-829 million) by 2040 [1]. The global economic burden of diabetes in adults aged between 20 and 79 was estimated to reach US$1.31 trillion (95% CI 1·28–1·36) in 2015 [2]. The total number of deaths attributed to diabetes is estimated to be 5 million in people with diabetes aged between 20 and 79 [1]. People with diabetes have a higher risk of getting infections as compared to the normal population, which potentially increases their morbidity and mortality [3]. The greater and frequent risk of infections is mainly correlated with a hyperglycemia environment [3, 4]. Moreover, studies suggests hypoglycemia episode could result in a higher hospitalization and mortality rate [5].

The individual’s BG dynamic is affected by various factors, which are mainly categorized as common, individual and unpredictable factors [6]. The common factors include such as amount of food intake, insulin intake, previous level of BG, pregnancy, drug and vitamin intake, smoking, alcohol intake. The individual factors include dawn phenomena, physical exercise load and menstruation. The unpredictable factors include stress, concomitant diseases and infections [6]. Swings in blood glucose dynamics, i.e., hypoglycemia and hyperglycemia, could be generally categorized under a normal cause variation and special cause variation. The normal cause variation is regarded as caused by those common and individual factors, whereas the special cause variation is caused by those unpredictable factors. The underlying reason of the special cause variations is difficult to understand and remains a challenge for the patient during the incidences. For instance, during stress and infections the patient usually struggles with hyperglycemia and injects frequent insulin to lower his/her BG levels.

BG anomalies could be defined as any undesirable reading either due to a precisely known (normal cause variation) or unknown reason (special cause variation) to the patient himself [7]. Even if the advancement in self-management applications and diabetes monitoring technologies have made things easier, the challenge of BG anomalies remains to be managed by the patient himself. There are some technological developments in the direction of personalized decision systems and BG event alarms to provide an alert and decision support to the patient in the time of these challenges. Techniques like classification and detection of glycemic variability, hypoglycemia, and hyperglycemia in particular and BG anomalies in general are central to the development of these diabetes technologies. The ubiquitous nature and widespread use of mobile health applications (mHealth apps), sensors and wearables and other POC devices for self-monitoring and management purposes have made possible the generation of automated and continuous diabetes related data, which brought an opportunity for the introduction of machine learning and its
application for an intelligent and improved systems, which is capable of solving complex tasks within a dynamic knowledge and environment. As far as our knowledge is concerned, there are no reviews conducted towards techniques of BG anomalies classifications and detections focusing on various approaches in general and machine learning applications in particular. However, there is review conducted to evaluate the significant effect of pattern management based on self-monitoring blood glucose (SMBG) in regard to clinical practices [8]. Therefore, we suggest that there is lack of reviews focusing on BG anomalies detection. The objective of this review is to identify, assess and analyze the state of the art machine learning strategies in BG anomalies detection including glycemic variability, hyperglycemia, and hypoglycemia detection in type 1 diabetes. Moreover, it presents the current modelling options of machine learning applications and its hybrid system focusing on these tasks. The review covers machine learning approaches pertinent to personalized decision support systems and BG alarm events applications in type 1 diabetes.

2. Categories of Machine Learning Tasks in Type 1 diabetes

Machine learning approaches (tasks) are generally categorized as regression (prediction), classification, detection, and clustering, which are grouped either in supervised, unsupervised, and reinforcement learning based on the type of learning employed. Generally, reinforcement learning is out of the scope of this review, where we mainly focus on the other two categories. Machine learning based data mining tasks could be categorized as descriptive or unsupervised (i.e., clustering, association, summarization) and predictive or supervised learning (i.e., classification, regression) [9]. In this regard, most widely used machine learning based data mining tasks in the literatures are blood glucose anomalies detection, blood glucose prediction, modelling blood glucose dynamics and decision making/Education, as shown in Figure 1. In this review, we will focus on the typical applications of classification and detection tasks in diabetes research, specifically in blood glucose anomalies detection within the context of personalized decision support system, and BG alarm events applications. The review considers various class of machine learning algorithms; artificial neural network, Bayesian neural network, decision tress, support vector machine and others. Moreover, a Hidden Markov Model (HMM) trained with a framework close to machine learning families is also considered.
Figure 1. Most widely used machine learning based data mining tasks based on self-recorded data in people with type 1 diabetes (Modified version of Figure 2 in [9]). The green shaded ellipse depicts the scope of this review.

2.1. Blood Glucose Anomalies Classification and Detection

Generally, anomalies could be defined as observation that deviates much from the other observations as to arouse suspicious that it could be generated by a different process [7, 10, 11]. There are terms that often used interchangeably with anomalies; such as outliers, deviations, exceptions, rare instances and irregularities. The problem of identifying and capturing anomalies in data can be a supervised, semi-supervised and unsupervised tasks [12, 13]. These strategies can be categorized as classifier or model-based approach. The semi-supervised (model) based approach is better when anomalous instances are not easily available whereas a supervised (classifier) based approach is more suitable when there are sufficient labeled instances of both normal and anomalous instances. Whereas the unsupervised approach doesn’t need any reference data labels, normal behaviors has to be determined dynamically, and the detections are mainly performed with respect to the entire datasets. The model-based strategies can be viewed as a diagnosis of system’s behavior during abnormal situations through modelling and adequately characterizing the system behavior during normal situations [12, 14]. It uses a system’s model to either estimate or predict of the underlying system (process) dynamics to capture anomalies in the data. The most important design requirement in using a model include discovering and characterizing what is to be considered a normal pattern of behaviors [15]. Unlike the classifier based strategies, the model-based strategies doesn’t require a rigorous knowledge of the underlying expected anomalies, i.e., to fully understand and characterize the shape and nature of the expected anomalies [15]. By simply defining what is the expected normal pattern the system should exhibit, the model-based anomaly detection is capable of detecting the abnormal behavior, which is not considered as normal behavior of the system. Defining and discovering what is “normal” is a challenging task especially for a dynamic and complex systems, e.g., BG dynamics. However, this is often tackled in a dynamic and complex system by relying on either a machine learning model trained on a large and enough datasets or using an explicit mathematical model of the system if it exists already.

