Fuzzy optimization to improve mobile health and wellness recommendation systems

Jozsef, Mezei; Nikou, Shahrokh

Published in:
Knowledge-Based Systems

DOI:
10.1016/j.knosys.2017.11.030

Publicerad: 01/01/2018

Document Version
(Referentgranskad version om publikationen är vetenskaplig)

Please cite the original version:
Jozsef, M., & Nikou, S. (2018). Fuzzy optimization to improve mobile health and wellness recommendation systems. *Knowledge-Based Systems*, *142*, 108–116. https://doi.org/10.1016/j.knosys.2017.11.030

Fuzzy optimization to improve mobile health and wellness recommendation systems

József Mezei[a,*], Shahrokh Nikou[b]

[a] *School of Business and Management, Lappeenranta University of Technology, Lappeenranta, South Karelia, Finland*
[b] *Faculty of Social Sciences, Business and Economics, Åbo Akademi University, Turku, Finland*

Abstract

In this article, we focus on mobile wellness and health-related applications from the perspective of the level of imprecision present in the data used in the recommendation systems. We propose a general fuzzy optimization model based on chanced constrained optimization to design recommendation systems that can take into consideration (i) the imprecision in the data and (ii) the imprecision by which one can estimate the effect of a recommendation on the user of the system. Our proposal is one of the first to use fuzzy optimization models in health-related decision making problems and the first to define a chance constrained optimization problem for interval-valued fuzzy numbers. The proposed approach identifies a set of actions to be taken by the users in order to optimize general health-related and/or wellness condition of the user from various perspectives. The model is illustrated through the example of walking speed optimization, with an additional numerical experiment offering a comparison with traditional methods.

Keywords:
Fuzzy Optimization, Mobile Health and Wellness Applications, Chance Constrained Programming, Linguistic Variables

---

*Corresponding author
Email addresses: jozsef.mezei@lut.fi (József Mezei), snikou@abo.fi (Shahrokh Nikou)

1. Introduction

One of the important components and features of mobile health and wellness applications can be a recommendation system that offers alternative actions to the user that can improve his/her health conditions [1]. Designing recommendation systems has also been an important problem in both academia and practice in recent years in the wellness domain [2, 3]. Recent advances in artificial intelligence can be observed in various practical applications in mobile and wearable devices offering health recommendations based on measuring different characteristics of the users and utilizing this data using various machine learning a d fuzzy techniques [1]. The wide variety of sensors available in different mobile and wearable devices makes it possible to collect large amount of personal health characteristics data supplemented by data collected in personal health record systems [4] that can be used as the basis of a wellness recommendation system [5]. The use of mobile technology can greatly improve the user's health status if the underlying recommendation tool properly takes into consideration the specific, context-dependent (and individual) requirements. A typical issue when dealing with data collected through sensors, specifically in mobile devices [6] or through manual input of users (a frequent case in mobile health/wellness applications) is that we have to deal with imprecise data. For example, a heart rate measurement for an individual: (i) provides an imprecise estimation of the actual heart rate and (ii) the same value for different individuals can indicate different health conditions. These observations illustrate the need for developing models that work appropriately with the imprecision present in the data [7]. As one of the most prominent tools to model imprecise information, fuzzy set theory has been introduced in [8]. The most important philosophical underpinning of the use of fuzzy sets in the medical context is Sadegh-Zadeh's characterization of health, illness and disease as fuzzy concepts [9].

By relying on various concepts from fuzzy set theory, in this article, *the core academic goal is to work out a set of general optimization models based on imprecise information that can be tailored to specific health-related (specifically wellness) decision problems in the form of a recommendation system.* In the optimization models, the parameters and the variables can take the form of crisp and fuzzy sets with various level of imprecision in order to address

[1] `https://www.indiegogo.com/projects/get-the-first-signs-live-longer-health#/`

the two main sources of imprecision discussed above. We formulate chance constrained models using fuzzy numbers and describe a procedure to extend a specific version of chance constrained optimization to interval-valued fuzzy sets. The use of the model is illustrated through the example of walking speed optimization by considering several health-related criteria. Furthermore, a numerical comparison is performed using a real world dataset to compare the performance of our model to traditional approaches in a binary classification problem.

The rest of this article is structured as follows. In Section 2, we briefly discuss the literature related to mobile health intervention, fuzzy sets in health-related decision making and present some preliminaries on modeling with interval-valued fuzzy sets. Section 3 presents an approach to personalized health and wellness decision support, starting from the representation of health data using fuzzy sets and moving on to the conceptual description of the optimization model. A numerical illustration of the general approach in the case of a recommendations system for running exercises and further numerical experiments are presented in Section 4. The conclusions and future work are discussed in Section 5.

2. Literature Review and Preliminaries

In this section, we discuss the most important contributions from the literature related to different aspects of the models presented in later sections regarding fuzzy sets in health-related decision making. Additionally, we define the basic preliminary concepts from possibility theory, and interval-valued fuzzy sets that are required later in developing the optimization models.

From clinical standpoint, mobile technology-based health interventions have the potential to provide supports and services to (i) healthcare providers (e.g., support in diagnosis or patient management) as well as (ii) communicate between healthcare services and patients (e.g., appointment reminders and test results notification)[10]. Digital devices, such as smartphones, smart wearable systems and tablets are becoming increasingly a potential medium for delivering health programs. In this article, we propose to utilize fuzzy optimization models to enhance health-related decision making and provide optimized recommendations to individuals. While research on this path is scarce, one related interesting example of the application of fuzzy logic is presented by [11], who use fuzzy ontology-based case-based reasoning and fuzzy semantic retrieval algorithm to answer complex medical queries related to semantic understanding of diabetes diagnosis.

2.1. Fuzzy sets for health related decision making

One important application of fuzzy set theory in decision making concerns the case of utilizing linguistic evaluations provided by experts. Medical or wellness-related decision making problems are typical examples of this phenomenon [12]. In the literature, different approaches relying on fuzzy sets have been frequently used in medical decision making problems [13, 14], including approximate reasoning [15], linguistic classification [16] and fuzzy clustering [17].

Additionally to traditional fuzzy sets, the number of contributions relying on interval-valued, intuitionistic, and hesitant fuzzy sets or vague sets [18], increases continuously. Particularly, interval-valued fuzzy sets have been applied extensively in health-related problems. For example, [19] develops a fuzzy rule-based classifier to estimate the risk of cardiovascular disease by utilizing extension of t-norms to interval-valued fuzzy sets. [20] designs an interactive group decision making method in order to handle the problem of patient-centered medicine concerning basilar artery occlusion.

