mathematics

Article

Leveraging ChatGPT and Long Short-Term Memory in Recommender Algorithm for Self-Management of Cardiovascular Risk Factors

Tatiana V. Afanasieva [1,*], Pavel V. Platov [2], Andrey V. Komolov [1] and Andrey V. Kuzlyakin [1]

[1] Department of Informatics, Plekhanov Russian University of Economics, 36, Stremyanny Lane, Moscow 109992, Russia; komolov_1995@mail.ru (A.V.K.); andrey-kuzliakin@yandex.ru (A.V.K.)
[2] Department of Information Systems, Ulyanovsk State Technical University, 32, Severny Venetz Street, Ulyanovsk 2432027, Russia; chipa.platov@gmail.com
* Correspondence: tv.afanasjeva@gmail.com or afanaseva.tv@rea.ru

Abstract: One of the new trends in the development of recommendation algorithms is the dissemination of their capabilities to support the population in managing their health, in particular cardiovascular health. Cardiovascular diseases (CVDs) affect people in their prime years and remain the main cause of morbidity and mortality worldwide, and their clinical treatment is expensive and time consuming. At the same time, about 80% of them can be prevented, according to the World Federation of Cardiology. The aim of this study is to develop and investigate a knowledge-based recommender algorithm for the self-management of CVD risk factors in adults at home. The proposed algorithm is based on the original user profile, which includes a predictive assessment of the presence of CVD. To obtain a predictive score for CVD presence, AutoML and LSTM models were studied on the Kaggle dataset, and it was shown that the LSTM model, with an accuracy of 0.88, outperformed the AutoML model. The algorithm recommendations generated contain items of three types: targeted, informational, and explanatory. For the first time, large language models, namely ChatGPT-3.5, ChatGPT-4, and ChatGPT-4.o, were leveraged and studied in creating explanations of the recommendations. The experiments show the following: (1) In explaining recommendations, ChatGPT-3.5, ChatGPT-4, and ChatGPT-4.o demonstrate a high accuracy of 71% to 91% and coherence with modern official guidelines of 84% to 92%. (2) The safety properties of ChatGPT-generated explanations estimated by doctors received the highest score of almost 100%. (3) On average, the stability and correctness of the GPT-4.o responses were more acceptable than those of other models for creating explanations. (4) The degree of user satisfaction with the recommendations obtained using the proposed algorithm was 88%, and the rating of the usefulness of the recommendations was 92%.

Keywords: large language model; multidimensional recommendation; predictive assessment; deep machine learning; official guidelines

MSC: 68T50

check for updates

Citation: Afanasieva, T.V.; Platov, P.V.; Komolov, A.V.; Kuzlyakin, A.V. Leveraging ChatGPT and Long Short-Term Memory in Recommender Algorithm for Self-Management of Cardiovascular Risk Factors. *Mathematics* **2024**, *12*, 2582. https://doi.org/10.3390/math12162582

Academic Editors: Silvia N. Schiaffino, Marcelo Gabriel Armentano and Ariel Monteserin

Received: 22 July 2024
Revised: 13 August 2024
Accepted: 17 August 2024
Published: 21 August 2024

1. Introduction

One emerging area of healthcare in which the power of artificial intelligence can have a significant positive impact is disease prevention, particularly cardiovascular diseases, which are the leading cause of mortality worldwide. Cardiovascular diseases (CVDs) are non-communicable diseases that are chronic in nature with a long asymptomatic period, which can lead to cardiovascular events requiring urgent medical care. The relevance of CVD prevention is associated with an increased life expectancy due to the significant contribution of CVDs to the total mortality of the working-age population. According to the 2021 European guidelines on CVD prevention [1], the main direction of CVD prevention is the management of risk factors not only in medical organizations but at home too, as CVD risk monitoring and assessment should be carried out regularly. The main obstacle in CVD risk management in adults is low medical literacy, in particular the lack of information about their risk factors, lack of knowledge of why and how to control them, and what risk factors require medical attention, including emergency care. This results in low adherence to healthy lifestyles and increased costs for individuals and the healthcare system. According to the WHO Global Strategy and Guidelines for Digital Health, the use of digital and information technologies, including artificial intelligence, data analytics, software tools, smartphone apps, and health websites, has proven its potential to improve health, improve health literacy, and prevent disease in individuals as well as in the population as a whole [2].

Today, health recommendation systems (HRSs) applied through AI that are focused on a person's self-management at home and promoting healthy lifestyles are promising for CVD prevention [3]. The main component of HRSs is recommendation algorithms that provide the selected recommended items, in which the selection conditions are determined by a user profile, by the attributes of the recommended items, or based on domain knowledge [4–6]. Utilizing knowledge-based HRSs for CVD prevention will ensure the completeness of the recommendations created and their safety and correctness, which is critical when supporting persons outside the clinical setting. To be effective, this support for adults must be based on official guidelines for managing multifactorial CVD risk, evidence-based medicine, explainable recommendations, and expert knowledge. It should be noted that systems that meet these requirements are not sufficiently represented in the knowledge-based HRS landscape. Currently, the sources of domain knowledge include not only experts and official documents but pre-trained machine learning (ML) models, among them large language models (LLMs). However, these models have not yet been widely used in HRSs for the self-monitoring of CVD risk factors due to insufficient research. To overcome these gaps, this article proposes a knowledge-based recommendation algorithm, named CaRiFaM, to support the management of CVD risk factors in adults using ChatGPT and the LSTM (Long Short-Term Memory) model.

The contributions of this study are described as follows:

(1) The user profile, in addition to the risk factors outlined in the 2021 European guidelines for CVD prevention [1], includes angina symptoms and two predictive assessments. The first predictive assessment estimates the possibility of the presence of CVD in adults with risk factors based on a pre-trained LSTM model. The second one assesses the total risk of CVDs over a 10-year period using SCORE models [7].
(2) Based on user profiles, we propose and study the CaRiFaM algorithm to support the management of CVD risk factors in adults at home. Our proposed algorithm, CaRiFaM, integrates several key enhancements, including an extended user profile and a modified structure of recommendation, leveraging LLM and deep ML models.
(3) The individual recommendations created have a three-dimensional structure and contain recommendation items of different types of meanings, namely targeted, informational, and explanatory. We leveraged the power of ChatGPT to generate explainable recommendations for managing CVD risk factors.

The rest of the article is organized as follows: Section 2 provides a brief overview of the field of HRSs for personal support and emerging trends in recommender systems. A model of multidimensional recommendations based on a formal representation of multiple CVD risk factors is introduced in Section 3. Section 4 describes a multifactorial user CVD risk model. Section 5 presents the proposed recommendation algorithm for supporting the management of CVD risk factors in an adult at home. A preliminary study of ChatGPT for explanatory item creation is described in Section 6. In Section 7, the experimental results are provided. The discussion is presented in Section 8, and the last section contains the conclusions of this study. The study design framework is shown in Figure 1.