BG readings are a time series data and anomalies in diabetes BG levels could be regarded as any undesirable readings, as shown in the Figure 1 above, either due to a predictable (normal cause variation) or unpredictable cause (special cause variation). A normal cause variation could be defined as any hypoglycemia or hyperglycemia incidences with the underlying cause known to the patient her-/himself and also referred as predictable (patient controllable) factors such as insulin injection, diet intake, physical activity and others. However, special cause variation refers to any hypoglycemia or hyperglycemia incidences with the underlying cause unknown to the patient and also called unpredictable (patient uncontrollable) factors such as stress, infections and others. The classifier, model and unsupervised based approach could be used to solve the challenge of capturing BG anomalies caused by both the predictable factors (normal cause variation) and unpredictable cause (special cause variation). However,
regarding the unpredictable cause (special cause variation), the classifier-based approach remains to be very challenging with limited feasibility since the classifier-based strategies requires a thorough understanding and characterization of the nature, size and shape of the anomalies, along with its inter and intra variability among the patients. With the same token, the unsupervised approach could face the same challenge since it doesn’t have any mechanisms of differentiating the special cause one from the normal cause variations. However, the model-based approach happens to be more appropriate given that it only requires to characterize what is considered to be normal to detect what is believed to be abnormal. For example, how do an infection (stress) related hyperglycemia and diet induced hyperglycemia treated according to the model-based anomalies detection strategies? According to the model-based strategies, diet related hyperglycemia could be considered as normal, since the model could describe the underlying cause, but infection related hyperglycemia is considered as anomalies because the model can’t describe the underlying cause based on patient controllable variables.

2.2. Glycemic Variability Classification and Detection

Glycemic variability (GV) could be defined as the degree of oscillations or fluctuation of a patient’s BG between high and low levels [16]. Glycemic variability measure could provide an all-inclusive information regarding Postprandial spikes in blood glucose, as well as episode of hypoglycemic and hyperglycemic events [16, 17], which are the factors blamed for an elevated chance of cardiovascular events in DM. The evaluation of GV helps to understand and estimate the effect of the patient daily action on the hypo and hyperglycemia incidence by associating out-of-target BG levels with patient-specific factors, such as activity, food, stress, illness, and medication [16]. However, there is no gold standard approaches for assessing GV, and despite its importance it remains to be challenging.

3. Method

The objective of this review is to identify, assess and analyze the state of the art machine learning strategies and its hybrid system focusing on blood glucose anomalies detection including glycemic variability, hyperglycemia, and hypoglycemia classification and detection in type 1 diabetes. The review covers machine learning approaches pertinent to personalized decision support systems and BG alarm events applications. Therefore, for the purpose of the study, a rigorous literature was conducted between September 1 and October 1, 2017, through various online databases including Google scholar, IEEE Xplore, DBLP Computer Science Bibliography, ScienceDirect, PubMed/Medline, Journal of Diabetes Science and Technology, Diabetes Technology & Therapeutics. Moreover, the reference list of the selected articles is used to extract additional articles to get a complete overview of the field. Peer reviewed journals and articles published between 2000 and 2017 were considered. The inclusion and exclusion criteria were setup through rigorous discussion and brainstorming among the authors. Different combination of terms like “Diabetes”, “Intelligent system”, “Hybrid system”, “Machine learning”, “BG event indicators (hypo and hyperglycemia prediction)”, “BG event alarm”, “BG personalized decision system”, “Clinical”, “Closed loop system”, “Hyperglycemia”, “Hypoglycemia”, “Glycemic variability”, and “Personalized profile” were used during the search. The terms were combined using “AND” and “OR” for a better searching strategy. Relevant articles were
first identified by reviewing the title, keywords, and abstracts for a preliminary filter with our selection criteria, and then reviewed full text articles that seemed relevant. Information from the selected literature was extracted based on some predefined categories, which were based on previous research and further elaborated through brainstorming.

### 3.1. Inclusion and Exclusion Criteria

To be included in the study the studies should develop, implement, test and discuss machine learning and any of its hybrid approach in type 1 diabetes focusing in one or more of the following application areas:

- Blood glucose anomalies detection
- Hypoglycemia prediction, classification or detection
- Hyperglycemia prediction, classification or detection
- Glycemic/ Blood Glucose variability classification or detection

Therefore, studies that resides outside of these stated scopes are excluded from the review including all articles written in other languages but English.

### 3.2. Data Categorization and Data Collection

Information were extracted from the selected studies based on a predefined parameters (variables) and categories. The categories were defined based on rigorous brainstorming and discussion among the authors. These categories were demarcated solely to collect the relevant data and to assess, analyze and evaluate the model’s characteristics and its experimental setup.

*Application scenario:* This category defines the type of applications the machine learning algorithm is exploited. It can be hypoglycemia and hyperglycemia prediction, classification or detection, and glycemic variability classification, and detection.

*Type of input:* This category is defined to assess, analyze and evaluate the type of inputs used to develop the algorithm. This includes the key diabetes parameters and other physiological parameters relevant for BG anomalies classification and detections; blood glucose, heart rate variability and others.

*Data Format or Type/Data source/Data size:* This category is defined to assess, analyze and evaluate the type of data format used as input to the algorithm. This depends based on the type of diabetes technologies, mobile applications and POC devices used for data collection and algorithm development. It includes different data format like from continuous glucose monitoring devices, i.e. CGM, mHealth applications, i.e. diabetes diary, heart rate monitoring devices and others.

*Input Preprocessing:* This category defines the kind of preprocessing algorithm the system implements so as to avoid missing, sparse and corrupted input data.

*Class of Machine Learning:* This category defines the class of machine learning algorithm used to train and test the BG anomalies classification and detection algorithm. It includes different class of machine learning algorithm; Artificial Neural Network, Support Vector Machine, Bayesian network, decision tree and others.