2.2. Possibility theory and fuzzy sets

Formally, possibility [21] can be defined as a maxitive normalized monotone measure, i.e., the possibility that any of a set of events will occur equals the maximum of the possibilities of individual events.

When we consider fuzzy numbers, possibility of an event can be defined as follows:

$$\mathrm{Pos}(A \text{ is in } B) = \sup_{x \in B} A(x),$$

where $A(x)$ is the membership function of the fuzzy number A. This formula highlights the main difference between applying possibility and probability. While the probability that a random variable takes its value from a set is the sum of probabilities for the individual values in that set (the integral of the density function in the continuous case), the possibility is the maximum membership value taken on by the fuzzy number (possibility distribution).

When we compare (triangular) fuzzy numbers with crisp values, the constraints require to determine the possibility that a fuzzy number A is smaller than the crisp value b, which can be formulated as the special case of the above formula as follows with $B = [-\infty, b]$:

$$\text{Pos}(A \leq b) = \text{Pos}(A \text{ is in } B) = \sup \{A(x) \mid x \in B\}.$$

In case of equality, the possibility that a fuzzy number A is equal to the crisp number b is simply the value of the membership function of b in A: $A(b)$.

One important property of probability that is not satisfied by possibility is auto-duality ($P(A) = 1 - P(A^C)$) where A^C is the complement of A; for this reason, a dual measure, necessity, is needed:

$$\text{Nec}(A) = 1 - \text{Pos}(A^C).$$

In this article, we only use the above mentioned definitions, for further properties the reader can consult the references [21] and [22].

From the perspective of the underlying meaning, Nec offers a more meaningful choice in various applications as it expresses the idea that the user want to be confident *at least* to the specified extent in a given statement. For example, a constraint formulated in terms of the Nec measure regarding an acceptable upper bound for an attribute would require the optimal solution to ensure that the attribute is not larger than the given bound with a given level of confidence (typically chosen to be 0.9).

2.3. Preliminary definitions on interval-valued fuzzy sets

Interval-valued fuzzy sets extend traditional fuzzy sets by introducing a second level of imprecision as we assume even the membership values to be ill-known and hence represented by intervals from $[0, 1]$.

Definition 1 ([23])**.** *An interval-valued fuzzy set (IVFS) A defined on X is given by*

$$A = \left\{(x, [A_L(x), A^U(x)])\right\}, x \in X,$$

where $A_L(x), A^U(x) : X \to [0,1]; \forall x \in X, A_L(x) \leq A^U(x)$, *and the ordinary fuzzy sets* $A_L(x)$ *and* $A^U(x)$ *are called lower fuzzy set and upper fuzzy set about A, respectively. Interval-valued fuzzy numbers (IVFN's) are IVFS's when* A^L *and* A_U *are both ordinary fuzzy numbers [24].*

The γ-cut of an ordinary fuzzy number A is defined as $[A]^\gamma = \{x \in X | A(x) \geq \gamma\}$. The γ-cut is an interval for every γ, and can be represented using its endpoints $[A]^\gamma = [a(\gamma), b(\gamma)]$. In case of IVFN's, the γ-cut of the lower and upper fuzzy numbers, A_L and A^U, are denoted as $[A_L]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $[A^U]^\gamma = [A_1(\gamma), A_2(\gamma)]$.

Example 1. *A triangular fuzzy number can be specified with the following membership function:*

$$A(x) = \begin{cases} \dfrac{x-a}{b-a} & \text{if } a \leq x \leq b \\ \dfrac{c-x}{c-b} & \text{if } b \leq x \leq c \\ 0 & \text{otherwise,} \end{cases}$$

where $a \leq b \leq c$, $[a, c]$ is the support of the fuzzy number and b is the centre. We use the notation (a, b, c) to refer to a triangular fuzzy number. From the definition of an IVFN it follows that if both the upper and lower fuzzy numbers are triangular, i.e., $A_L = (a^L, b^L, c^L)$ and $A^U = (a^U, b^U, c^U)$, then the center values should be the same: $b^L = b^U$.

3. Personalized health and wellness decision support with fuzzy optimization

In this part of the article, we motivate the use of fuzzy sets, and particularly interval-valued fuzzy sets, in modelling health-related/wellness decision making problems. We then propose optimization models incorporating different levels of imprecision in the form of (i) crisp, (ii) fuzzy and (iii) interval-valued fuzzy parameters.

3.1. Representing wellness related data with (interval-valued) fuzzy sets

According to [25], a specific system (e.g. the human body) can be described by a set of characteristic properties, i.e., weight, height, blood pressure, body temperature, etc. The overall health state of a person is then determined as the function of the characteristics. Additionally, there are several possible actions that the individual can perform which affect the actual value of the characteristics, and in turn his overall health. In other worlds, we are dealing with a type of input-output system [26]. The inputs refer to the actions to be taken and the initial value of the characteristics (attributes of the system), while the output includes the modified values of the characteristics and the resulting overall health state. Formally, the system can be specified using a set of possible actions, $X = \{x_1, \ldots, x_n\}$, a set of characteristics/attributes, $P = \{p_1, \ldots, p_m\}$, and a function $F : X \to P$ specifying the impact of the actions on the characteristics.

On a practical level in our context, it means that a mobile wellness application monitors different performed activities and associated attributes of the user and based on analyzing the collected data can recommend taking a specific action to improve some of the attributes. Most of the components and overall goals relevant to this problem can only be specified in an imprecise manner as they depend on several pieces of information that by themselves are imprecise. According to [27], in the context of health and wellness-related decision making, three different levels of imprecision can be specified: (i) measurements that are approximately stable in a short time-horizon (e.g., age or weight of the person); (ii) measurements that can change rapidly (e.g., blood pressure or blood glucose level); (iii) subjective modifier of the previous measurements (e.g., low blood pressure).

The first set of characteristics in general can be described by using crisp values: age is 55 years or weight is 68 kilograms. Naturally, even these characteristics change in a short time period, but the magnitude of these changes does not affect the overall health state of the person. Using set theoretical notions, we can say that a person either belongs to the set of people who are aged 55 years (membership value 1) or does not belong (membership value 0).

