Artificial Intelligence Methodologies and Their Application to Diabetes

Mercedes Rigla, MD, PhD, Gema García-Sáez, PhD, Belén Pons, MsC, and Maria Elena Hernando, PhD

Abstract
In the past decade diabetes management has been transformed by the addition of continuous glucose monitoring and insulin pump data. More recently, a wide variety of functions and physiologic variables, such as heart rate, hours of sleep, number of steps walked and movement, have been available through wristbands or watches. New data, hydration, geolocation, and barometric pressure, among others, will be incorporated in the future. All these parameters, when analyzed, can be helpful for patients and doctors’ decision support. Similar new scenarios have appeared in most medical fields, in such a way that in recent years, there has been an increased interest in the development and application of the methods of artificial intelligence (AI) to decision support and knowledge acquisition. Multidisciplinary research teams integrated by computer engineers and doctors are more and more frequent, mirroring the need of cooperation in this new topic. AI, as a science, can be defined as the ability to make computers do things that would require intelligence if done by humans. Increasingly, diabetes-related journals have been incorporating publications focused on AI tools applied to diabetes. In summary, diabetes management scenarios have suffered a deep transformation that forces diabetologists to incorporate skills from new areas. This recently needed knowledge includes AI tools, which have become part of the diabetes health care. The aim of this article is to explain in an easy and plane way the most used AI methodologies to promote the implication of health care providers—doctors and nurses—in this field.

Keywords
artificial intelligence, decision support, diabetes, machine learning

Artificial intelligence (AI) has been defined in many ways. Currently, the most accepted definition is the one made by Boden: the ability to make computers do things that would require intelligence if done by humans. It is also not trivial to define intelligence. Intelligence is usually defined as a group of abilities such as understanding, learning and reasoning to make decisions and to solve problems. AI emulates these aspects of human intelligence by means of a number of tools. The aim of this review is to list and explain the most frequently used AI tools in simple words to facilitate understanding. AI methodologies and techniques have been applied to medicine and health in general over the past decades. Diagnosis, classification, therapy and robotics, among others, are common AI medical applications. Among the variety of different AI technologies, neural networks and fuzzy logic (FL) are the most often used ones to date. However, there are other techniques and methodologies, which have been also selected and included in this review due to their relevance. In addition, a glossary of useful terms has been included (Table 1) as well as a list of examples of the most representative publications on AI applied to diabetes (appendix).

AI Methodologies

Expert Systems in Medicine
Expert systems (ES) correspond to the most common type of AI system in routine clinical use. They are defined as systems with the ability to capture expert knowledge, facts and reasoning techniques to help care providers in routine work. ES attempt to mimic clinician’s expertise by applying inference methods to help in decision support or problem solving. ES have the ability to manage data to come up with reasoned
Computer interpretable guidelines (CIGs) are clinical practice guidelines formalized in a computer-based system.

Computational process to extract information and knowledge from a large dataset and to transform it into an understandable structure.

In FL, defuzzification is the process of converting a combined output of fuzzy rules into a numerical values. The input for the defuzzification process is the aggregated set and the output is a single number.

In FL, fuzzification is the process of mapping numerical inputs into fuzzy inputs: degree to which the inputs belong to the respective fuzzy sets according to a membership function.

Exploratory algorithms that shorten the time to find a reasonably good solution that would otherwise be excessively time-consuming.

Systems which integrate a combination of AI techniques; for example, neuro-fuzzy, fuzzy-expert systems, etc.

Key feature of an expert system in charge of the reasoning process whereby the expert system reaches a solution based on the expert's knowledge contained in the rule base and the facts contained in the database.

In FL, a membership function is required to convert input parameters to a fuzzy set. These membership functions can have different shapes. The most common are triangular shape; however bell, trapezoidal, sinusoidal, and exponential can be also used.

Algorithmic framework that provides a set of strategies to develop heuristic algorithms. GA are examples.

ANN structure for supervised learning. Neurons in the perceptron are classifiers that aggregate inputs and assign a binary value (either 0 or 1).

Technique for classification and prediction based on the Bayes theorem. The algorithm constructs models that estimate the posterior probability of each class, given a set of input attributes. Bayes's theorem, which allows calculating the probability of A given B, from knowing the probability of observing event B given that A is true, and the probabilities of A and B. The term "Naive" refers to the assumption that given a class, all the features or attributes are conditionally independent of each other. That premise simplify very much the calculations.

Describes the objects, concepts and their relationships in a domain of knowledge.

A stochastic program operates using probabilistic methods to solve problems.

Mathematical algorithm able to learn from a dataset where the desired output is already known. It generalizes a function that maps the available inputs to their corresponding desired output. Classification and prediction algorithms are supervised learning algorithms.