*Training/Learning Method/Algorithm:* This category defines the class of learning algorithms used to train the hypoglycemia, hyperglycemia, glycemic variability classification and detection algorithms. It includes different training algorithms like the backpropagation algorithm, kernel, optimization techniques and others.
Performance Metrics/Evaluation Criteria: This category defines the type of evaluation metrics used to determine the accuracy of the classification and detection algorithm implemented. It includes different performance metrics like specificity, sensitivity, ROC curves and others.

3.3. Literature Evaluation

The included literatures were analyzed and evaluated based on the above defined categories and variables in order to uncover the state of the art machine learning applications in hyper/hypoglycemia prediction, classification and detection, and glycemic variability classification and detection. It also tries to pinpoint their characteristics along with the experimental setup used to implement and test the algorithms. The first evaluation and analysis was carried out based on the type of input used to develop the algorithms, in order to uncover the state of the art inputs used in these circumstances. The second evaluation and analysis was carried out based on the various class of machine learning used to develop these algorithms in order to uncover the rate of adoption and their suitability to the task. The third evaluation and analysis was carried out based on the performance metrics used to evaluate the performance of these algorithms.

4. Results

4.1. Relevant Literatures

The initial hit was vetted using the title, abstract, and keywords, and retrieved a total of 467 papers (DBLP Computer Science (17), Diabetes Technology and Therapeutics (23), Google scholar (150), IEEE (211), Journal of Diabetes Science and Technology (21), PubMed Medlin (19), ScienceDirect (26)) (see Figure 3 below). After removing duplicates from the list, 389 records remained. Then, we did an independent assessment of the articles and screening based on the inclusion and exclusion criteria, which eliminated another 210 papers, leaving 179 relevant papers. After a full-text assessment, 45 articles were left (hyperglycemia = 5, glycemic variabilities = 3, hypoglycemia = 37), which were critically analyzed, as shown in the Figure 2 below. The inter-rater agreement was measured using Cohen Kappa test, and disagreements were resolved through discussion.

![Bar chart showing the distribution of relevant papers by time period and type of classification and detection.]
Figure 2: The number of articles published per year of publication

Figure 3: Flow Diagram of the review process.

Table 1: Reported Input features, Machine learning class and Accuracy.

<table>
<thead>
<tr>
<th>Study</th>
<th>Features</th>
<th>Type of Machine Learning</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skin Impedance, Glucose response</td>
<td>MLP, NARX &amp; GARCH</td>
<td>76% 51% 93%</td>
</tr>
<tr>
<td>[18]</td>
<td>Diet</td>
<td>SVM, RNN, ELM, DBN, VTN, NN</td>
<td>79% 52% 90%</td>
</tr>
<tr>
<td>[19] &amp; [20]</td>
<td>Heart Rate (HR)</td>
<td>Rule Based, Hybrid</td>
<td>76% 51% 92%</td>
</tr>
</tbody>
</table>
4.2. Evaluation and Analysis of the Literatures

The literatures, as described previously, were evaluated based on the type of machine learning used to develop the algorithm, the type of input used to train the system and the performance metrics used to evaluate the algorithm performance based on Appendix 1 (See Table 1 & 2).

4.2.1. Data characteristics and Input Parameters

4.2.1.1. Input parameters

Selecting the proper types of input parameters is one of the crucial design strategies for successful classification and detection algorithm development. In this regard, the outer bigger ring, middle ring, and the inner ring, in the Figure 4 depicts the types of input used in hypoglycemia, hyperglycemia, glycemic variability classification and detection algorithm respectively. According to the outer ring, hypoglycemia classification and detection algorithm, blood glucose (BG), heart rate and QT interval are the most used types of input parameters (65%), as shown in the Figure 4 below. Blood glucose alone is the second most used types of input parameter (11%). Blood glucose and insulin are the third most used types of input parameters (8%). Blood glucose, insulin, diet, physical activity and others are the fourth most used types of input parameters (5%). Blood glucose, insulin, diet, heart rate, galvanic response, skin impedance, are the fifth most used types of input parameters (3%). Blood glucose and diet alone are the sixth most used types of input parameters (2%). According to the middle ring, hyperglycemia classification and detection algorithm, Blood glucose alone, and Blood glucose, insulin represent the most used types of input parameters (40%), as shown in the Figure 4 below. Blood glucose, heart rate, and QT interval represents the second most used types of input parameters (20%). According to the inner ring, glycemic variability classification and detection algorithm, blood glucose alone (50%) and blood glucose, & insulin (50%) are the most used types of input parameters, as shown in the Figure 4 below.
4.2.1.2. Data Characteristics

Data Sources

Different kinds of data sources ranging from blood glucose monitors, physical activity, electrocardiogram, and heart rate sensor have been used in the reviewed articles for hyperglycemia, hypoglycemia and glycemic variability classification and detection algorithms. The reviewed articles relied on different kinds of data format including SMBG (finger sticks), CGM, and ECG signals, as shown in the Table 2 below. Generally, ECG signal is the most used type of data format (53%), followed by CGM (38%) and SMBG (9%). Specifically, hypoglycemia classification and detection involves (CGM (n=10), ECG (n=24), and SMBG (n=4)). Regarding hyperglycemia classification and detection (CGM (n=5), and ECG (n=1)) and glycemic variability classification and detection (CGM (n=3)).

<table>
<thead>
<tr>
<th>Data type/format</th>
<th>Count (n)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGM</td>
<td>18</td>
<td>38%</td>
</tr>
<tr>
<td>SMBG</td>
<td>4</td>
<td>9%</td>
</tr>
<tr>
<td>ECG Signal</td>
<td>25</td>
<td>53%</td>
</tr>
</tbody>
</table>

Regarding Blood glucose monitoring, different devices and brands have been exploited for hypo/hyperglycemia and glycemic variability classification and detection, as shown in Table 3 below. Generally, Yellow Spring Instruments is the most used device (50%) followed by Gaurdian Real Time (MinMed, CGM) (30%). GlycoMark is the third most
used device followed by HemoCue Glucose 201 (5%) and Self-Monitored Blood Glucose (SMBG) (5%). Specifically, for hypoglycemia classification and detection, Gaurdian Real Time (MinMed, CGM) (n=6), Yellow Spring Instruments (n=20), HemoCue Glucose 201 (n=2), Dexcom CGM system (n=1), Self-Monitored Blood Glucose (SMBG) (n=2), Medtronic Insulin pump (n=3), and SensorWear armband (physical activity) (n=2). As to hyperglycemia classification and detection, Gaurdian® Real Time (MinMed, CGM) (n=2) and Medtronic Insulin pump (n=2) had been used. In regard to glycemic variability classification and detection, GlycoMark (n=3), Gaurdian Real Time (MinMed, CGM) (n=3), and Medtronic Insulin pump (n=3) had been used.