To represent characteristics belonging to the second class, we can make use of fuzzy sets. As the blood pressure of a person can change significantly briefly after a measurement, the recorded data only approximates the imprecise value of the real blood pressure level. A realistic representation should assign non-zero membership value not only to the actual measurement, but also to values close to the observed value. The most commonly used fuzzy sets are triangular fuzzy numbers: we assign membership value 1 to the observed value, and using linear functions, assign memberships to numbers that are in the neighbourhood of the observation.

In general, we can specify a fixed set of triangular fuzzy numbers, usually associated with a linguistic scale, and after transforming the range of possible values of a characteristics to the $[0, 1]$ interval, we can use a fixed linguistic label set. For example, using a 5-scale evaluation, we can classify a specific value (with a corresponding triangular fuzzy number) as: *very low* ($(0, 0, 0.25)$), *low* ($(0, 0.25, 0.5)$), *average* ($(0.25, 0.5, 0.75)$), *high* ($(0.5, 0.75, 1)$), *very high* ($(0.75, 1, 1)$).

Linguistic modelling can also be applied to represent the third set of imprecise values to model expressions such as “high blood pressure" or “low blood glucose level". The linguistic labels are modelled as fuzzy sets with support in the unit interval and then applied on the interval of the feasible values of the specific health-related concept. All of the above quantities are special cases of interval-valued fuzzy numbers that provide a more complete modelling in our context.

The main idea of the models presented in the following sections is that we require that the "possibility" of achieving a goal is sufficiently high, or equivalently we require that the "necessity" that an event happens (the health condition or wellness of a person improves) is higher than a predefined threshold. It is important to mention here that, beside a few cases [28], there is a lack of research on applying fuzzy optimization tools in wellness or health-related decision problems. For overcoming this lack of research, we propose to use possibilistic chance constrained programming models [22] in developing health-related decision support systems.

To motivate this modeling choice, as we discussed above, one can claim that the data in health-related problems can be more classified as imprecise rather than random. According to this, in specifying the constraints of optimization models aimed at identifying the best action to improve the health/wellness state of a person, one can only require to be confident enough (for example, to the degree of 0.9), that the heart rate will go down from 120 and stabilizes around 80. In this example the imprecision that has to be taken into consideration has several components: (i) 120 is measured imprecisely by the device, offering an approximation of the real value; optimal heart rate depends on the person, e.g., for some 120 is already a very high value, for some it is only moderately high as their average is around 90; (iii) the threshold for the confidence level depends on the problem context (medical decision support requiring higher confidence versus wellness applications with lower acceptable confidence levels).

3.2. Model specifications

The characteristics relevant to the underlying health-related decision problem will be denoted by $p_1, p_2, \ldots, p_m$. These attributes can be quantified in terms of an acceptable range of values specified as an interval, $[p_i^l, p_i^u]$ (l stands for lower and u for upper bound). The acceptable range can be assumed to be estimated using general available statistics concerning the entire population under consideration when specifying the model. The identified attributes are used to formulate constraints on the choice of optimal values for the variables, $x_1, x_2, \ldots, x_n$. These variables of the optimization models are the actions in the terminology of the previous section that need to be selected in a way to improve the values of the attributes. For example, a typical problem is to specify the walking speed (the variable of the model) of a person in order to achieve optimal heart rate, calories brunt and fitness level (attributes of the model). As it is a typical feature of multiobjective optimization problems, the value of the variable that results in an "ideal" state in terms of one attribute, can have deteriorating impact on another attribute. For example, while increasing walking speed can result in more calories burnt, at the same time it can increase the heart rate beyond an acceptable limit.

When formulating the optimization models, we will assume a set of initial values of the variables, denoted by $x_1^0, x_2^0, \ldots, x_n^0$. The objective then becomes identifying new and optimal values for the variables that result in overall improved condition in terms of the attributes, while minimizing the effort to reach the improvements (in terms of change from the x_i^0 values to the x_i values). This could be formulated as a bi-objective optimization problem: (i) closeness to an ideal state in terms of the attributes as the first objective (measured by calculating the distance from this ideal state) and (ii) the effort required to reach this state as the second objective (loss function measuring the cost of change in terms of the variables). As it is commonly done, we can apply a scalarization procedure by taking the weighted average of the two objectives. According to this, the objective function of the problem is the following:

$$\begin{aligned} \text{minimize} \quad & w_I \text{dist}(I(p_1, p_2, \ldots, p_m), A(x_1, \ldots, x_n, x_1^0, \ldots, x_n^0, p_1, \ldots, p_m)) + \\ & w_E E(x_1, x_2, \ldots, x_n, x_1^0, x_2^0, \ldots, x_n^0) \end{aligned} \tag{1}$$

where $I(p_1, p_2, \ldots, p_m) = (I_1, I_2, \ldots, I_m)$ is the ideal point based on a set of m attributes (I_k is the ideal value for attribute k), $A = (A_1, A_2, \ldots, A_m)$ determines the state of attributes we achieve by changing the initial variable values $x_1^0, x_2^0, \ldots, x_n^0$ to $x_1, x_2, \ldots, x_n$ (A_k determines the state for attribute k), while the function E measures the cost of this change. The weights w_I and w_E specify the preference of the user between improving the attribute values and increasing the total effort of changing the variables.

The functions A, E and dist can be chosen according to the available knowledge in the specific context. As a basic option, one can utilize linear functions for A and E, implying that the improvement/deterioration in an attribute is directly proportional to the change in terms of the variables (this is the choice we will use in the later numerical examples). The use of linear functions in the objective ensures that the model is either directly linear, or can simply be transformed into a linear model, mitigating computational complexity. Alternatively, if there is available information on the more complex relationship between attributes and variables, for example in terms of a quadratic function, one can obtain a non-linear optimization problem. The distance function can also be specified in various ways, with one natural choice being the Euclidean distance, which is used in our application presented in the next section.

To expand on the function A in more details, we emphasize again that the ith component, A_i captures the change in the value of attribute i caused by the adjustment in the variables. The effect of the changes of the variable j on the attribute i can be captured through a predefined transformation function $f(x_j, x_j^0, p_i)$. The output of f is a (fuzzy) number, specifying what the change in the value of attribute i after the value of variable j changes from x_j^0 to x_j. Then the final value of A_i is the sum of all the changes caused by the individual variables as $A_i = \sum_j f(x_j, x_j^0, p_i)$. In the simplest case, this transformation can be a linear function, e.g., a specific amount of increase in the variable can be simply proportional to the change in the value of the attribute. In more complex cases, the transformation can be logarithmic, implying that changing the variable in the beginning causes a significant change in the attribute, however, this change becomes less relevant when the variable is increased further.