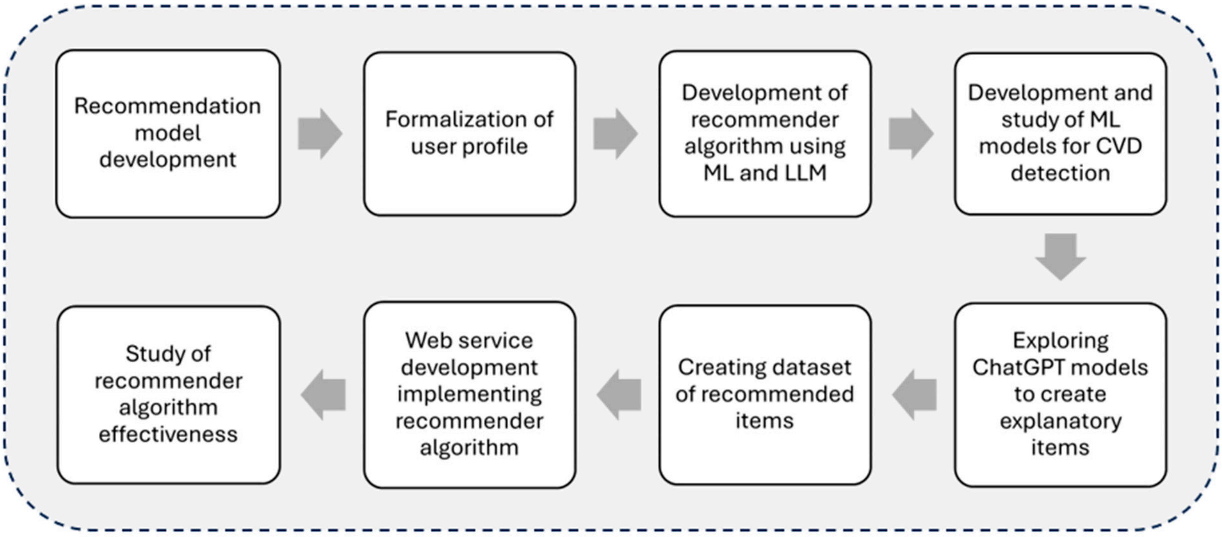

Figure 1. Framework of study design.

2. Literature Review

Among the recommendation systems for health, there are recommendation systems based on collaborative filtering, content filtering, and knowledge-based algorithms [8–10].

Healthcare, as a new area for using the capabilities of recommender systems, has specific features, in particular safety, high-risk decisions, face-to-face communication, and poor formalization. Since healthcare follows the principles of evidence-based medicine and is an area that actively uses professional knowledge, knowledge-based recommendation algorithms are, therefore, the most appropriate when creating HRSs. This fact is confirmed by the number of approaches used in the HRSs presented in prior reviews [3,9,10]. A distinctive feature of a knowledge-based HRS is the usage of expert knowledge, knowledge extracted from official documents, and knowledge about the user's profile [11]. For example, in [12], standard questionnaires for assessing the level of health are used. The authors of another review [10] found a clear trend toward HRSs that provide well-being recommendations, which are aimed at disease prevention, but do not directly intervene with the user's medical status. The results of [13] demonstrate that the personalized lifestyle modifications recommended by the proposed methodology have the potential to effectively reduce the CVD risk. The authors described, based on the CVD risk prediction model, a lifestyle recommendation algorithm incorporating ML and utility function models. The set of 30 challengers was created by experts, and the user profile includes physical, mental, and nutritional aspects. By "challenger", the authors mean a brief recommendation, for example, "Sleep eight hours a day" for people from the group of sleep and wakefulness disorders. The lack of explanation as to why these recommendations should be implemented is a limitation of the proposed recommendation system.

Article [14] notes the limitations of knowledge-based HRSs for the range of lifestyle risk factors used, namely physical activity and diet. They proposed a knowledge-based recommender algorithm for stress management, expanding multiple lifestyle risk factors from a multidimensional viewpoint. The proposed system was designed to be used only at the beginning of the health coaching program, which limits its usefulness as a stand-alone HRS for the long-term support of health behavior change. However, the studies [12,14] were aimed toward persons without any medical conditions, while patients with CVD need supportive advice as well.

The integration of rule-based and natural language processing (NLP) approaches to create knowledge-based recommendations on personalized educational materials for chronic disease patients in China was proposed in [11]. The patient profile was presented using ontology, and then, a user vector was generated using the SWRL rules, which were previously defined. The ontology structure includes patient characteristics in five areas:

chronic diseases, lifestyle and biological risk factors, demographic features, and medication use. NLP technology was applied to obtain text vectors based on the keywords of each educational document and Word2Vec embeddings to map the extracted keywords to the ontology vector space. However, in this study, patients' opinions of usefulness are not provided, and the constructed ontology remains to be further validated using official guidelines for consistency, correctness, and completeness. Unlike other recommender algorithms, the fields of application of HRSs are those in which irrelevant recommendations can cause serious negative consequences [4]. Therefore, it is crucial to ensure that the development of an HRS is guided by the principles of trust and explainability [15,16]. In [10,17], user needs for the explanation of the recommendations received from recommender systems are studied, including knowledge-based recommendations.

Recent advancements in NLP have introduced LLMs that exhibit remarkable capabilities in understanding and generating human-like text [18]. These models use a combination of neural networks and machine learning algorithms to process language in a way that is similar to the way humans do. One of the key advantages of LLMs is their ability to perform a wide range of NLP tasks without the need for task-specific training [19]. Article [20] presents a general framework for integrating LLMs into decision support systems in medical practice, and the authors highlight that LLMs have the potential to change the way healthcare is delivered to provide better and more efficient care. The implementation of LLMs in knowledge-based recommender systems is beginning to play an important role in the context of various tasks, including the creation of explanations in specific recommendation contexts. These models have big potential in cardiology to improve medical diagnosis and decision support as well as provide educational materials for patients [21]. A previous study [22] aimed to evaluate ChatGPT's capacity for ongoing clinical decision support via its performance on standardized clinical vignettes. The results demonstrate that ChatGPT achieved an overall accuracy of 71.7% across all 36 clinical vignettes, with the highest performance of 76.9% in making a final diagnosis and the lowest performance in generating an initial differential diagnosis with an accuracy of 60.3%. Another promising application of ChatGPT in healthcare is helping patients manage their health, as highlighted in [23]. According to [19], ChatGPT can be used to develop systems that can assist with medical education and provide medical information to patients in an easily understandable format. By analyzing patient data and providing personalized recommendations, ChatGPT can help patients manage mental health conditions and improve patient engagement and adherence. In [24], a study of the ability of GPT-3.5 and GPT-4 to perform phenotype concept recognition using The Human Phenotype Ontology was conducted. The experiments included seven prompts of various levels of specificity and showed that GPT-4 surpasses the state-of-the-art performance if the task is constrained to a subset of the target ontology where there is prior knowledge of the terms.