Table 3: Types of devices used for monitoring of blood glucose levels.

<table>
<thead>
<tr>
<th>BG Devices</th>
<th>Count (n)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guardian Real Time (MinMed, CGM)</td>
<td>12</td>
<td>30%</td>
</tr>
<tr>
<td>HemoCue Glucose 201 (HemoCue)</td>
<td>2</td>
<td>5%</td>
</tr>
<tr>
<td>Yellow Spring Instruments</td>
<td>20</td>
<td>50%</td>
</tr>
<tr>
<td>Dexcom CGM system</td>
<td>1</td>
<td>2.5%</td>
</tr>
<tr>
<td>GlycoMark</td>
<td>3</td>
<td>7.5%</td>
</tr>
<tr>
<td>Self-Monitored Blood Glucose (SMBG)-unknown device</td>
<td>2</td>
<td>5%</td>
</tr>
</tbody>
</table>

Various brands of physiological monitoring (heart rate and ECG signals) devices have been exploited in the reviewed articles. Generally, as shown in the Table 4 below, Compumedics system is the most used system (52%) followed by a customized device like battery powered chest belt worn device (24%). HypoMon is the third most used device (14%) followed by Basis Peak and a self-designed portable apparatus (5%). Specifically, for hypoglycemia classification and detection purpose various devices have been used like HypoMon (n=3), Basis Peak (n=1), Compumedics system (n=10), A battery-powered chest belt worn (n=5), and Self-designed portable apparatus (n=1). Regarding hyperglycemia classification and detection, only one articles has used the Compumedics system (n=1), which indicates that heart rate and ECG signals have a limited use in this case.

Table 4: Types of devices used for monitoring of Physiological parameters (heart rate and ECG signals)

<table>
<thead>
<tr>
<th>HR Devices</th>
<th>Count (n)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>HypoMon</td>
<td>3</td>
<td>14%</td>
</tr>
<tr>
<td>Basis Peak</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Compumedics system</td>
<td>11</td>
<td>52%</td>
</tr>
<tr>
<td>A battery-powered chest belt worn (customized)</td>
<td>5</td>
<td>24%</td>
</tr>
<tr>
<td>Self-designed portable apparatus (customized)</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

Data preprocessing

Data preprocessing is an important stage of any machine learning strategies. In this regard, there are various kinds of data preprocessing strategies used in the reviewed articles. The reviewed articles had relied on both blood glucose and other physiological (heart rate, ECG, skin impedance and others) data, which of course involves different preprocessing strategies depending on the data type under consideration. Regarding the blood glucose data, various preprocessing approach had been used including differencing (derivative) BG values [21, 29], CGM data reconstruction or smoothing using different methods, i.e. spline interpolation [26, 27, 67, 73, 74], a rough feature elimination, i.e. fast SEPCOR algorithm [26, 29], representing BG temporal change information [31], feature selection and feature ranking [65], filtering using Pearson’s Correlation Coefficient (PCC) and the t-test, and the wrapper approach using greedy backward elimination [74]. The other physiological (heart rate, ECG, skin...
impedance and others) parameters had been preprocessed using different methods like normalization [56, 58, 70], feature extraction and selection [51, 53], feature extraction using Fast Fourier Transform (FFT) [49], unsupervised restricted Boltzmann machines (RBM) based feature representation [63], filtering, i.e. using IIR high pass filter [47, 49], correlation analysis [33, 38, 41], and transformation of frequency domain into time domain (FFT) [32].

4.2.2. **Class of machine learning**

4.2.2.1. **Hypoglycemia Classification & Detection**

Different class of machine learning techniques have been adopted in hypoglycemia classification and detection algorithms to predict and detect the incoming hypoglycemia incident in people with type 1 diabetes. Feed Forward Artificial Neural Network (ANN) is the most adopted class of machine learning, which is used in 25% of the studies, as shown in the Figure 5 below. Hybridization of machine learning techniques with other approach like time series, fuzzy logic and others are the second most adopted approach (19%). Support Vector Machine (SVM) ranked the third most adopted class of machine learning (14%). Decision tree ranked the fourth most adopted techniques (6%). Genetic Algorithm, Time delay ANN, Time sensitive ANN, Block based neural network, and Adaptive neural fuzzy inference system (ANFIS) are the fifth most used class of machine learning (5%). Nonlinear Autoregressive Network with exogenous inputs (NARX), and Nonlinear Autoregressive Network (NAR), Gaussian process regression, and Combinational neural logic network ranked as the sixth most used class of machine learning (3%). Radial basis function neural network, variable translation wavelet neural network, and deep belief network ranked the seventh most used class of machine learning (2%). Bayesian neural network (BNN) is the eighth most used class of machine learning (1%).

![Machine Learning Class Diagram](image-url)
4.2.2.2. Hyperglycemia Classification & Detection

Hyperglycemia classification, prediction, and detection has been practiced less as compared to hypoglycemia, which might be linked due to its less severe short term complication as opposed to hypoglycemia incidences. However, irrespective of this limitation, different types of machine learning techniques have been adopted. For example, Recurrent neural network (RNN) is the most used machine learning techniques (23%) as shown in the Figure 6 below. The second most adopted class of machine learning are genetic programming and Hidden Markov Model (HMM) (22%). feed forward ANN, genetic algorithm, and hybrid are the third most adopted machine learning techniques.