Additionally to the above objective function, some constraints need to be specified to complete the optimization model. For every attribute, we can require through two constraints that the effect of changing variable values does not change the value of an attribute in a way that moves it outside its acceptable region as $f(x_j, x_j^0, p_i) \geq p_i^l$ and $f(x_j, x_j^0, p_i) \leq p_i^u$, and this should also hold for the overall aggregated change expressed using A_i.

Example 2. *We can exemplify the presented general ideas to optimize walking speed, the case that will be used in the later part of the article in a more detailed numerical example. In this example, there is one possible action, changing the walking speed (x_1). The characteristics of the person performing the walking exercise include heart rate (p_1), number of calories burnt per minute (p_2), and fitness level (p_3). According to this, the function F can be* *specified using three component functions,* $F = (f_1, f_2, f_3)$*, where* f_i *quantifies how a change in* x_1 *(Walking speed) impacts the three characteristics. Using our above terminology, the input is walking speed, and when we apply this input to the system (human body), it results in specific output (heart rate, burnt calories, fitness) values. The task then is to determine the input value that results in optimal output values. This is naturally a multiple objective problem, and this is what we will formulate and discuss as the main part of the article in later sections.*

To take a concrete representation, the actual value of x_1 *can be a real number specifying to what extent the actual speed value is changed. For example* $x_1 = 0.05$ *would mean a* 5% *increase in speed, while* $x_1 = -0.05$ *would mean a* 5% *decrease in speed. In order to simplify the description, here we assume that there is only one attribute,* p_1*, the heart rate of the person. One possible complete specification to form the optimization model can be the following:*

- *the feasible interval for the attribute,* $[p_i^l, p_i^u] = [60, 140]$*;*
- *the ideal value for the attribute,* $I = 120$*;*
- *transformation function specifies that a given percentage of change in walking speed results in the same percentage of change in heart rate,* $f(x_1, x_1^0, p_1) = p_1 * (x_1 - x_1^0)/x_1$
- *feasible interval for the change in the variable,* $x_1 \in [-0.5, 0.5]$*;*
- *weights of the two terms in the objective function are equal,* $w_I = w_E = 0.5$.

3.3. Chance constrained optimization with interval-valued fuzzy sets

In this section, we will describe three different approaches that differ mainly in the form of the input data used as the value of parameters and variables in the optimization models. We will consider the following three cases:

- *Crisp optimization model*: only crisp values are used in the optimization model (potentially obtained as transforming IVFN's into crips values).
- *Possibilistic chance constrained optimization*: ordinary fuzzy numbers and traditional possibilistic chance constrained programming is applied.

- *Interval-valued chance constrained optimization*: Two ordinary fuzzy numbers (maximum and minimum embedded fuzzy numbers) are obtained from the interval-valued parameters and used separately in different constraints of the model.

3.3.1. Model 1: Optimization with crisp values

In the first model, the parameters and the variables are represented by crisp numbers. According to this, we assume that the data collected is accepted as precise enough to use it in the analysis. In order to work with a completely crisp model, the fuzzy parameters (linguistic labels) have to be transformed into crisp values. For this purpose, we utilize the possibilistic mean value of $A \in \text{IVFN}$, which is defined in [29].

$$E(A) = \int_0^1 \gamma \frac{a_1(\gamma) + a_2(\gamma) + A_1(\gamma) + A_2(\gamma)}{2} \mathrm{d}\gamma \tag{2}$$

The crisp model can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & w_I \text{dist}(I(p_1, p_2, \ldots, p_m), A(x_1, x_2, \ldots, x_n, p_1, p_2, \ldots, p_m)) + \\ & w_E E(x_1, x_2, \ldots, x_n, x_1^0, x_2^0, \ldots, x_n^0) \\ \text{subject to} \quad & f(x_j, x_j^0, p_i) \geq p_i^l, \qquad \forall i, j \\ & f(x_j, x_j^0, p_i) \leq p_i^u, \qquad \forall i, j \\ & E(x_1, x_2, \ldots, x_n, x_1^0, x_2^0, \ldots, x_n^0) \leq E_{\max} \\ & x_j \in [x_j^l, x_j^u], \qquad \forall j \end{aligned} \tag{3}$$

where $[p_i^l, p_i^u]$ is the acceptable region for the parameter i, $E_{\max}$ is the maximum energy to be spent on changing the values of the variables and $[x_j^l, x_j^u]$ is the feasible interval of variable j. The output of this model is a recommendation on how to adjust specific actions in order to achieve the optimal combination of attributes. The first two set of constraints ensure that the value of attributes remains in the acceptable interval, while the last constraint puts a limit on the maximum energy that can be utilized in order to enforce changes.

3.3.2. Model 2: Chance constrained programming with fuzzy numbers

As we discussed in previous sections, in practice, the data utilized by wellness applications relying on sensors built in the mobile/wearable devices can rarely be considered as an exact value, rather an approximation of the real value of a health characteristics. To tackle this problem, chance constrained programming with imprecise parameters can be utilized, making use of the concept of possibility and necessity measures recalled in the previous section.

Fuzzy values can be incorporated in the optimization model (3) on several levels, with the two most important being the following: (i) the values of attributes and variables and (ii) the feasible interval for attributes and variables of the model (a general average value describing the population of users can be combined with user-specific values to construct an interval of possible values).

When using fuzzy values, a distance measure has to be used in order to calculate the deviation from the optimal attribute values. For this purpose, the possibilistic mean value can be utilized: the distance of two fuzzy numbers, A and B can be calculated as the absolute value of the mean value of $A - B$ [29].