Algorithm devoted to discover relationships or structures in a dataset. The desired output is unknown. Examples of unsupervised algorithms are clustering and association.

Rule-based reasoning (RBR), case-based reasoning (CBR), and fuzzy systems are the most common ES used in the diabetes domain.

RBR. RBR is based on the transfer of knowledge from an expert to a computer. As a consequence, the computer has to be able to find solutions to problems that otherwise should be solved by an expert. Knowledge is represented in statements in the form “if-then,” in such a way that the line of reasoning can be explained. The process of knowledge acquisition starts with a number of interviews between the expert and the knowledge engineer who will end up building and testing the ES. During these interviews, the domain expert establishes all the possible options and the engineer encodes this knowledge to become “computer interpretable.”

CBR. CBR finds solutions to new problems by adapting previously good solutions to similar problems. Case studies features need to be specified to be helpful in retrieving other cases. At the same time, features have to be discriminative enough to avoid the retrieval of cases studies which could lead to wrong solutions because of being too different. Unlike RBR, CBR does not require an explicit domain
model, but just to identify new cases with significant features, which is in fact the way CBR "learns."

CBR procedures are usually explained as the so called "CBR working cycle", which includes five steps: (1) current problem description; (2) search for a successful solution of a similar case; (3) adaptation and reuse of the solution to the new problem; (4) evaluation; and (5) confirmed solution storage. The main CBR limitations are related to the need to get huge case studies databases, which could include nonrelevant information and make the retrieval often excessively time-consuming.

**FL.** Fuzzy ES are used for representing, in a computer-understandable way, expert knowledge that uses ambiguous terms. Thinking in terms of conventional logic, a blood glucose range >180 mg/dl is high and a range <80 mg/dl is low. This classification is not particularly useful for making decisions. In real life a blood glucose value of 181 mg/dl in most cases deserves a different action with respect to 281 mg/dl. In other words—in fuzzy words—181 mg/dl is high but almost acceptable while 281 mg/dl is very high and far from being acceptable. FL expresses this ambiguity assigning a certain degree of membership to different categories. In our previous example, we could say that 181 mg/dl pertains 70% to the category of "high" but only 30% to the category of "very high."

**Machine Learning**

Machine learning (ML) algorithms are characterized by the ability to learn over time without being explicitly programmed. The main features of ML are problem solving usually based on a classification of data. There has been a gradual switch from heuristic approaches toward ML techniques. In the field of data mining, ML algorithms are being used to discover valuable knowledge from large databases such as in electronic medical records, which might include implicit regularities. Also ML can be applied to domains where a computer program needs to dynamically adapt to changing conditions. For example ML algorithms are useful to learn from each patient monitoring data and adapt along time in an artificial pancreas system.

ML draws on results from AI, probability and statistics, computational complexity theory, control theory, information theory, philosophy, psychology, neurobiology, and so on.³

Methods in ML include decision trees (DT), artificial neural networks (ANN), genetic algorithms (GA), or support vector machines (SVM). All of them have been successfully applied in the field of diabetes.

**ANN.** ANN are based on the human brain function, that means, interconnected neurons. Each neuron, the simple unit, receives several inputs and generates only one output. Each connection has assigned a weight related to the importance of the output. The neural network "learns” by training with known inputs, comparing actual output with the known one and using the error to adjust weights. Thus, the links which produce right answers are strengthened and those which generate wrong answers, weakened.

When using a library of existing neural networks, the most common is that in the training process we obtain information about how the algorithm works in the form of a mean square error (MSE). For each example, the ANN evaluates the error in all its output neurons, raises each of those numbers squared, and finally the average is calculated. Using MSE, errors are always positive and the errors of some neurons do not nullify those of others.

**Deep learning.** Deep learning is a new branch of ML based on neuron behavior inside of human brains. It can be considered as a so called “evolutionary computation” and were defined by John Holland almost 50 years ago.⁴ GA simulate natural selection by creating a population of individuals (solutions) for optimization problems. The new solutions are obtained from operating “genetically” the initial population. The chromosome (set of “genes”) is represented as a string of 0s and 1s.

Once an initial population of chromosomes is generated, the first step is just to calculate the fitness of each chromosome. The fitness function value quantifies the optimality of a solution ranking it against the other solutions. If the solution created is not optimal, then a pair of chromosomes is selected for exchanging parts (crossover) and creates two offspring chromosomes. In the next step, a mutation randomly changes at least one gene in the chromosomes. The initial population is replaced with the new population and a new iteration starts. GA iterations end when one of the termination criteria (usually a predefined number of iterations) is satisfied. In the end, the more fit chromosomes survive.