An evaluation of ChatGPT's performance on the United States Medical Licensing Exam (USMLE) was provided in [16]. The results of this study show that ChatGPT performed at a >50% accuracy across all examinations, exceeding 60% in some analyses, and demonstrated a high level of concordance and insight in its explanations.

The authors of [25] created 25 questions for ChatGPT (GPT-3.5 version), addressing fundamental preventive concepts, including risk factor counseling, test results, and medication information, based on guideline-based prevention topics and clinical experience in tertiary care preventive cardiology clinics. Their findings suggest the potential of interactive AI to assist with clinical workflows by augmenting patient education and patient–clinician communication around common CVD prevention queries. However, to the best of our knowledge, no studies to date have concentrated on combining knowledge-based recommendation algorithms with ChatGPT to support individuals at home for the management of the set of CVD risk factors outlined in official guidelines [1].

3. Model of Multidimensional Recommendation

When developing recommendations for the management CVD risk factors at home, it is necessary to consider that they should be intended both for individuals without and with CVD risk factors, in particular having CVD, which may cause acute cardiovascular (CV) events requiring urgent medical attention. An analysis of HRS research [3,8,9,16], the problems in effective strategies for the prevention of CVDs at the individual level [26], and the process of consulting persons on the prevention of CVDs, carried out by preventive medicine doctors, allows us to determine the following requirements for the recommendation content:

1. Usefulness: The recommendations should help adults to manage CVD risk factors more effectively.
2. Safety: The recommendations should be based on the principles of evidence-based medicine and, at the same time, without drug interventions. The recommendations should not lead to an adverse event for users.
3. Completeness: The recommendations should be based on most of the CVD risk factors and their assessment, as set out in the current clinical guidelines for the prevention of CVDs. At the same time, the symptoms of angina pectoris as the main predictor of coronary heart disease should be included in the complex of CVD risk factors.
4. Accuracy: The recommendation items must be correct and free of errors. The text of recommendations should provide information about a person's CVD risk factors only if they are identified. For the persons without CVD risk factors, the text of recommendations should present general information supporting their healthy lifestyle.
5. Explainability: The recommendation items should not only inform about identified CVD risk factors but also explain their impact on the person's CV health.
6. Timeliness: The recommendations should include information about the need and degree of urgency to seek medical help, depending on the assessment of the person's risk of CVD.

Based on the above-defined requirements, a recommendation model will be understood as a text, the content of which combines recommendation items of different dimensions of meaning and contains the individual characteristics of the CV health of each person, grouped by the dimensions. We propose a three-component structure of multidimensional recommendations for supporting the management of CVD risk factors: targeted Es, informational Inf, and explanatory $Expl$ recommendation components:

$$R = \{Es, Inf, Expl\}, \{Inf = Inf(1), Inf(2), \ldots, Inf(N)\}, Expl = \{Expl(1), Expl(2), \ldots, Expl(N)\}, \quad (1)$$

where N is the number of CVD risk factors.

The targeted component Es includes goals at the strategic level and tactical level of CVD risk management. The Es component answers the question "What is recommended to be done?" and contains information about the degree of urgency and type of medical care. The informational component Inf answers the question "What are your CVD risk factors and how to manage them?" and contains a list of individual CVD risk factors. The explanatory component $Expl$ answers the question "Why is it recommended to manage this particular risk factor for CVD?".

Below is an example of multidimensional recommendations for a 49-year-old woman with an elevated systolic blood pressure of 160/90 mmHg and habit of smoking tobacco:

Es: You are recommended to plan consultation with a primary care physician this year as some CVD risk factors have been identified.

Inf: Your CVD risk factors are

* Elevated systolic blood pressure of 160/90 mmHg. Target values for your age no more than 130 mmHg. Measure your blood pressure daily. The correct way to measure blood pressure is to sit at rest three times and determine the average reading. To reduce systolic blood pressure gradually increase your physical activity. It has been proven that varied physical activity provides a moderate reduction in blood pressure. Limit the total

amount of salt consumed in any form to 5 g per day (1 teaspoon), give preference to the Mediterranean diet.

** Smoking. Imagine what would happen if you didn't smoke at all. Would you be reaching for a cigarette right now? Probably not. Make a decision to quit smoking and stick to it. Set a date in advance when you will quit. Give up smoking for one day, tomorrow—another one, the day after tomorrow—another one, etc. Find a partner with whom you will quit smoking together. Try to help someone else quit smoking.

Expl: Do you know that

* Systolic blood pressure is the pressure in the arteries at the moment when your heart contracts and pushes blood into the arteries. Increased systolic pressure can cause increased heart function and damage to blood vessels. Over time, this can lead to heart attacks, strokes and other CVDs. It is important to pay attention to the level of systolic pressure and consult with your doctor about appropriate methods of its control and treatment.

** Smoking is harmful because it damages blood vessels and the heart, increasing the risk of cardiovascular disease. Nicotine and other chemicals in tobacco contribute to the formation of plaques in the arteries, which can lead to their narrowing and blocking. This increases the likelihood of heart attack and stroke. Quitting smoking improves vascular health and reduces the risk of cardiovascular events. Therefore, quitting smoking is an important measure to preserve the health of the heart and blood vessels.

The proposed multidimensional recommendation model (1) has a structure that is as close as possible to the structure of clinical recommendations for the prevention of CVD [1] regarding recommendations for management health indicators, CVD risk factors, and goals without recommendations for treatment. The contents of the recommendation items Es and Inf were manually extracted from official guidelines [1] and added by experts with knowledge of counseling on the management of CVD risk factors, while the set of explanatory items $Expl$ was generated by ChatGPT, since this model is able to generate human-like text [18], including the creation of explanations, in particular in cardiology [21].

The basic concepts used for generating multidimensional recommendations for CVD risk management are depicted in Figure 2.