Based on these considerations, we can formulate the chance constrained optimization model as follows (˜indicates that the function/variable/parameter is fuzzy-valued):

$$\begin{aligned}
\text{minimize} \quad & w_I \text{dist}(I(\tilde{p_1}, \tilde{p_2}, \ldots, \tilde{p_m}), A(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{p_1}, \tilde{p_2}, \ldots, \tilde{p_m})) + \\
& w_E E(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{x_1^0}, \tilde{x_2^0}, \ldots, \tilde{x_n^0}) \\
\text{subject to} \quad & \text{Nec}(\tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i}) \geq p_i^l) \geq \eta_{i,j}^l, \qquad \forall i, j \\
& \text{Nec}(\tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i}) \leq p_i^u) \geq \eta_{i,j}^u, \qquad \forall i, j \\
& \text{Nec}(\tilde{E}(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{x_1^0}, \tilde{x_2^0}, \ldots, \tilde{x_n^0}) \leq E_{\max}) \geq \eta_E \\
& \text{supp}\, \tilde{x_j} \subset [x_j^l, x_j^u], \qquad \forall j
\end{aligned} \tag{4}$$

where, additionally to the previously introduced notations, $\eta_{i,j}^l$ and $\eta_{i,j}^u$ are confidence values specified by the user and can be based on some general medical knowledge; η_E stands for the confidence level regarding the energy used for performing actions according to the recommendations. Following the same reasoning as the previous simpler model, the first two sets of constraint ensure with a given confidence that the final value of attributes is in the acceptable interval.

As $\tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i})$ is a fuzzy number, if we denote its membership function by $\tilde{f}_{i,j}(x)$, the left-side of the two constraints involving the transformation function can be rewritten as

$$\begin{aligned}\text{Nec}(\tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i}) \geq p_i^l) = 1 - \sup\{f_{i,j}(x) | x \leq p_i^l\} \\ \text{Nec}(\tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i}) \leq p_i^u) = 1 - \sup\{f_{i,j}(x) | x \geq p_i^u\}\end{aligned} \tag{5}$$

These equalities follow from the definition of the necessity function. Using the formula specifying the possibility of a fuzzy set smaller than a crisp value and the definition of necessity as one minus the possibility of the complement, we can easily see that the above reformulations follow from the original form of the constraints.

The same reformulation can be done for the constraint limiting the used energy with the membership function $\tilde{E}(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{x_1^0}, \tilde{x_2^0}, \ldots, \tilde{x_n^0})$.

3.3.3. Model 3: Chance constrained programming with interval-valued fuzzy numbers

While transforming IVFN's into ordinary fuzzy numbers or crisp values clearly simplifies the problem formulation and the computational complexity of optimization models (for a detailed discussion on fuzzy and possibilistic optimization models see [30]), it can lead to the loss of important information as a consequence of reducing the level of imprecision specific to the user.

Note 1. *Regarding the complexity of the models, we note that using for example triangular fuzzy numbers triples the number of variables and parameters from Model 1 to Model 2, and doubles from Model 2 to Model 3 (instead of a single fuzzy number, we deal with upper and lower fuzzy numbers). While the complexity in terms of the parameters increases 6-fold, as we are dealing with linear optimization models in general, solving the fuzzy models still remains feasible in real time with traditional optimization algorithms.*

In order to make use of interval-valued fuzzy numbers in the optimization problems, chance constrained optimization has to be extended to this wider class of fuzzy sets. While the constructs of possibility and necessity are not generalized to the case of higher level fuzzy numbers in the literature, the use of these measures in the proposed models are restricted to the specific case of comparing a fuzzy number to a crisp value. In this case, as we have shown in equation (5), we can directly operate on the membership values. The problem of extending the model requires the use of a function to compare intervals and crisp values. For this purpose, we can define the following function, F_l, to quantify the extent to which an interval, $[a, b]$ is larger than a crisp value, c:

$$F_l(c|a,b) = \begin{cases} 1 & \text{if } c \leq w \\ \dfrac{b-c}{b-w} & \text{if } w \leq c \leq b \\ 0 & \text{otherwise,} \end{cases}$$

where $w = w_1 a + w_2 b$, with $w_1 + w_2 = 1, w_1, w_2 > 0$. The weights in our context can represent the preferences of the user with respect to the upper and lower estimations of the membership value.

Similarly, we can define the following function, F_s, to quantify the extent to which an interval, $[a, b]$ is smaller than a crisp value, c:

$$F_s(c|a,b) = \begin{cases} 0 & \text{if } c \leq a \\ \dfrac{c-w}{b-w} & \text{if } w \leq c \leq b \\ 1 & \text{otherwise,} \end{cases}$$

Note 2. *We note here that this procedure corresponds to fitting a possibility distribution to the $[a, b]$ interval and calculating the necessity measure corresponding to the defined possibility. In other words, we define the constraint in terms of a possibility distribution on both considered levels of imprecision (on the values and on the membership of the values).*

Based on the above considerations, the extension of the chance constrained model (4) based on our proposal can be formulated as:

$$\begin{aligned} \text{minimize} \quad & w_I \text{dist}(I(\tilde{p_1}, \tilde{p_2}, \ldots, \tilde{p_m}), A(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{p_1}, \tilde{p_2}, \ldots, \tilde{p_m}))+ \\ & w_E E(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{x_1^0}, \tilde{x_2^0}, \ldots, \tilde{x_n^0}) \\ \text{subject to} \quad & F_l(p_i^l | \tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i})) \geq \eta_{i,j}^l, \qquad \forall i, j \\ & F_s(p_i^u | \tilde{f}(\tilde{x_j}, \tilde{x_j^0}, \tilde{p_i})) \geq \eta_{i,j}^u, \qquad \forall i, j \\ & F_s(E_{\max} | \tilde{E}(\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}, \tilde{x_1^0}, \tilde{x_2^0}, \ldots, \tilde{x_n^0})) \geq \eta_E \\ & \text{supp}\, \tilde{x_j} \subset [x_j^l, x_j^u], \qquad \forall j \end{aligned} \tag{6}$$

Note 3. *Model (6) is naturally a generalization of the models presented in previous sub-sections. If we replace every interval-valued fuzzy parameters* *and variables with a triangular fuzzy number or a crisp value in (6), the model collapses into (4) and (3), respectively. This is the consequence of operations performed on triangular fuzzy numbers when moving from Model 1 to Model 2 (as crisp numbers can be considered as special cases of triangular fuzzy numbers when the center is the same as the left and right end-points of the support) and the consequence of the fact that the possibilistic mean of an IVFN is the same as the possibilistic mean of an embedded triangular fuzzy number [29].*

4. Numerical examples

In this section, we will exemplify the specification and use of the proposed general models in the case of a wellness recommendation system that helps the user to keep the optimal pace of the walking. Additionally, a numerical experiment is conducted to compare the performance of the presented approach to traditional techniques.