![Figure 6: Class of machine learning used in hyperglycemia classification and detection.](image)

4.2.2.3. Glycemic Variability Classification & Detection

Glycemic variability detection is a recent development, which has a great importance in quantifying factors associated with hypo/hyperglycemia incidence. In this regard, there are some research and development involving machine learning techniques. For example, feed forward ANN are the most used class of machine learning (37%), as shown in the Figure 7 below. Naïve Baye classifier, and support vector machine (SVM) are the second most adopted techniques of machine learning (25%). Decision tree is the third most used class of machine learning (13%).

![Figure 7: Class of machine learning used in glycemic variability classification and detection.](image)
4.2.3. Performance Metrics

The performance metrics used in the evaluation of hypoglycemia, hyperglycemia and glycemic variability classification and detection algorithms are depicted in the outer ring, middle ring and inner ring respectively, as shown in the Figure 8 below. According to the outer ring, hypoglycemia classification and detection, sensitivity and specificity are the most used performance metrics (64%), as shown in the Figure 8 below. Accuracy and precision are the second most used performance metrics (14%). Root mean square error and mean square error are the third most used performance metrics (7%). Geometric mean is the fourth most used performance metrics (5%). Correlation coefficient is the fifth most used performance metrics (4%). Time lag (TL), recall, and ROC curve are sixth most used performance metrics (2%). According to the middle ring, hyperglycemia classification and detection, Accuracy and precision are the most used performance metrics (14%), as shown in the Figure 8 below. Root mean square error, mean square error, Time lag (TL), correlation coefficient, recall, false positive rate are the second most used performance metrics (13%). ROC curve, geometric mean, sensitivity and specificity are the third most used performance metrics (7%). According to the inner ring, glycemic variability classification and detection, accuracy and precision are the most used performance metrics (60%), as shown in the Figure 8 below. Sensitivity and specificity are the second most used performance metrics (40%).

Figure 8: Depicts the type of performance metrics used in the studies. The outer ring, middle ring and inner rings depicts the type of performance metrics used in hypoglycemia, hyperglycemia and glycemic variability classification and detection algorithms.
5. Discussion

5.1. Principal findings

The objective of this review is to identify, assess and analyze the state of the art machine applications in blood glucose pattern classifications and anomalies detection; hyperglycemia, hypoglycemia and glycemic variability classification and detection. According to the reviewed literatures, the anomalies classification and detection approach could be roughly categorized as either as a classifier based or model based approach [12, 14]. The classifier-based approach mainly relies on using either a specified threshold or some kinds of rules to classify the BG levels as either normal or abnormal. The difference is that unlike the model-based approach, the classifier-based approach requires a rigorous and deeper knowledge regarding the nature, size and shape of the underlying anomalies under consideration so as to develop the necessary threshold or rule to capture them. However, the model based approach requires only to demarcate the boundary of what is known to be normal so to capture what is believed to be abnormal [14]. The model-based approach doesn’t require a rigorous knowledge of the underlying expected anomalies, i.e., to fully understand and characterize the shape and nature of the expected anomalies [15]. By simply defining what is the expected normal pattern the system should exhibit, the model-based approach is capable of detecting the abnormal behavior, which is not considered as normal behavior of the system. Defining and discovering what is “normal” is a challenging task especially for a dynamic and complex systems, e.g., BG dynamics. However, this is often tackled in a dynamic and complex system by relying on either a machine learning model trained on a large and enough datasets or using an explicit mathematical model of the system if it exists already [14].

Various class of machine learning algorithms have been adopted for the task. Regarding hypoglycemia classification and detection, feed forward artificial neural network, hybrid systems, support vector machine (SVM), decision tree, genetic Algorithm, adaptive neural fuzzy inference system (ANFIS), nonlinear autoregressive network with exogenous inputs (NARX), nonlinear autoregressive network (NAR), Gaussian process regression, deep belief network, Bayesian neural network (BNN) have been developed and tested. These techniques have explored various kinds of input parameters notably blood glucose (BG), heart rate, QT interval, insulin, diet, physical activity, galvanic response, and skin impedance. Concerning hyperglycemia classification and detection, recurrent neural network (RNN), genetic programming, Hidden Markov Model (HMM), feed forward artificial neural network, genetic algorithm, and hybrid systems have been developed and tested exploring various types of input parameters including blood glucose, insulin, heart rate, and QT interval. Glycemic variability detection is a recent development, which has a great importance in quantifying factors associated with hypo/hyperglycemia incidence. In this regard, there are some research and development involving machine learning techniques. For example, feed forward artificial neural network, Naïve Baye classifier, support vector machine (SVM), and decision tress have been tested up to the task using blood glucose and insulin delivery profiles. Moreover, all those studies have relied on either indirect indicators variable such as heart rate, QT interval and others or subset of input parameters that affects blood glucose dynamics. The patient’s contextual information, e.g. meals, physical activity, insulin and sleep, have a significant effect on blood glucose dynamics and a proper anomalies classification and detection algorithm should
consider the effects of these parameters. In this regard, the individual patient is expected to record meal, insulin and physical activity data. One of the main limitation is the meal modelling, where most of the algorithm depends on the individual estimation of carbohydrate, which is prone to errors and further aggravates the degradation to detection performance. Regarding physical activity, there are various wearables and sensors that can record the individual’s physical activity load and durations. However, there is limitation on the way these signals are employed in the detection algorithms. For example, there are some studies which considers descriptive features by summarizing the number, intensity, steps, exercise durations and others to better quantify the effect of physical activities. Moreover, recording insulin dosage has its inherent limitations, which might affect the detection performance. For example, blockage of insulin flow from the insulin pump due to the infusion set failure and error incurred during manual registrations might pose significant challenge in the performance of the detection system. Furthermore, CGM is becoming one of the most important components in these classification and detection algorithms. However, even if CGM advancement has enabled patients to have a continuous estimation of their subcutaneous glucose levels, it has limitations when used in a personalized detection system (an alarm). In this regard, recent studies have showed that autocorrelation of the CGM reading vanishes after 30 minutes making the detection performance to degrade afterwards. These findings suggest that any detection algorithms aiming for a better lead time should consider other patient’s contextual information and various feature of the CGM itself. There are some studies, which develop a model by assessing several features of the CGM signal so as to compensate for its inaccuracy. Moreover, CGM is found to be inaccurate during hypoglycemia episodes, i.e., insulin-induced hypoglycemic vs. spontaneous hypoglycemia. In this regard, insulin-induced hypoglycemia is found to be difficult to detect as compared to the spontaneous hypoglycemia. Fast occurring hypoglycemia is difficult to detect due to the blood-interstitial delay, which makes them important features to be detected by a given model. Furthermore, CGM calibration frequency and timing also affects the performance of the detection algorithm.