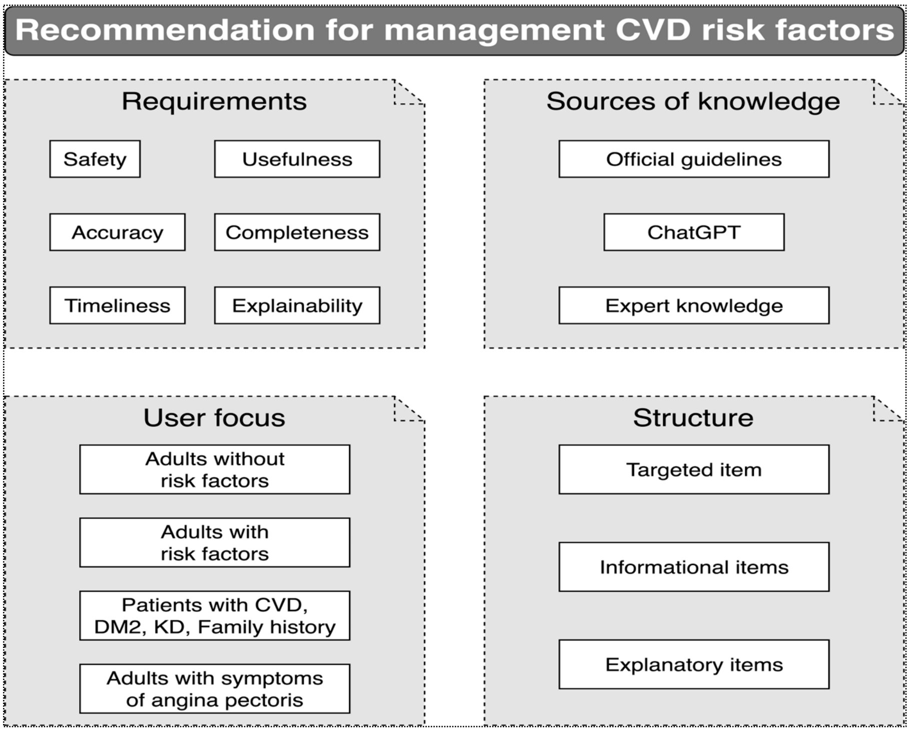

Figure 2. Basic concepts of proposed recommendation model for CVD risk management.

4. Multifactorial User CVD Risk Model

As a result of the analysis of the guidelines for CVD prevention, the following conclusion was made: the user profile should be based on a modern model of the CVD risk factors outlined in official recommendations; the 2021 European guidelines [1] were selected in this study.

By "CVD risk factor", we mean a characteristic of the user's condition that may be the cause of the development of CVD or an acute CVD event, for example, myocardial infarction. Therefore, the model should include modifiable, non-modifiable, and specific CVD risk factors, as described in [1], as well as an additional risk factor in the form of angina pectoris symptoms, leading to a significant deterioration in well-being. Taking into consideration the multifactorial structure of CVD risk factors, it is important to determine the formal model before defining a user profile on this basis. Let us define a multifactorial model of CVD risk factors in the form of the following expressions:

$$Z = \{S, NonModF, Biol, Beh, Predict\},\ NonModF = \{CVD, KD, DM2, GEN\},$$
$$Biol = \{OBS, CH, NHL, SBP, GL\},\ Beh = \{Sm, Fa, Diet\}.$$

The designation of the components of this model is presented in Table 1. Note that the groups of factors S, $NonModF$, $Biol$, and Beh describe the current state of CV health, and the factors of the $Predict$ group represent prognostic estimates of the overall risk of CVD over a 10-year perspective (Total) and the possibility of having CVD (Pre-CVD). We propose the user profile as a binary vector factor of descriptive and predictive assessments of individual CVD risk, for which the initial values of each component are set to 0:

$$Factor = \{f(i),\ f(i) \in \{0, 1\}, i = 1, 2, \ldots, 15\}. \quad (2)$$

Table 1. Designation of CVD risk factors. If there is a risk factor, its value will be set to "1", or else "0".

Factor No.	CVD Risk Factor Designation and Description
Type 1: Symptoms (S)	
1	S—a sign of symptoms of angina pectoris with a significant deterioration in health
Type 2: Non-modifiable factors (NonModF)	
2	CVD—a sign that the person has documented one of the CVDs
3	KD—a sign of chronic kidney disease
4	DM2—a sign of type 2 diabetes
5	GEN—a sign of family history of early CV diseases in close relatives: in men under the age of 55 years and in women under 65 years
Type 3: Modifiable biological factors (Biol)	
6	OBS—a sign of obesity if the body mass index is more than 30 kg/m
7	CH—a sign of high cholesterol; the threshold is 5 mmol/L
8	NHL—a sign of high levels of non-high-density lipoprotein cholesterol; the threshold is 4 mmol/L
9	SBP—a sign of raised systolic blood pressure; the threshold is 130/90 mmHg
10	GL—a sign of raised glucose levels in the blood serum; the threshold is 7 mmol/L
Type 4: Modifiable behavioral factors of lifestyle (Beh)	
11	Sm—a sign of tobacco smoking
12	Fa—a sign of insufficient physical activity, no more than 2 h of household activity or moderate-intensity aerobic physical activity per week
13	Diet—a sign of unhealthy food abuse
Type 5: Predictive estimates of CVD risk (Predict)	
14	Total—a sign of moderate and higher total CVD risk in 10-year perspective, based on clinical model
15	Pre-CVD—a sign of the possibility of existing CVD, based on machine learning

Here, the first 13 components are descriptive, and the last ones are predictive, and i denotes CVD risk factor number, as presented in Table 1.

To determine the user profile, we use the set of CV health indicators in the form of a vector:

$$\mathrm{X} = \{x_j, x_j \in \mathbb{R}, x_j \geq 0, j = 1, 2 \ldots 17\}.$$

Table 2 shows the set of CV health indicators, grouped by the types of CVD risk factors.

Table 2. Indicators of a person's CV health. Here, 1 denotes "Yes", 0 denotes "No".

X	Indicator	X	Indicator
X(1)	Gender: male (1)/female (0)	X(10)	Total cholesterol (mmol/L)
X(2)	Age (years)	X(11)	Non-high-density lipoprotein cholesterol (mmol/L)
X(3)	Height (cm)	X(12)	Systolic blood pressure (mmHg)
X(4)	Weight (kg)	X(13)	Glucose level (mmol/L)
X(5)	Family history of CVD (1/0)	X(14)	Physical inactivity (1/0)
X(6)	Presence of CVD (1/0)	X(15)	Smoking (1/0)
X(7)	Chronic kidney disease (1/0)	X(16)	Unhealthy diet (1/0))
X(8)	History of CV events (1/0)	X(17)	Symptoms of angina pectoris, with a significant deterioration in health (1/0)
X(9)	Type 2 diabetes mellitus (1/0)		

5. Recommender Algorithm to Support Self-Management of CVD Risk Factors

Let us present the statement of the problem to be solved using the CaRiFaM algorithm. We assume that we are given a set of CV health indicators X for a person Y (see Table 2) and the sets of items $Es^* \subset R^*$, $Inf^* \subset R^*$, and $Expl^* \subset R^*$ corresponding to the multidimensional model of recommendation R. It is required for person Y to generate recommendation R^{x}, that is, to construct a mapping:

$$\mathrm{X} \times R^* \to \mathrm{R}^{\mathrm{x}},$$

where $\times$ is the sign of the Cartesian product $\mathrm{R}^X = \{Es^X, Inf^X, Expl^X\}$, $Es^X \subseteq Es^*$, $Inf^X \subseteq Inf^*$, $Expl^X \subseteq Expl^*$.