4.1. Use scenario: recommending walking speed

Without loss of generality, we can assume that the ranges of all the variables and parameters are in the $[0, 1]$ interval, with 1 indicating the best value of that attribute for the user and 0 indicating the worst. Naturally, this representation is not necessarily obtained through a linear transformation of the original values. For example, very high and very low heart rate values will be mapped to 0, while heart rates close to 70 or 80 will be mapped to a value close to 1.

This example is a typical case which is already supported on various mobile and wearable devices, although not necessarily with a recommendation system using advanced mathematical tools. All the parameters that we consider in this example can be measured by or retrieved on the mobile device in real-time manner. The parameters in this specific application can include the following:

- weather, including, for example, temperature and the intensity of rain and wind;
- route, by modeling the slope and quality of the road and its effect on the energy required for a specific walking speed;
- age, weight and height of the user;

- other health characteristics such as heart rate, blood pressure, blood glucose level, etc.;
- time elapsed from the last meal and number of calories consumed during the day;
- more parameters, depending on the special needs of the user.

The variable of the decision model is walking speed: the system provides recommendation to the user on increasing, decreasing or maintaining the actual speed in order to optimize the value of some health characteristics. The parameters of the model (the attributes of the user) are heart rate, number of calories burnt per minute, and a physical fitness level indicator. The ideal solution in this example is simply obtained as the weighted average of the ideal values of the individual attributes (we will use the values $(0.3, 0.4, 0.3)$ for the three attributes in this example). In Table 1, the actual and ideal values are listed. The parameters for all the three models (crisp, fuzzy and interval-valued fuzzy) are listed, with the triplets representing triangular fuzzy numbers (upper and lower fuzzy numbers in the interval-valued case). Based on these numbers and weights, the ideal point would be $I^1 = 0.7$, $I^2 = (0.6, 0.7, 0.8)$ and $I^3_L = (0.6, 0.7, 0.8), I^3_L = (0.65, 0.7, 0.75)$ for Model 1, 2, and 3, respectively.

Actual speed of the user is specified as 0.6 for Model 1, $(0.5, 0.6, 0.7)$ for Model 2, and $(0.5, 0.6, 0.7)$, $(0.55, 0.6, 0.65)$ are the upper and lower fuzzy numbers in Model 3. Increasing the speed increases the heart rate, improves on the calories burnt, and improves the physical fitness. For simplicity, we assume that change in speed affects all the attributes to the same extent: 1 % change in speed results in 1 % change in the attribute; this means that the f transformation functions take a simple linear form as shown in Example 3. The constraints require, with $\eta = 0.9$ confidence level in the fuzzy models, that the distance from the ideal attribute value should not be more than 0.25. In the objective function, the weight of the two components are equal. To summarize in terms of the notations specified in the previous section, we have the following components in the model:

- one variable, x, the change in the walking speed, and the initial speed value x^0;
- three attributes, p_1, p_2, p_3, are heart rate, number of calories burnt per minute, and fitness indicator with corresponding acceptable intervals;

- the ideal point I specified above;
- the weights of the two components in the objective function as 0.5;
- the functions A_1, A_2, A_3, specifying the change in the value of the three attributes after the variable is changed, as specified simply multiplying the actual value of the attribute by $(x - x^0)/x^0$;
- dist is specified as the Euclidean distance.

Value/attribute	Heart rate	Calories	Fitness
Model 1 actual	0.7	0.4	0.9
Model 1 ideal	0.8	0.7	0.6
Model 2 actual	(0.6, 0.7, 0.8)	(0.3, 0.4, 0.5)	(0.8, 0.9, 1.0)
Model 2 ideal	(0.7, 0.8, 0.9)	(0.6, 0.7, 0.8)	(0.5, 0.6, 0.7)
Model 3 actual	(0.60, 0.70, 0.80) (0.65, 0.70, 0.75)	(0.30, 0.40, 0.50) (0.35, 0.40, 0.45)	(0.80, 0.90, 1.0) (0.85, 0.90, 0.95)
Model 3 ideal	(0.70, 0.80, 0.90) (0.75, 0.80, 0.85)	(0.60, 0.70, 0.80) (0.65, 0.70, 0.75)	(0.50, 0.60, 0.70) (0.55, 0.60, 0.65)

Table 1: Parameters of the optimization models

The optimization problem was implemented in R [31] utilizing the *lpSolve* package to find the optimal solution. The optimal solution was identified as increasing the walking speed by $S^1 = 0.27$, $S^2 = (0.21, 0.26, 0.32)$, $S^3_U = (0.22, 0.27, 0.33)$, $S^3_L = (0.24, 0.27, 0.31)$ in the three models, respectively. As one can observe, the higher the level of imprecision in the model, the more realistic the recommendation provided by the system is. This resulted in the attribute values listed in Table 2. As one would expect, by increasing (decreasing) the η values, the solution moves closer to (further from) the solution of Model 1.

4.2. Numerical experiment

In this section, a numerical evaluation of the developed models is presented. In contrast to the previous section, a different type of utilization of the model for a discrete optimization problem is described. As the developed model has not been implemented yet as an actual mobile application to allow for a direct validation, we opted for comparing the performance of the method on a real-life dataset used in various machine learning studies.

Attribute/value	Model 1	Model 2	Model 3
Heart rate	0.89	(0.77, 0.89, 1.0)	(0.77, 0.89, 1.0) (0.83, 0.89, 0.95)
Calories	0.51	(0.38, 0.51, 0.64)	(0.38, 0.51, 0.64) (0.43, 0.51, 0.57)
Fitness	0.97	(0.92, 0.97, 1.0)	(0.92, 0.97, 1.0) (0.95, 0.97, 0.99)

Table 2: Final attribute values

Our goal is to show that the performance of our proposal is comparable to established classification methods in case of this specific dataset. When considering the additional benefits of the model regarding modelling imprecision, ease of use and interpretability, these factors together illustrate the potential of the developed models.

For testing the presented method, one would need a multivariate time-series in order to appropriately assess the impact of changes caused by one or several variables on several outputs, which is rarely the case in publicly available datasets. Specifically, medical datasets describing the effect of a treatment or a drug given to a patient on a regular basis are usually confidential and not available in public databases. For these reasons, we found the dataset describing the occupancy of a room as the outcome variable appropriate for validation purposes in case of the discrete optimization based recommendations[2]. The dataset includes the following attributes: temperature, relative humidity, light, CO_2, and humidity ratio [32].