**DT.** DT constitute a graphical representation of a dataset that describes the data by tree-like structures, which provides a very intuitive way of representing and understanding rules. A decision tree is composed of nodes, branches and leaves. A node represents a decision while a leaf represents an outcome. The DT always starts from the root node and grows
down by splitting the data at each level into new nodes. DT are particularly good at solving classification problems.

DT are most often created based on a learning algorithm able to extract the knowledge accumulated in a specific dataset. After the DT structure has been defined, the knowledge can also be represented as sets of if-then rules to improve human readability. Some of the most widely used algorithms are ID3 and C4.5.6

DT have been successfully applied in diabetes to a broad range of tasks such as screening in type 2 diabetes and blood glucose classification.8

SVM. SVM are currently one of the most popular, flexible and powerful ML algorithms used for classification.

SVM are maximum-distance classification algorithms. They define an hyperplane to separate two classes above and below it, providing the maximal distance between the classifying plane and the closest data points. The points that are closest to the border are called “support vectors.” In its most basic formulation, SVM can only work with binary classification problems but, with a relatively simple extension, they can also solve multiclass classification tasks.

In the field of diabetes, SVM have been used to predict prediabetes and diabetes disease9 and in diabetes diagnosis.10

AI Technologies and Diabetes: Areas of Application

All the AI technologies explained in section 1 have been applied to different areas of diabetes management (see the appendix). In this section some examples of these applications are going to be summarized to improve the overall understanding of their utility.

Decision Support for Patients Using CBR

One of the most relevant experiences on the application of ES to decision support for patients has been performed by researchers from the Imperial College in London. They have developed and tested a bolus calculator algorithm based on CBR. This system uses continuous glucose monitoring data and is implemented in the patients smartphone. A pilot feasibility study has been published11 showing the potential benefits of this tool over conventional bolus calculators.

Closed-Loop Systems Based on FL

Apart from proportional derivative integral (PID) and model predictive control (MPC), FL-based algorithms have been successfully used for closed-loop studies, even in the ambulatory setting.12

One of the first publications including FL for closed-loop system was done by Mauseth et al in 2010. The controller used as inputs BG and the rate of change of glucose.13 Using a matrix the system assigned a coefficient which after defuzzification proposed insulin microbolus. Three years later the system was tested in an pilot study with good results.14

Computer Interpretable Guidelines (CIGs) Applied to Gestational Diabetes Management

Clinical practice guidelines are worthy instruments for quality of care improvement. Through formalization as CIGs using a complex RBR system, decision-support tools can be developed.

Clinical experience with gestational diabetes CIGs used for patients and doctors decision support is shown in another article included in this special section of the journal.15 In brief, a pilot study shown a high degree of patients’ satisfaction and higher compliance with blood glucose monitoring in comparison with usual care based on face-to-face visits.

Retinopathy Detection Using ANN

Recently, deep learning ANN has shown to identify diabetic retinopathy or diabetic macular edema in retinal fundus images with high sensitivity and high specificity.16 The authors have developed an algorithm that computes diabetic retinopathy severity from the intensity of the pixels in a fundus picture. The function was trained with a large set of images and then evaluated at one operating point selected for high specificity and a second operating point for high sensitivity obtaining very high scores.

Conclusion

Diabetology needs to suffer an adaptation process to incorporate new tools for diabetes management. Technology and particularly sensors and computer applications have become a key instrument in diabetes management for health care providers and patients. Although modern diabetes care units should include a diabetes technologist17 for dealing with technology, doctors and nurses cannot ignore the basics to better find solutions to each patient circumstances. Knowledge on insulin pumps and more recently on glucose sensors has been increasing progressively; however, comprehension about AI and smart applications performance remains largely inadequate. This article provides a general overview of the elementary concepts, definitions, and terminology frequently used in AI-related applications as well as a list of relevant publications of AI applied to diabetes.
Appendix

Examples of the Most Representative Publications on AI Applied to Diabetes.