The CaRiFaM algorithm for each person includes the following stages:

1. Creating the user profile *Factor* using multifactorial CVD risk model Z:

$$X \times \mathrm{Z} \to Factor.$$

2. Generating the recommendations:

$$Factor \times R^* \to \mathrm{R}^{\mathrm{x}}.$$

5.1. Creating User Profile Factor

Determining the 15 components of a user profile (2) includes the following steps:

Step 1. Defining descriptive assessments of individual CVD risk through an analysis of CV health indicators X using the multifactorial model of CVD risk Z:

For the assessment of the biological CVD risk factor obesity $f(6)$, the body mass index BMI is calculated:

$$BMI = \begin{cases} 10^4 \cdot \frac{x(4)}{x(3)^2}, \ if\ x(3) \neq 0, \\ \quad 25, \ if\ x(3) = 0, \end{cases}$$

and then, the $f(6)$ component is determined using the expression corresponding to [1]:

$$f(6) = \begin{cases} 0, \ if\ BMI < 30, \\ 1\ , if\ BMI \geq 30. \end{cases}$$

Then, the binary health indicators are assigned to the corresponding binary components of the *Factor* vector:

$$f(1) = x(17), f(2) = x(6) \vee x(8), f(3) = x(7), f(4) = x(9), f(5) = x(5),$$
$$f(11) = x(15), f(12) = x(14), f(13) = x(16).$$

For biological CVD risk factors, the components $f(7)$, $f(8)$, $f(9)$, and $f(10)$ of the *Factor* vector will be determined using the rules for comparing the values x_q with the thresholds p_j specified in the clinical guidelines [1] and presented in Table 1:

$$f(j) = \begin{cases} 0,\ if x_q \le p_j, \\ 1\ , if x_q > p_j, \end{cases} \quad j = 7, 8, 9, 10; q = 10, 11, 12, 13.$$

Step 2. Calculating the predictive component of the *Factor's* vector: an assessment of the total CVD risk over a 10-year period using the SCORE/SCORE2-OP models [1,7] and the detection of possible CVD using a classification model based on ML algorithms.

The SCORE (Systematic Coronary Risk Evaluation) model is designed to estimate (probably predict) the total risk of fatal CVDs over a 10-year perspective, outlined in the guidelines for CVD prevention [1]. To assess the total CVD risk of a person using SCORE, the following are used: a person's CV health profile (Table 2): $x(1)$—gender, $x(2)$—age, $x(10)$—total cholesterol, $x(12)$—systolic blood pressure, $x(15)$—sign of smoking. Subsequently, the SCORE2/SCORE2-OP model presented an expanded, updated prediction model to estimate the 10-year risk of fatal and nonfatal CVDs. Unlike the SCORE model, the SCORE2/SCORE2-OP model uses a personal profile indicator of $x(11)$ (non-high-density lipoprotein cholesterol) instead of $x(10)$. The updated SCORE2/SCORE2-OP model is applicable to individuals without prior CVD or diabetes aged 40–89 years in Europe. The probability of the person's total CVD risk in a ten-year perspective is calculated using Rule 1, in which SCO denotes the SCORE2/SCORE2-OP algorithm [7]:

$$Rule\ 1 : IF\ f(1) \wedge f(2) \wedge f(4) \wedge f(5) = 0\ THEN\ Total = SCO(x(1), x(2), x(11), x(12), x(15)).$$

The SCORE2/SCORE2-OP algorithm, as an external function SCO, calculates the predictive assessment Total, that is, $f(14)$, following the guidelines [1]:

$$f(14) = \begin{cases} 0,\ Total \le 2.5\%, \\ 1,\ Total > 2.5\%. \end{cases}$$

We assume that often a person who already has CVD does not know about it and, therefore, assign 0 to the CVD indicator $x(6)$. However, even the possibility of CVD is critical to make proper recommendations for the prevention of CV events. Therefore, for a person who indicates $x(6) = 0$, we include one more predictive estimate (Pre-CVD) based on the ML model. For CVD detection, pre-trained AutoML and deep ML will be tested in Section 7. To calculate the predictive estimate $f(15)$, Rule 2 was developed for detecting CVD:

$$Rule\ 2 : IF\ (\bigwedge_{i=6}^{10} f(i) = 1) \wedge (f(2) = 0\ ) THEN\ f(15) = ML(S_x),$$
$$S_x = \{x(1), x(2), x(3), x(4), x(7), x(10), x(12), x(13), x(14), x(15)\}.$$

The use of not only descriptive but also predictive estimates is a key advantage of the proposed user profile for creating recommendations with greater information capacity.

5.2. Generating Individual Recommendations

The content of recommendation items of the *Es* type, which sets a target for CVD prevention, is determined by constraining rules, presented in Table 3, where the first column denotes the number of CVD risk levels. Constraint-based rules were formalized using official guideline knowledge, expert knowledge, and the user profile *Factor*. According to the above rules presented in Table 3, we consider nine levels of individual CVD risk.

Table 3. Constraint-based rules for selecting targeted item of individual recommendations.

Level	Constraint-Based Rules	Targeted Item Es^*
1	$\bigwedge_{i=1}^{14} f(i) = 0$	CVD risk factors have not been identified. Continue to follow a healthy lifestyle
2	$(f(1) = 0) \wedge (\bigwedge_{i=11}^{13} f(i) = 1) \wedge (\bigwedge_{i=6}^{10} f(i) = 0)$	Behavioral risk factors for CVD have been identified. Follow healthy lifestyle and **regularly** control these factors.
3	$(f(1) = 0) \wedge (f(2) = 1) \wedge (\bigwedge_{i=6}^{10} f(i) = 0)$	You have reported having CVD, which is a high-level risk factor. To prevent cardiovascular events, it is recommended that you consult your doctor **regularly** and adhere to his recommendations.
4	$(f(1) = 0) \wedge (f(14) = 1) \wedge (Total < 5) \vee (\bigwedge_{i=6}^{10} f(i) = 1) \vee f(15) = 1$	You are recommended to plan consultation with a primary care physician **this year** as some CVD risk factors have been identified.
5	$(f(1) = 0) \wedge (f(14) = 1) \wedge (5 \leq Total < 7.5) \wedge (\bigwedge_{i=6}^{10} f(i) = 1)$	You are advised to consult with a primary care physician **within six months** because some of your risk factors for heart disease need professional help.
6	$(f(1) = 0) \wedge (\bigwedge_{i=6}^{10} f(i) = 1)) \wedge (\bigwedge_{i=3}^{5} f(i) = 1)$	You have reported a health status that is a specific risk factor for CVD, moreover some additional CVD risk factors have been identified. Therefore, it is recommended to consult with a physician within **next month.**
7	$(f(1) = 0) \wedge (f(14) = 1) \wedge (Total \geq 7.5) \wedge (\bigwedge_{i=6}^{10} f(i) = 1)$	Based on your CVD risk level, you are **strongly advised** to consult with physician or cardiologist. Do it **within a month**—a good decision.
8	$(f(1) = 0) \wedge (f(2) = 1) \wedge (\bigwedge_{i=6}^{10} f(i) = 1)$	You have reported the presence of CVD and some indicators are out of the norm. To clarify your treatment plan, you **strongly advised** to consult with your doctor **this month**.
9	$f(1) = 1$	Symptoms of angina pectoris have been identified. To prevent CV events, it is **extremely important** for you to consult a primary care physician or cardiologist as soon as possible **without delay**.