This model can be utilized to test our model as follows. The assumption of a single variable impacting the attributes is fulfilled in this case: a person entering or leaving the room causes the attributes to change. The task is to specify whether the occupancy status of the room should be changed in order to attain the required attribute values, which is a binary classification problem. In order to utilize the developed models, the functional relationship between occupancy and other variables and the value of an ideal point need to be specified. As there is no information on the cost of somebody entering/leaving the room, this component of the model is left out in this experiment. The two required components have been identified following

[2] http://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+

a traditional training-testing process used in machine learning: part of the dataset is used to specify the models and check in-sample classification performance, and other datapoints were used to test out-of-sample performance of the optimal model.

We assumed a linear relationship between individual attributes and the occupancy variable, resulting in two parameters per attributes to be estimated. Additionally, we assumed that it takes some time for the change in occupancy to take effect on the variables. According to this, when determining the optimal class for each case in the training phase, we considered the average of the attributes in the next 20 time periods to capture the resulting changes. Based on these specifications, in the model building we tried to find the optimal coefficients of the linear relationship and the optimal occupancy class such as the resulting attribute values minimize the distance from the calculated average of the following 20 datapoints. This procedure resulted in the optimal relationship coefficients to be used on the test dataset. In the out-of-sample evaluation, the mean value of occupancy observations from the training data is used as the ideal point to ensure the real-time nature of the experiment.

In the numerical comparison, we utilized the most widely used classification algorithms, including classification trees (CT), support vector machines (SVM) and neural networks (NN), which in general found to be the optimal choice of classification algortihms in most or real-life applications [33]. The experiments were implemented in R [31].

The evaluation of binary classification methods usually relies on the elements of the contingency matrix: (i) true positive (TP), i.e., positive cases classified correctly as positive; (ii) true negative (TN), i.e., negative cases classified correctly as negative;(i) false positive (FP), i.e., negative cases classified incorrectly as positive;(iv) false negatives (FN), i.e., positive cases classified incorrectly as negative. The simplified version of the problem in the previous section could be a classification problem, by predicting only whether the person should decrease or increase speed.

Utilizing these notations, the most widely used performance measures is accuracy (ACC), formally calculated as

$$ACC = \frac{TP + TN}{TP + TN + FN + FP}.$$

This measure simply describes how frequently the predictions are correct.

While this is a straightforward and intuitive measure, it can offer wrong indication regarding performance; for this reason, other measures, such as true positive rate (TPR, also called sensitivity or recall)

$$TPR = \frac{TP}{TP + FN},$$

true negative rate (TNR, also called specificity)

$$TNR = \frac{TN}{TN + FP},$$

and Cohen's κ [34] can be used. In classification tasks one usually differentiates between in-sample and out-of-sample evaluation. In the first case, all the datapoints that are used to build the model are also used in assessing the performance of the model. A more general procedure relies on randomly selecting a subset of the available data for model building and then testing the model on the remaining datapoints (out-of-sample evaluation).

Table 3: In-sample classification performance with the proposed approach

	Accuracy	**Specificity**	**Sensitivity**	**Cohen's κ**
Model 1	0.95	0.94	0.96	0.94
Model 2	0.95	0.95	0.95	0.95
Model 3	0.97	0.97	0.98	0.97

Table 4: Out-of-sample classification performance with the proposed approach

	Accuracy	**Specificity**	**Sensitivity**	**Cohen's κ**
Model 1	0.91	0.89	0.92	0.88
Model 2	0.93	0.95	0.91	0.92
Model 3	0.96	0.95	0.97	0.94

The results of the experiments are presented in Tables 3, 4, 5 and 6, respectively. As one can observe, the performance of the proposed approach for each of the models is sufficiently close to that of the traditional approaches. This is particularly the case for the most general Model 3, which outperforms neural networks. There is a larger decrease in performance in the out-of-sample analysis for our model, which can be attributed to the choice

Table 5: In-sample classification performance with traditional approaches

	Accuracy	**Specificity**	**Sensitivity**	**Cohen's κ**
CT	0.98	0.99	0.98	0.96
NN	0.96	0.97	0.95	0.94
SVM	0.98	0.97	0.99	0.97

Table 6: Out-of-sample classification performance with traditional approaches

	Accuracy	**Specificity**	**Sensitivity**	**Cohen's κ**
CT	0.95	0.96	0.94	0.91
NN	0.93	0.91	0.94	0.89
SVM	0.97	0.96	0.97	0.95

of specifying the ideal point as mean values from the training set. In case of using a proper recommendation dataset as described in the first example, this would not cause a problem as the ideal point could be defined more generically. As a summary, considering the additional benefits of our proposed algorithm discussed above, these results highlight the potential of the approach in real-life applications.

5. Conclusion

In this article, we present a new way of looking at health-related data used in recommendation systems and utilize it specifically in mobile wellness applications. The aim of the presented models is to identify the best course of action in a given situation that can help to improve (or simply maintain) the wellness condition of the user. Throughout the article, we reasoned why there is a need for models to incorporate imprecise information in health-related decision making problems and how fuzzy set theory provides a tool for decision support. Contributions such as the one presented in the paper can further pave the way for artificial and computational intelligence methods into intelligent mobile and wearable devices that continuously collect data on customers and offering health and wellness recommendations [3].

The main difference compared to traditional fuzzy models is that we do not rely on IF-THEN rules, but define the ideal solutions and identify opti-

[3] `https://maphealthwatch.com/technology/ai/`

mal actions to be performed by: (i) minimizing a distance (in a possibilistic sense) from the ideal solution, while (ii) minimizing the effort necessary to perform the action. The main contribution of this article is that it is one of the first approaches (i) utilizing fuzzy optimization models in health-related decision making models and (ii) building wellness recommendation systems aimed at mobile devices by making use of fuzzy set theory-based models, and (iii) offering an approach to extend chance constrained programming to interval-valued fuzzy numbers. The framework has been exemplified through the case of walking pace recommendation. Additionally, a numerical comparison illustrated that the performance of the proposed approach is comparable to traditional algorithms in case of a binary classification task. It is important to mention from the perspective of usability of our proposal, that while the presented models are general in the sense that they can work with any membership function, in the numerical examples we only used triangular fuzzy numbers. Still, the results show that even with this simplest membership function that can be used and can be easily understood and interpreted by practitioners, we achieved sufficient performance in the model comparisons.

