

РАСЧЕТ РИСКА ЗАБОЛЕВАНИЯ НА ОСНОВЕ ОБЪЕДИНЕННОГО ПОДХОДА Т. СААТИ И МАМДАНИ

Дедина Анастасия Валерьевна

*студент,
Белорусский государственный университет
информатики и радиоэлектроники,
Республика Беларусь, г. Минск*

Герман Юлия Олеговна

*научный руководитель, канд. техн. наук, доц.
Белорусского государственного университета
информатики и радиоэлектроники,
Республика Беларусь, г. Минск*

В последние годы все больше внимания уделяется применению математических методов для оценки риска заболеваний. В работе рассмотрен подход, объединяющий технику принятия решений Т.Саати и нечеткий вывод Мамдани.

Метод анализа иерархий (Саати) представляет собой систему оценки, которая позволяет иерархически организовать различные факторы риска и присвоить им весовые коэффициенты в соответствии с их важностью. Этот метод помогает оценить риск заболевания, связанный с различными факторами, такими как генетическая предрасположенность, питание, образ жизни, возраст и другие. Саати-анализ помогает определить, какие факторы риска являются наиболее критическими, и планировать меры по снижению уровня заболевания.

Не менее важным методом принятия решений является метод нечеткого вывода Мамдани, который позволяет оценить риск заболевания с использованием нечетких правил [1]. Это помогает сделать более точные выводы о риске заболеваний и определить подходящую терапию. На основе объединения указанных подходов была разработана методика оценки риска заболеваний.

Итак, первый этап анализа состоит в выборе факторов риска, которые влияют на вероятность развития заболевания. Для каждого из них формируются возможные показатели, которые могут варьироваться в зависимости от контекста и конкретных условий. Составляются графики функций полезности по критериям. Например, для критериев «Уровень физической активности» (рисунок 1) и «Возраст» (рисунок 2) используются следующие графики.

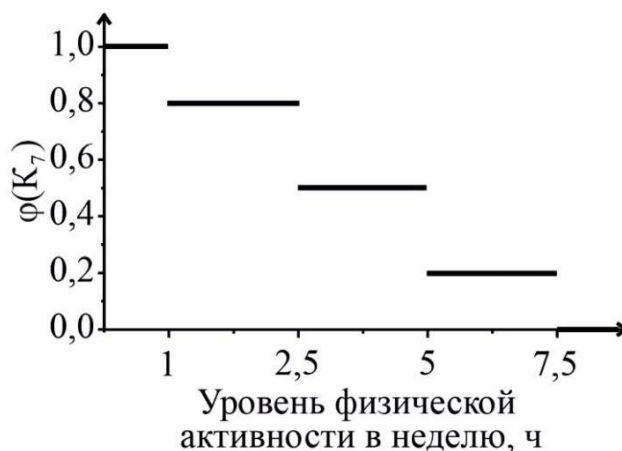


Рисунок 1. Функция полезности по физической активности

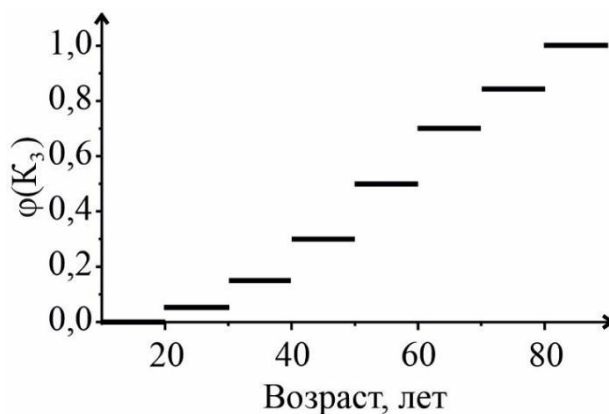


Рисунок 2. Функция полезности по критерию «Возраст»

Для некоторых критериев, например, «Уровень глюкозы» и «Холестерин» вводится нечеткая функция полезности. Это делается для того, чтобы минимизировать ошибку в постановке диагноза.

Нечеткая функция полезности – это нечеткая функция, для значений которой указываются меры истинности этих значений в контексте решаемой задачи. Мера (степень) истинности – числовое значение, которое определяет степень соответствия высказывания реальности.

Для составления базы правил значения критерия «Уровень глюкозы» были разбиты на четыре множества: высокий уровень содержания глюкозы в крови, уровень с содержанием глюкозы в крови выше среднего уровня, нормальный уровень и низкий.