The first level is associated with persons without CVD risk factors, and the remaining ones are for persons with various combinations of CVD risk factors and clinical conditions. Note that the higher the level number, the higher the risk of CVD in a person and the more strictly the targeted item of the recommendation is formulated. Moreover, the recommended urgency of seeking medical care rises with an increase in the number of CVD risk levels. If more than one level of CVD risk for a person is determined, the targeted item with the maximum level number will be selected. For obtaining the informational Inf and explanatory $Expl$ components of individual recommendations, the rules corresponding to non-zero values of the vector *Factor* (see Expression 2) are used:

$$Rule\ 3i : IF\ f(i) = 1\ THEN\ Inf(i) \cup Expl(i), i = 1, 2, \ldots, 14.$$

Here, the sign "∪" denotes the operation of concatenation of the items, and i designates the risk factor number in Table 1.

Figure 3 shows a framework of the CaRiFaM algorithm.

To obtain initial data about a person's CV health, a questionnaire was developed, the questions of which correspond to the CVD health indicators presented in Table 2. To assess the symptoms of angina pectoris, the Rose questionnaire was used [27].

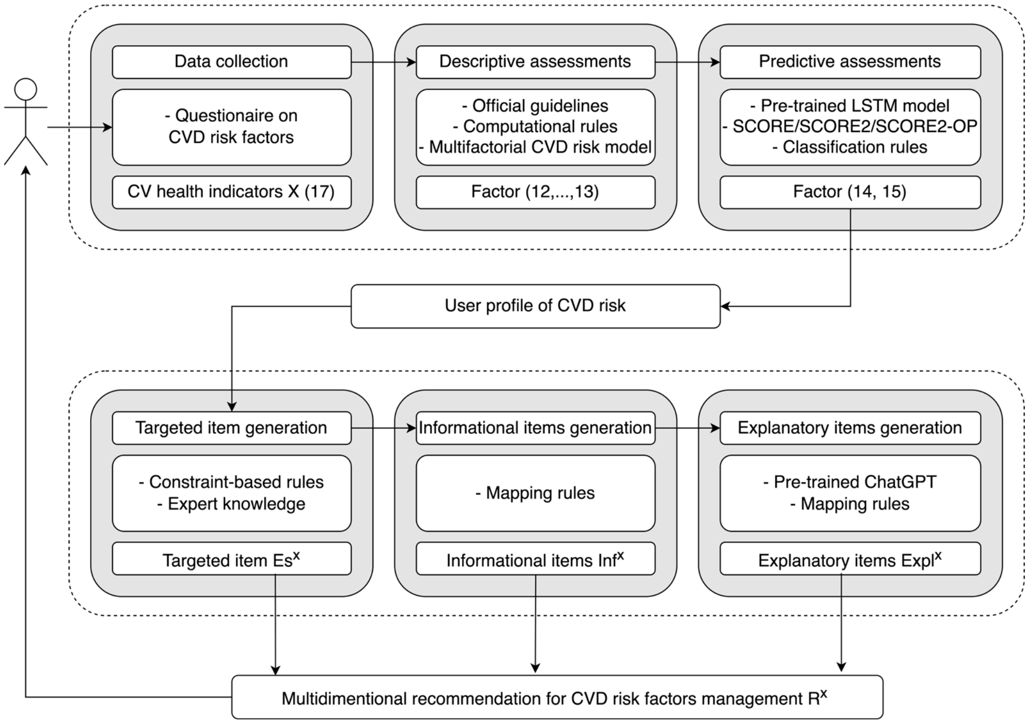

Figure 3. Basic steps of the proposed algorithm.

6. Preliminary Analysis of ChatGPT Explanations for CVD Risk Factors

Since the explanations in the guidelines [1] about why CVD risk factors should be self-managed are not simple enough for individuals to understand, LLM models promise to fill this gap. An LLM typically uses over a billion parameters and is pre-trained on large volumes of text corpora containing up to trillions of tokens obtained from various Internet sources. Therefore, we decided to explore the capabilities of ChatGPT, developed by OpenAI [28], to generate explanatory items in the proposed algorithm. Three versions of the GPT model were explored, namely GPT-3.5, GPT-4, and GPT-4o. The GPT-3.5 model provides more accurate and relevant answers compared to the previous versions, namely GPT-3, due to the large number of parameters and improved natural language processing skills. Moreover GPT-3.5 was trained using documents published in PubMed [19]. Not only is the GPT-4 model a further development compared to GPT-3.5, but it appears to be a significant step forward in the direction of improved text generation and context understanding [29].

Announced in May 2024, the GPT-4o multimodal model combines the high performance of GPT-4 with the ability to work not only with text but also with other types of information.

We turned to ChatGPT with a set of prompts to obtain explanations for why identified CVD risk factors should be controlled by humans. Since the purpose of the recommendations is to support individuals in the self-management of CVD risk factors, the GPT responses received are considered as explanations for the recommendations provided. The following general form of prompts to obtain ChatGPT responses was used:

$$Response = ChatGPT(task, params, constraints, [Optional : viewPoint]), \tag{3}$$

where *task* is a text expressing the content of what we want to receive, *constraints* denote some restrictions for the response, and *params* define the parameters that must be taken into account to clarify the *task*, namely the list of CVD risk factors of the person.

The optional *viewPoint* component defines some scopes for expressing the results of the *task* or some of an analogy. For example, the explanation of why a person should manage a risk factor could be created using the analogy of the maintenance of car functionality, from the point of view of a long active life, or from the point of view of reducing the financial costs of future treatment. In our study, an optional component in addition to the above may include the point of view of a traveler and a gardener. Component *viewPoint* is used to diversify the explanations in repeated recommendations for the same risk factor. In Table 4, the prompts to ChatGPT and their responses are shown. The first sentence of the prompts combines the task text and person's CVD risk factors, while restrictions on the reasoning for explanatory items are indicated in the last sentence of the prompt. The prompt parameters *params* are a person's CVD risk factors; they are highlighted in bold. The second and third prompts contain an optional *viewPoint* parameter, namely an analogy with a garden and a car, respectively. A review of the responses generated by ChatGPT models showed the following: When responding to a prompt, ChatGPT creates a plausible and stylistically correct explanation of the specified risk factor within the specified restrictions. If the prompt contains a list of CVD risk factors, the ChatGPT can create an explanation of them, united by one metaphor, as follows in examples 2 and 3 depicted in the Table 4:

Table 4. Examples of three prompts and the explanatory item lists received as ChatGPT responses.