The reviewed studies are limited to and could be roughly be categorized by age groups (children, young, adult and old), time of the day (diurnal vs nocturnal) and, configurations (online vs offline). For example, most of the studies considers nocturnal hypoglycemia detection considering the fact that most of hypoglycemia crises occurred during night time and also the crises during this time has a bad consequence as compared to the diurnal period. Moreover, it is a fact that nocturnal detection is simpler as compared to the diurnal considering the dynamics of the patients. However, irrespective of these challenge, there are also studies, which considers the diurnal period. However, there are limited study that attempts to develop an algorithm that could detect anomalies in both contexts. Most surveys reported that age group has a great effect on blood glucose dynamics, which is typically related with the dynamics and active lifestyle adopted by each group. Therefore, it deems a necessary approach to consider a personalized algorithm for each age group. Regarding the configuration, there are fewer attempt of online algorithm, where almost all of the algorithms were tested and implemented as offline mode. In this regard, the most crucial issues concerning machine learning strategies could be the necessity of frequent retraining when subject to a real time and dynamic task. In addition, the most important component in classification and detection algorithms is the threshold used to differentiate the normal from the abnormal. In this regard, almost most of the studies have used a static threshold based on suggestion either form literatures or physicians and other concerned bodies like the American
diabetes association. However, the critical issues in this approach is that it might vary from patient to patient and also some patients might not feel any symptoms at the specified threshold (when using indirect indicators such as heart rate, QT interval and others). However, there are some studies, which employed a fuzzy logic-based approach by having a continuous decision space.

In principle, any future blood glucose anomalies classification and detection algorithm should be expected to detect any upcoming anomalies as soon as possible (lead time - giving more response time), avoid any false alarm at any cost, perform in real time (in an online fashion), adapt with the dynamics of blood glucose evolution (learn continuously), automatically tune its parameters without user intervention, be able to perform throughout the day in a free living condition (diurnal and nocturnal periods), incorporate as many input variables to better capture the dynamics. In this regard, for example, the most crucial issues concerning a real time (online) machine learning algorithm could be the necessity of frequent retraining when subject to a real time and dynamic task. Moreover, developing a model that considers a real time and adaption to free living condition needs to incorporate a wide range of parameters that affects blood glucose dynamics. Furthermore, it should properly consider and address the inherent technological limitation that affects the performance of the detection algorithm. Almost all of the studies need a proper clinical validation to be integrated into a smartphone and CGM for a real time application. This can be better described by looking at the number of sample used and its validation strategies (See Appendix 1). Therefore, future studies should give more emphasis on clinical validation by taking sufficient number of subjects in the development and testing phase so as to better quantify the intra and intra variability among patients. In addition, the most crucial concept of justifying and reporting the underlying cause, as either due to patient controllable or patient uncontrollable parameters, for the detected anomalies is not addressed in any of the reviewed literatures. For example, the underlying cause of hyperglycemia incidences could be patient controllable parameters such as diet or patient uncontrollable parameters such as stress and infections. Therefore, in this regard, a proper hyperglycemia classification and detection system might be expected to be able to identify and report the underlying cause, which has a greater significance to the patient especially during infections crises.

5.2. Summary of Existing Efforts - Machine Learning Techniques

5.2.1. Artificial Neural Network

Artificial neural network is computational model consisted of a set of interconnected neurons, and a scaled connection between them called weights [75]. Based on network topology, ANN is mainly categorized as a feed forward artificial network (SLP, MLP, and Radial Basis function) and a recurrent neural network (Elman net, Kohonen’s SOM, and Hopfield Networks) [75]. There are various types of ANN used in solving blood glucose classification and detection tasks; hypoglycemia, hyperglycemia, glycemic variability classification and detection. Regarding hypoglycemia classification and detection, for instance, Eljil et al. [22], had proposed a special type of Artificial Neural Networks (ANN) known as Time- Sensitive ANN (TS-ANN) and compared the result with a Time Delay Neural Network (TDNN), Nonlinear Autoregressive Network with exogenous inputs (NARX), Distributed Time Delay Neural Network (DTDNN), and Nonlinear Autoregressive Network (NAR). San et al. [58], proposed an evolvable block based neural network (BBNN) and compared the result with feedforward neural network (FWNN) and multiple regression (MR). Moreover, San et al. [63], proposed a deep belief network (DBN) and compared the
result with a wavelet neural network (WNN), feedforward neural network (FFNN), and multiple regression (MR) models. Some of the studies has investigated the advantage of having a separate feature extraction and classification unit. In this regard, for example, both Laione et al. [32] & Nguyen et al. [49], have proposed an artificial neural network (ANN) using fast Fourier transform (FFT) for data extraction. Nguyen et al. [49], has further trained the network through a two-step process that combines the advantage of genetic algorithm and Levenberg Marquardt algorithm. Chan et al. & Yan et al. [19, 20], also proposed a neural network-based rule discovery system consisted of a neural network-based classification unit and a rule-based extraction unit. There are some studies, which optimized the ANN parameters through a particle swarm optimization technique. For example, Ling et al. [37], Phyo et al. [55, 56], & San et al. [61], proposed a new hybrid rough neural network, a variable translation wavelet neural network (VTWNN), a normalized radial basis function neural network (NRBFNN), and a combinational neural logic network (NLN) with the neural-Logic-AND, -OR and -NOT gates respectively, where the design parameters of the network were optimized through a hybrid particle swarm optimization with wavelet mutation (HPSOWM) operation. Moreover, Nguyen et al. [47, 48] also proposed an artificial neural network, which is optimized through a standard particle swarm optimization strategy. Furthermore, some studies have investigated extreme learning machines. For instance, Ling et al. [39] & San et al. [64], proposed a feed forwarded neural network trained through extreme learning machine (FFNN-ELM) and compared the result with, feed forward neural network optimized through particle swarm optimization (FFNN-PSO), multiple regression based fuzzy inference system (MR-FIS), fuzzy inference system (FIS) and Linear multiple regression (LMR). Mo et al. [42], has also used an extreme learning machines (ELM) and regularized ELM (RELM) on CGM data. In addition, Nguyen et al. [44-46], had proposed an optimal Bayesian neural network algorithm using a feed forward neural network architecture. Concerning hyperglycemia classification and detection, there is one study Nguyen et al. [70], which uses a feed forward multi-layer neural network using different training algorithm, i.e. gradient descent, gradient descent with momentum, scaled conjugate gradient, and resilient back propagation. Regarding glycemic variability classification and detection, the reviewed studies had been performed for either detection purpose or for an automated metrics purposes. For detection purpose, for example, Wiley et al. [74], proposed a Naive Bayes (NB), Multilayer Perceptron (MP), and Support Vector Machine (SVM) models to detect excessive glycemic variability on a CGM data and compared the accuracy of the result with other two diabetes experts. Regarding the automated metrics, Marling et al. [72], had developed naive Bayes classifier (probabilistic reasoning), a multilayer perceptron (ANN), and a logistic model tree (decision tree built using logistic regression), which could be used to monitor continuous glucose monitoring (CGM) data. Moreover, Marling et al. [73], also proposed multilayer perceptron (MPs) and support vector regression (SVR) to develop a consensus perceived glycemic variability (CPGV) metric.