Нормальный уровень глюкозы в крови через два часа после приема пищи для здорового человека составляет от 70 до 130 мг/дл. Уровень глюкозы в крови более от 130 мг/дл до 200 мг/дл составляет уровень выше среднего и соответствует преддиабетическому состоянию. Уровень содержания глюкозы более 200 мг/дл говорит о диагнозе сахарного диабета и соответствует высокому уровню. Низкий уровень глюкозы в крови называется гипогликемией и определяется, когда уровень глюкозы составляет ниже 60 мг/дл.

Далее каждое множество разделяется на меньшие подмножества для достижения большей точности. Здесь же осуществляется механизм фазификации: значение X отображается на нечеткое множество функций полезности, а для каждой из функций задается своя нечеткая мера истинности.

Пусть в отдельное подмножество выделены значения уровня глюкозы в окрестности значения 120 мг/дл. Данное значение входит в множество нормального уровня глюкозы, поэтому функция полезности для него находится в районе 0,3. Однако для достижения большей точности этому уровню содержания глюкозы присваивается также значения функции полезности 0,4.

Далее каждому значению функции полезности присваивается нечеткая мера истинности, отражающая степень доверия выбранному значению функции полезности. Таким образом, для значения 120 мг/дл был получен нечеткий входной вектор $X = \langle 0,4(0,8); 0,3(0,9) \rangle$.

База правил, сформированная в соответствии с данными этапами для критерия «Уровень глюкозы», представлена в таблице 1.

Таблица 1.

Функции полезности и нечеткие меры истинности для глюкозы

Величина	Уровень глюкозы, мг/дл										
	30	60	80	100	120	140	160	190	220	250	270
Значение функции полезности	0,30	0,30	0,30	0,30	0,40	0,40	0,50	0,70	0,80	0,90	1,00
Мера истинности	0,90	0,90	0,85	0,85	0,80	0,80	0,70	0,80	0,90	0,80	0,90
Значение функции полезности	0,20	0,20	0,20	0,20	0,30	0,30	0,40	0,60	0,70	0,80	0,90
Мера истинности	0,85	0,85	0,90	0,95	0,90	0,80	0,70	0,90	0,75	0,80	0,90

Для определения показателей пользователя в начале необходимо получить нечеткие меры истинности заключений для каждого значения функции полезности по глюкозе (холестерин рассчитывается по аналогии). Вычисления производятся по формуле 1:

$$R_j = \frac{\sum_{i=1}^n \alpha_i \cdot X_{ij} \cdot Y_j}{\sum_{i=1}^n \alpha_i}, \quad (1)$$

где: α_i – функция принадлежности введенного пользователем параметра i нечеткому множеству;
 n – число множеств;
 X_{ij} – значение меры истинности рассчитываемого параметра в базе правил;
 Y_j – значение функции полезности рассчитываемого параметра в базе правил.
 R_j – нечеткая мера истинности для соответствующего значения функции.

Для определения, к какому из диапазонов относятся показатели пользователя, найдем расстояние a_i до границ диапазона по формуле 2.

$$a_i = \frac{|V - V_i|}{V_2 - V_1}, \quad (2)$$

где: a_i – функция принадлежности введенного пользователем параметра i нечеткому множеству;
 V_i – показатель границы критерия, для которой рассчитывается функция принадлежности;
 V – показатель критерия, для которого рассчитываются функции принадлежности.
 Например, для введенного значения глюкозы 95 мг/дл, расстояние до граничного значения 100 мг/дл составляет 0,75, а до значения 80 мг/дл – 0,25.

Расчет по формуле 1 для значений глюкозы 95 мг/дл:

$$R_1 = \frac{0,75 \cdot 0,3 \cdot 0,85 + 0,25 \cdot 0,3 \cdot 0,85}{0,75 + 0,25} = 0,255.$$

$$R_2 = \frac{0,75 \cdot 0,2 \cdot 0,95 + 0,25 \cdot 0,2 \cdot 0,9}{0,75 + 0,25} = 0,188.$$

Следующим этапом является переход от нечеткой меры к четкой (дефаззификация). Для этого используется формула 3.

$$\varphi = \frac{\sum_{j=1}^n R_j \cdot X_j}{\sum_{j=1}^n R_j}, \quad (3)$$

где: φ – четкое значение меры истинности;
 R_j – значение нечеткой функции полезности;
 X_j – нечеткая мера истинности.

Свертка всех критериев производится по формуле для интегральной функции полезности Саати по формуле 4:

$$W = \sum_{i=1}^n V_i \cdot \varphi(K_i), \quad (4)$$

где: $\varphi(K_i)$ – значение функции полезности по критерию;
 V_i – вес критерия.

Если интегральная функция полезности $W > 0,5$, то делаем вывод, что риск заболевания есть.

Список литературы:

1. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.:Мир, 1976.-165 с.