Prompts	ChatGPT Responses
Explain, please, why human with diagnosed **CVD** should manage **cholesterol**? The explanation must be understandable to the person, include only person CVD risk factors and contain no more 4 propositions.	ChatGPT-4 response: Managing cholesterol is crucial for individuals diagnosed with cardiovascular disease (CVD) because cholesterol plays a significant role in the development of atherosclerosis, which is a common underlying cause of CVD. **Plaque Formation.** High levels of low-density lipoprotein (LDL) cholesterol can lead to the buildup of plaque in the arteries. This plaque narrows the arteries and can reduce or block blood flow to the heart and other organs. **Risk of Heart Attack and Stroke.** Plaque buildup increases the risk of heart attacks and strokes. If a plaque ruptures, it can form a clot that blocks the flow of blood. **Reducing Inflammation.** High cholesterol can contribute to inflammation, which is associated with a higher risk of heart disease. Managing cholesterol levels can help reduce this inflammation. **Improving Heart Health.** Lowering high cholesterol, especially LDL cholesterol, can slow, reduce, or even stop the buildup of plaque. This helps maintain healthy blood vessels and improves overall heart health.
Explain, please, why it is recommender to manage CVD risk factors? Use analogy of a garden. The explanation must be understandable to the person, include only person CVD risk factors and contain no more 4 propositions.	ChatGPT-3.5 response: **Think of your heart as a garden** that you want to keep flourishing with vibrant flowers and plants. The risk factors for CVD—like high blood pressure, high cholesterol, smoking, obesity, and diabetes—are akin to weeds, pests, and harsh weather conditions that can harm your garden. **Just as you would pull out weeds**, protect your plants from pests, and prepare for changes in weather, managing your CVD risk factors involves adopting a healthy diet, exercising regularly, and following medical advice to keep your heart healthy. **Neglecting these risk factors** is like letting weeds grow wild, which can choke out the healthy plants. Over time, this can lead to a garden—or in this case, a heart—that isn't as healthy or vibrant as it could be. Managing these risks is like nurturing your garden, ensuring that it remains a beautiful and serene place to enjoy for years to come.
Explain, please, why human need to control her/his CVD risk factors such as **high blood pressure**—160/90 mmHg, **physical inactivity, smoking, unhealthy diet**. Use analogy of a car. The explanation must be understandable to the person, include only person CV risk factors and contain no more 4 propositions.	ChatGPT-4.o response: **High Blood Pressure** is like putting too much air in a balloon—it strains the walls and can cause them to burst. In the body, this can lead to damage in the arteries, heart, and other organs. It's often called the "silent killer" because it doesn't have symptoms but can lead to serious problems like heart attacks and stroke. Keeping blood pressure lower reduces the strain on your heart and the risk of heart disease. **Physical Inactivity**. Being inactive is like leaving a car unused for a long time; eventually, it won't run as well. Regular physical activity keeps your heart and body running smoothly and can help lower high blood pressure, manage weight, and reduce stress on your heart. It can also improve your mood and energy levels. **Smoking** damages your blood vessels and can make them narrow, making it harder for blood to get around your body. Think of it as a traffic jam in your arteries. Stopping smoking can help to clear the traffic and reduce your risk of heart disease, as well as improve your overall health. **Unhealthy Diet.** Eating foods high in fat, salt, and sugar is like using low-quality fuel in a car—it won't run properly and will break down sooner. A healthy diet is like high-quality fuel; it makes sure everything runs smoothly and reduces the risk of blockages in your arteries. It's good for your whole body, including your heart.

The explanations received by the ChatGPT models seem to be acceptable for creating explanatory items of the recommendations in the proposed recommender algorithm. To compare the capabilities of different ChatGPT versions and choose the most effective one from ChatGPT-3.5, ChatGPT-4, and ChatGPT-4o, in the next section, we will examine ChatGPT responses from different points of view to make sure that the GPT models are applicable to generate explanatory items in the proposed recommendation algorithm.

7. Experimental Results

7.1. Evaluation of CVD Prediction Using ML Models

In this article, firstly, we decided to use AutoML, which provides an automated process for building and optimizing machine learning models and the AutoGluon-Tabular framework [30], which is focused on processing tabular data. Key features of AutoGluon-Tabular include the stability of the classification results on heterogeneous tabular data, the use of modern ML models including artificial neural networks, and the automatic assembly of these models based on multi-layer stacking and re-bagging. To test AutoML, we use the Cardiovascular Disease dataset on Kaggle [31], which contains 70,000 records with attributes, the range of which is similar to those used in our study: age, gender, blood pressure, cholesterol level, glucose level, smoking, physical activity, height, weight, and indicator of CVD. For training the AutoML model, 80% of the data were used, and the remaining 20% were for testing.

The following data preprocessing was conducted:

Data cleaning: This stage involved removing missing values and anomalies. Missing values were replaced with mean values for the corresponding variables, and anomalies such as incorrect or extreme values were identified and removed.

Data normalization: All numerical variables were scaled to a single scale using standardization, which improved the model's training efficiency. Standardization involved subtracting the mean and dividing the value by the standard deviation for each continuous variable.

In addition to the data preprocessing stages described above, clustering of the dataset into two clusters was performed, numbers of which, as well as the calculated body mass index, were added to the analyzed attributes.

As a result, after 300 epochs of training, the AutoML model achieved an accuracy of 0.72, CVD detection precision of 0.75, and F1-score of 0.71 on the test part of the dataset. Although AutoML provides minimal manual effort to obtain an optimal prediction and achieved an accuracy of 0.72, consistent with the results of other authors [32–34], it was decided to apply a deep ML model, namely LSTM, to detect CVD on the same dataset [31].

The LSTM is a type of Recurrent Neural Network model that involves LSTM neurons containing the memory cell states (c_t) and the hidden states (h_t), which are controlled by three gates: the input gate (i), the forget gate (f), and the output gate (o). These gates decide what information to add to, remove from, and output. The computational process in LSTM neurons can be mathematically expressed as

$$
\begin{gathered}
f = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
i = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
c_t = (c_{t-1} \odot f) + (i \odot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)), \\
o = \sigma\big(W_o \cdot [h_{t-1}, x_t] + b_o\big), \\
h_t = \tanh(c_t) \odot o,
\end{gathered}
$$

where x_t and h_t denote the input and output of LSTM neuron, respectively; σ and tanh are sigmoid and *tanh* activation functions; $[h_{t-1}, x_t]$ expresses the concatenation of h_{t-1} and x_t; $\odot$ denotes element-wise multiplication; W_f, W_i, W_c, and W_o are the weight matrices; and b_f, b_i, b_c, and b_o designate the biases.