5.2.2. Support vector machines (SVM), Kernel Function (KF) & Gaussian process regression

Support vector machine has been widely exploited across a wide range of applications such as pattern identification and recognition, categorization or classification, regression and prediction [76]. Support vector machine, kernel function and Gaussian process regression have been exploited for hypoglycemia classification and detection purposes in the reviewed literatures. For example, Georga et al. [23], developed a support vector regression (SVR)
for hypoglycemia prediction and compared the performance with feed-forward multilayer perceptron (MLP) and Gaussian processes (GP) regression. Georga et al. [24], also proposed support vector regression (SVR) and Gaussian processes (GP) regression for prediction of BG as indication of a daily hyper/hypoglycemia incidences to the patients as well as provision of support to physicians in decision making about treatment and risk of complications. Moreover, Jensen et al. [26, 27], developed a retrospective automatic pattern recognition system so as to detect hypoglycemia incidences using CGM and to foster a thorough evaluation of past events and discussion with their caregivers. Jensen et al. [28, 29], also developed a real-time pattern classification model by using several features from the CGM data, which is able to detect hypoglycemia in a real time. Furthermore, Marling et al. [40], proposed a hypoglycemia detection algorithm that incorporate noninvasive data sensor data from fitness bands and also compared different kernel for the task; linear, Gaussian and quadratic kernels. Nuryani et al. [50], also proposed a swarm-based support vector machine (SVM) algorithm with inputs of the repolarization variabilities so as to detect hypoglycemia incidences.

5.2.3. Genetic Programming and Genetic algorithm

An evolutionary algorithm (EA) is a biologically inspired approach to problem solving [77]. The two most used variants of EA are Genetic programming (GP) and Genetic algorithm (GA). In this regard, there is little visibility of using genetic programming and genetic algorithm in their non-hybrid form. However, there are some studies that uses these techniques in their hybrid form. For example, Ling et al. [33-35], developed a hypoglycemia detection algorithm using a genetic algorithm (GA)-based multiple regression with fuzzy inference system (FIS). The study exploited GA so as to optimize the fuzzy rules, membership function of FIS and also the model parameters of regression method.

5.2.4. Random Forest (RF)

Random forest or random decision forests are an ensemble approach of learning for classification and regression applications, which learns by constructing a multitude of decision tress generating the mode of the class or mean of prediction. In this regard, there are some studies, which uses decision tress in the context of hypoglycemia classification and detection tasks. For example, Eljil et al. [21], proposed a decision trees using different techniques, namely, C4.5, J4.8, REPTree, Bagging and the cost sensitive version of J4.8. Jung et al. [30], also proposed a decision trees using a new predictor variables using continuous glucose monitoring (CGM) data. Moreover, Jung et al. [31], proposed a decision tree and support vector machine based prediction model using self-monitored blood glucose. Zhang et al. [65], also developed a new method using classification tree-to predict the occurrences of acute hypoglycemia during intravenous (IV) insulin infusion before the actual hypoglycemic events take place.

5.2.5. Hidden Markov Model (HMM)

Hidden Markov model (HMM) is a variant of the statistical Markov model, where the system being modeled is assumed to follow a Markov property with unobserved states [69]. There are some studies, which uses HMM to develop a model-based blood glucose anomalies classification and detection. For example, Zhu et al. [71] & [69]
studied an approach for automatic detection of anomalies in an individual BG data, where the daily normal measurements are used for training the model. The method tries to analyze the incoming BG data with the trained Markovian world based on the individual’s historical BG data.