<table>
<thead>
<tr>
<th>Method</th>
<th>Application</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBR</td>
<td>Decision support</td>
<td>Artif Intell Med&lt;sup&gt;16&lt;/sup&gt;</td>
</tr>
<tr>
<td>+FL</td>
<td>Automated control</td>
<td>Diabetes Technol Ther&lt;sup&gt;17&lt;/sup&gt;</td>
</tr>
<tr>
<td>CBR</td>
<td>Bolus calculator</td>
<td>Diabetes Technol Ther&lt;sup&gt;11&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Insulin dose recommendation</td>
<td>J Biomed Inform&lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Type 2 DM treatment suggestions</td>
<td>Comput Methods Programs Biomed&lt;sup&gt;21&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>BG pattern detection in pump users</td>
<td>J Diabetes Sci Technol&lt;sup&gt;22&lt;/sup&gt;</td>
</tr>
<tr>
<td>+FL</td>
<td>Risk of complications</td>
<td>J Diabetes Sci Technol&lt;sup&gt;23&lt;/sup&gt;</td>
</tr>
<tr>
<td>FL</td>
<td>DM diagnosis</td>
<td>Methods Inf Med&lt;sup&gt;24&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Automated control</td>
<td>Artif Intell Med&lt;sup&gt;25&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Peripheral neuropathy assessment</td>
<td>Diabetes Obes Metab&lt;sup&gt;26&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Albuminurinia screening</td>
<td>J Diabetes Sci Technol&lt;sup&gt;27&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Diabetes diagnosis</td>
<td>Diabetes Technol Ther&lt;sup&gt;27&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Hypoglycemia detection</td>
<td>N Engl J Med&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Decision support</td>
<td>Diabetes Technol Ther&lt;sup&gt;14&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Blood glucose classification</td>
<td>J Med Eng Technol&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>+ANN</td>
<td>Renal failure prediction</td>
<td>Gait Posture&lt;sup&gt;29&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Retinopathy assessment</td>
<td>Comput Methods Programs Biomed&lt;sup&gt;30&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Glucose prediction</td>
<td>Australas Phys Eng Sci Med&lt;sup&gt;31&lt;/sup&gt;</td>
</tr>
<tr>
<td>ANN</td>
<td>Hypoglycemia detection</td>
<td>Artif Intell Med&lt;sup&gt;32&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Liver cancer prediction in type 2 DM</td>
<td>IEEE Trans Syst Man Cybern B Cybern&lt;sup&gt;33&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>GFR prediction</td>
<td>IEEE Trans Biomed Eng&lt;sup&gt;34&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>PreDM/DM screening</td>
<td>IEEE Trans Biomed Eng&lt;sup&gt;35&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Prediction of DM regression after surgery</td>
<td>AIME&lt;sup&gt;36&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Retinopathy detection</td>
<td>Comput Math Methods Med&lt;sup&gt;37&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Foot ulcers risk</td>
<td>Med Biol Eng Comput&lt;sup&gt;38&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Glucose prediction</td>
<td>Med Biol Eng Comput&lt;sup&gt;39&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Bone mineral density prediction in type 1 DM</td>
<td>Conf Proc IEEE Eng Med Biol Soc&lt;sup&gt;40&lt;/sup&gt;</td>
</tr>
<tr>
<td>GA</td>
<td>Diabetic retinopathy detection</td>
<td>Compu Math Methods Biomed&lt;sup&gt;41&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Estimation of model parameters</td>
<td>J Med Eng Technol&lt;sup&gt;42&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Australas Phys Eng Sci Med&lt;sup&gt;43&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diabetes Res Clin Pract&lt;sup&gt;44&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diabetes Technol Ther&lt;sup&gt;45&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Obes Res Clin Pract&lt;sup&gt;46&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Appendix. (continued)

<table>
<thead>
<tr>
<th>Method</th>
<th>Application</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT</td>
<td>Prediction of macrosomia and gestational DM</td>
<td>Biomark Med^6,63</td>
</tr>
<tr>
<td></td>
<td>Foot ulcer prediction</td>
<td>J Biomed Opt^41</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular risk</td>
<td>Diab Vasc Dis Res^55</td>
</tr>
<tr>
<td></td>
<td>Retinopathy assessment</td>
<td>IEEE Trans Inf Technol Biomed^66</td>
</tr>
<tr>
<td></td>
<td>Peripheral neuropathy prediction</td>
<td>Biomed Eng Online^66</td>
</tr>
<tr>
<td></td>
<td>Type 2 DM screening</td>
<td>Artif Intell Med^69</td>
</tr>
<tr>
<td>SVM</td>
<td>Prediction of prediabetes and diabetes</td>
<td>Chin Med J (Engl)^70</td>
</tr>
<tr>
<td></td>
<td>Diagnosis of diabetes</td>
<td>Eur J Endocrinol^71</td>
</tr>
</tbody>
</table>

Abbreviations

AI, artificial intelligence; ANN, artificial neural network; CBR, case-based reasoning; CIG, computer interpretable guidelines; DT, decision trees; ES, expert systems; FL, fuzzy logic; GA, genetic algorithms; ML, machine learning; MLP, multilayer perceptron; MSE, mean square error; RBR, rule-based reasoning; SVM, support vector machines.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