The most important future research direction would be to perform user studies to test the models in different contexts to further evaluate the performance, particularly in the most general form of the model for prediction purposes. As the presented numerical study evaluates the performance on a dataset with moderate size and with a simplified version requiring a binary output, it will be an important task in the future to test the model on a larger data instances, particularly taking into consideration the specific environment of running the models on mobile devices. Furthermore, as it is pointed out in the article, in many practical situations we have to deal with different levels of imprecision: in the future we can improve the model by incorporating different families of fuzzy sets.

References

[1] M. Wiesner, D. Pfeifer, Health recommender systems: Concepts, requirements, technical basics and challenges, International Journal of Environmental Research and Public Health 11 (3) (2014) 2580–2607.

[2] W. Husain, L. T. Pheng, The development of personalized wellness therapy recommender system using hybrid case-based reasoning, in: Computer Technology and Development (ICCTD), 2010 2nd International Conference on, IEEE, 2010, pp. 85–89.

[3] T. P. Lim, W. Husain, N. Zakaria, Recommender system for personalised wellness therapy, International Journal of Advanced Computer Science and Applications 4 (9) (2013) 54–60.

[4] I. Genitsaridi, H. Kondylakis, L. Koumakis, K. Marias, M. Tsiknakis, Towards intelligent personal health record systems: Review, criteria and extensions, Procedia Computer Science 21 (2013) 327–334.

[5] Z. Yumak, P. Pu, Survey of sensor-based personal wellness management systems, BioNanoScience 3 (3) (2013) 254–269.

[6] S. B. Ahire, H. K. Khanuja, Healthcare recommendation for personalized framework, International Journal of Computer Applications 110 (1).

[7] J. Mezei, S. Nikou, Fuzzy optimization to improve mobile wellness applications for young-elderly, in: Proceedings of the 28th Bled eConference (eWellBeing 2015), 2015, pp. 263–275.

[8] L. A. Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) 338–353.

[9] K. Sadegh-Zadeh, Fuzzy health, illness, and disease, Journal of Medicine and Philosophy 25 (5) (2000) 605–638.

[10] C. Free, G. Phillips, L. Watson, L. Galli, L. Felix, P. Edwards, V. Patel, A. Haines, The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis, PLoS Medicine 10 (1) (2013) e1001363.

[11] S. El-Sappagh, M. Elmogy, A. M. Raid, A fuzzy-ontology oriented case-based reasoning framework for semantic diabetes diagnosis, Artificial Intelligence in Medicinedoi:10.1016/j.artmed.2015.08.003.

[12] N. H. Phuong, V. Kreinovich, Fuzzy logic and its applications in medicine, International Journal of Medical Informatics 62 (2) (2001) 165–173.

[13] M. Mahfouf, M. F. Abbod, D. A. Linkens, A survey of fuzzy logic monitoring and control utilisation in medicine, Artificial Intelligence in Medicine 21 (1) (2001) 27–42.

[14] J. H. Bates, M. P. Young, Applying fuzzy logic to medical decision making in the intensive care unit, American Journal of Respiratory and Critical Care Medicine 167 (7) (2003) 948–952.

[15] M. Pota, M. Esposito, G. De Pietro, Designing rule-based fuzzy systems for classification in medicine, Knowledge-Based Systems 124 (2017) 105–132.

[16] D. Nauck, R. Kruse, Obtaining interpretable fuzzy classification rules from medical data, Artificial Intelligence in Medicine 16 (2) (1999) 149–169.

[17] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, T.-J. Chen, Fuzzy c-means clustering with spatial information for image segmentation, Computerized Medical Imaging and Graphics 30 (1) (2006) 9–15.

[18] P. Wang, K.-M. Chao, C.-C. Lo, Satisfaction-based web service discovery and selection scheme utilizing vague sets theory, Information Systems Frontiers 17 (4) (2015) 827–844.

[19] J. A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, H. Bustince, Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system, Applied Soft Computing 20 (2014) 103–111.

[20] T.-Y. Chen, An interactive method for multiple criteria group decision analysis based on interval type-2 fuzzy sets and its application to medical decision making, Fuzzy Optimization and Decision Making 12 (3) (2013) 323–356.

[21] L. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1) (1978) 3 – 28. doi:10.1016/0165-0114(78)90029-5. URL `http://www.sciencedirect.com/science/article/pii/0165011478900295`

[22] B. Liu, K. Iwamura, Chance constrained programming with fuzzy parameters, Fuzzy Sets and Systems 94 (2) (1998) 227–237.

[23] M. B. Gorzałczany, A method of inference in approximate reasoning based on interval-valued fuzzy sets, Fuzzy Sets and Systems 21 (1) (1987) 1–17.

[24] G. Wang, X. Li, The applications of interval-valued fuzzy numbers and interval-distribution numbers, Fuzzy Sets and Systems 98 (3) (1998) 331–335.

[25] R. Seising, From vagueness in medical thought to the foundations of fuzzy reasoning in medical diagnosis, Artificial Intelligence in Medicine 38 (3) (2006) 237–256.

[26] C. Carlsson, M. Brunelli, J. Mezei, A soft computing approach to mastering paper machines, in: 46th Hawaii International Conference on System Sciences (HICSS), IEEE, 2013, pp. 1394–1401.

[27] D. L. Hudson, M. E. Cohen, Fuzzy logic in medical expert systems, Engineering in Medicine and Biology Magazine, IEEE 13 (5) (1994) 693–698.

[28] T. Aribarg, S. Supratid, C. Lursinsap, Optimizing the modified fuzzy ant-miner for efficient medical diagnosis, Applied Intelligence 37 (3) (2012) 357–376.

[29] C. Carlsson, R. Fuller, J. Mezei, Project selection with interval-valued fuzzy numbers, in: 12th International Symposium on Computational Intelligence and Informatics (CINTI), IEEE, 2011, pp. 23–26.

[30] J. Kacprzyk, S. A. Orlovski, Optimization models using fuzzy sets and possibility theory, Vol. 4, Springer Science & Business Media, 2013.

[31] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2016). URL `https://www.R-project.org/`

[32] L. M. Candanedo, V. Feldheim, Accurate occupancy detection of an office room from light, temperature, humidity and co 2 measurements using statistical learning models, Energy and Buildings 112 (2016) 28–39.

[33] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.

[34] J. Cohen, A coefficient of agreement for nominal scales, Educational and psychological measurement 20 (1) (1960) 37–46.