5.2.6. Hybrid and Ensemble Models

Hybridization is the process of combining two or more different approaches either in parallel or in serious connection, either at preprocessing stage, feature extraction or learning stage, when looking for an improved performance. In this regard, there are some attempts in the reviewed articles, which tries different approaches for an enhanced performance in hypoglycemia classification and detection. For example, hybridization of neural network with other techniques are demonstrated in some of these studies. Chan et al. [18] developed a hybrid system consisted of neural networks and genetic algorithm and also compared the performance with MLP network and classical statistical algorithms. Ghevondian et al. [25], proposed a novel hybrid system of a fuzzy neural network (FNNE) estimator to predict the glycaemia profile and hypoglycemia incidences. San et al. [60], proposed a hybrid system using an adaptive neural fuzzy inference system (ANFIS) and also compared the performance with wavelet neural network (WNN), feedforward neural network (FWNN) and multiple regression (MR). There is also some approach to hybridize support vector machine. Nuryani et al. [51], proposes a hybrid Fuzzy Support Vector Machine (FSVM) and investigated the applicability of three kernel function radial basis function (RBF), exponential radial basis function (ERBF) and polynomial function for the task. Moreover, Nuryani et al. [53, 54], proposed a novel strategy using a hybrid particle swarm - based fuzzy support vector machine (SFisSvm) technique. Fuzzy reasoning models are also tested in some of the studies. For example, Ling et al. [36], developed a hybrid particle-swarm-optimization-based fuzzy-reasoning model, where the fuzzy rules and the fuzzy-membership functions are optimized through a hybrid particle-swarm-optimization with wavelet mutation. The FRM are also compared with Feed-Forward Neural Network (FFNN) and Multiple-Regression (MR). Mathews et al. [41] developed a hybrid model using a fuzzy inference system with multiple regression, where the fuzzy rules are optimized through genetic algorithm. The study also compares the performance of the developed system (FIS + Genetic Algorithm) with a neural network whose parameters are optimized through particle swarm optimization (NN+PSO). In addition, San et al. [62], proposed a hybrid system based on rough sets concepts and neural computing. The study has compared various hybrid approaches trained through hybrid particle swarm optimization with wavelet mutation including Rough-Block-Based Neural Network (R-BBNN), BBNN, rough feedforward neural network (R-FWNN), wavelet neural network (WNN)], SVM with a radial basis function and conventional feedforward neural network (FWNN). Ling et al. [38], also proposed an alarm system based on hybrid neural logic network with multiple regression.

Due to the complexity of blood glucose dynamics, it remains difficult to achieve an accurate result in every circumstances. One model can have a better accuracy some circumstances and the other model can achieve better accuracy where the first model fails to achieve comparable result. Therefore, it is natural to look for possibilities to exploit the strengths from these different models to achieve a better accuracy in most of the circumstances, which lead to ensemble approaches. Ensemble approach is generally favored when one is interested to merge two or more different models for an improved performance. In this regard, there are some studies, which tries to combine two different models looking for performance improvement in the overall system. In this regard, Daskalaki et al. [68],
proposed an early warning system (EWS), for both hyperglycemia and hypoglycemia, using recurrent neural network (RNN) and autoregressive with an output correction module (cARX) models. Moreover, the study investigated the performance improvement from the combined use of both recurrent neural network (RNN) and autoregressive with an output correction module (cARX). Moreover, Botwey et al. [67], proposed combining an autoregressive model with output correction – cARX, and a recurrent neural network – RNN based on different data fusion schemes including Dempster-Shafer Evidential Theory (DST), Genetic Algorithms (GA), and Genetic Programming (GP).

**Conclusion**

Despite the complexity of blood glucose dynamics, there are many attempts to capture hypoglycemia, hyperglycemia incidences and the extent of an individual glycemic variability using different approaches. Recently, due to the ubiquitous nature of self-management mHealth apps, sensors and wearables has paved the way for the continuous accumulation of self-collected health data, which in turn contributed for the widespread research of machine learning applications in these tasks. In the reviewed articles, generally, the anomalies classification and detection approaches could be categorized as either process based or rule-based approaches. Hypoglycemia classification and detection has been given more attention than hyperglycemia and glycemic variability detection, which might be due to its serious complication and the comparable complexity involved. The state of the art indicates that various class of machine learning have been developed and tested in these tasks. Regarding hypoglycemia classification and detection, Feed Forward Artificial Neural Network, Hybrid systems, Support Vector Machine (SVM), Decision tree, Genetic Algorithm, Adaptive neural fuzzy inference system (ANFIS), Nonlinear Autoregressive Network with exogenous inputs (NARX), and Nonlinear Autoregressive Network (NAR), Gaussian process regression, Deep belief network, Bayesian neural network (BNN) have been developed and tested. These techniques have explored various kinds of input parameters notably blood glucose (BG), heart rate, QT interval, insulin, diet, physical activity, galvanic response, and skin impedance. Concerning hyperglycemia classification and detection, Recurrent neural network (RNN), Genetic programming, Hidden Markov Model (HMM), Feed forward ANN, Genetic algorithm, and hybrid systems have been developed and tested exploring various types of input parameters including blood glucose, insulin, heart rate, and QT interval. Glycemic variability detection is a recent development, which has a great importance in quantifying factors associated with hypo/hyperglycemia incidence. In this regard, there are some research and development involving machine learning techniques. For example, Feed forward ANN, Naïve Baye classifier, and Support vector machine (SVM). Most of these studies have used a theoretical threshold either suggested by literatures or physician and various concerned bodies like the American diabetes association. However, the problem here is that some patient might feel no symptoms at the specified threshold and it may vary from patient to patient. Therefore, a model should consider such difference among the patients (intra and inter variability) and also track its temporal change overtime. Moreover, the studies should give more emphasis on the time lag (TL) and the various types of inputs used. Furthermore, researchers should give proper emphasis to develop anomalies classification and detection models, which is capable of justifying and reporting the underlying cause, as either due to patient controllable or patient uncontrollable parameters. Generally,
we foresee these developments might encourage researchers to further develop and test these systems on a large-scale basis.

**Acronyms**

- Decision Tree (DT)
- Support vector machine (SVM)
- Recurrent neural network (RNN)
- Multi-layer perceptron (MLP)
- Non-linear Autoregressive with and without Exogenous inputs (NARX & NAR)
- Adaptive neural fuzzy inference system (ANFIS)
- Naïve Bayes classifier (NBC)
- Rule Based (RB)
- Extreme learning machine (ELM)
- Hidden Markov model (HMM)
- Deep belief network (DBN)
- Gaussian Process (GAP)
- Evolvable block based neural network (BBNN)
- Variable translation wavelet neural network (VTWNN)
- Bayesian Neural Network (BNN)
- Combinational neural logic network -CNLN
- Genetic Programming (GP)
- Genetic algorithm (GA)
- Blood Glucose (BG)
- Heart Rate (HR)

**References**