The LSTM architecture contained three layers: the first LSTM layer with eight neurons, the second LSTM layer with four neurons, and a fully connected layer with a softmax activation function. Binary cross_entropy was chosen as the loss function, and the Adam optimizer was used for optimization. The data were split into training and testing sets in a ratio of 80:20, and the LSTM model was trained on 300 epochs. The result was a more exact model with an accuracy of 0.88, CVD detection precision of 0.9, and F1-score of 0.87. Thus, the pre-trained LSTM model was used in the proposed recommender algorithm for CVD detection in adults with risk factors.

7.2. Exploring the Effectiveness of ChatGPT for the Explanation of CVD Risk Factors

In order to use a specific ChatGPT version in a recommendation algorithm to generate explanations for CVD risk factors, a study of the responses of three versions of ChatGPT from the sets of ChatGPT-3.5, ChatGPT-4, and ChatGPT-4o was carry out. In the following, for brevity, the analyzed versions will be designated as GPT-3.5, GPT-4, and GPT-4o, since their models are the core of ChatGPT.

We use single-turn tasks for ChatGPT models in the prompting structure as presented in (3) and zero-shot setting, a technique in which we prompt GPT models without any examples [35]. As a result, 56 explanations across 14 CVD risk factors were generated by GPT-3.5, GPT-4, and GPT-4o in different sessions. In this study, we assessed the correctness of the GPTs' responses in two ways: based on metrics and based on expert knowledge of physicians. The following metrics for assessing the GPT explanations of CVD risk factors were used: Answer Semantic Similarity, accuracy, and Diversity.

Answer Semantic Similarity (ASS) pertains to the assessment of the semantic resemblance between the generated answer and the ground truth. This metric calculates the degree of similarity in texts of GPT explanations to the texts of the explanations given by an expert as a ground truth. A higher ASS score signifies a better alignment between the generated answer and the ground truth:

$$ASS\_GPTj = \frac{1}{N}\sum_{i=1}^{N} Sim(j,i), ASS\_GPTj \in [0,100],$$
$$Sim(j,i) = 100\% \cdot Cosine\_similarity\left(V_{GPTj}(i), V_E(i)\right), Sim(j,i) \in [0,100],$$

Here, the version of GPT is designated as j ∈ {3.5,4,4o}, the number of GPT responses to prompts is denoted by N, i = 1,2,. . .,N, and the *Cosine_similarity* function calculates the similarity between the GPT answer and the expert answer, which have been converted into term-frequency vectors, namely $V_{GPTj}(i)$ and $V_E(i)$, respectively [36].

By accuracy (ACC), we mean a measure that is share of responses having Cosine Similarity $V_{GPTj}(i)$ with $V_E(i)$ above a given threshold:

$$\begin{gathered} ACC\_GPTj = \tfrac{100\%}{N}\textstyle\sum_{i=1}^{N} f(Sim(j,i)), Acc\_GPTj \in [0,100], \\ f(Sim(j,i)) = \begin{cases} 0, & Sim(j,i) < Threshold, \\ 1, & Sim(j,i) \geq Threshold \end{cases} \end{gathered} \tag{4}$$

To calculate the Diversity (DIV) of the GPT explanations of CVD risk factors, the Jaccard similarity index was used to compare words in the GPT response $R_{GPTj}(i)$ and the expert response $R_E(i)$ [37]. The Jaccard index expresses similarity between two texts as a ratio of intersection of their words to their union. The higher the Jaccard index, the more similar the two texts. In contrast, the Diversity coefficient ranges from 0 to 1, where 1 denotes that there is no overlap between the words in the texts, and 0 indicates that the texts are identical:

$$\begin{gathered} DIV\_GPTj = \tfrac{100\%}{N}\textstyle\sum_{i=1}^{N} f(Diversity(j,i)), DIV\_GPTj \in [0,100], \\ Diversity(j,i) = 100\% \cdot \left(1 - Jaccard\left(R_{GPTj}(i), R_E(i)\right)\right), \\ f(Diversity(j,i)) = \begin{cases} 0, & Diversity(j,i) < Threshold, \\ 1, & Diversity(j,i) \geq Threshold \end{cases} \end{gathered} \tag{5}$$

Since the GPT answers may be plausible and similar to the ground truth, nevertheless, the opinion of a doctor is necessary to confirm the safety, coherence, and accuracy of the given GPT explanations of CVD risk factors.

The framework of the ChatGPT study for the explanation of CVD risk factors in the proposed recommender algorithm is presented in the Figure 4.

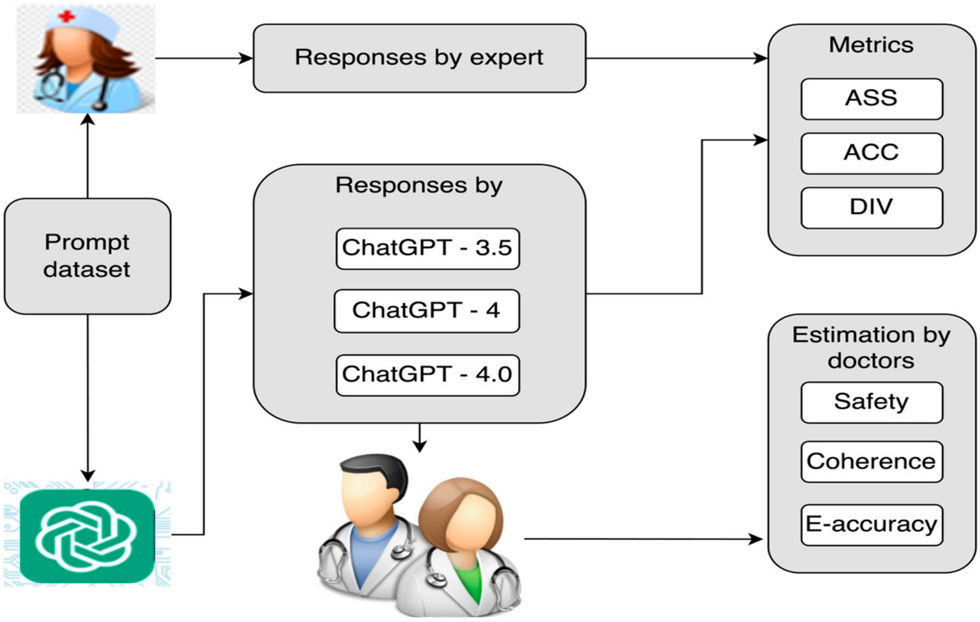

Figure 4. The scheme of the ChatGPT study for the explanation of recommendations for CVD risk factor management.

The method of comparison of GPT models according to Table 5 is based on the formal metrics introduced: ASS_GPTj, ACC_GPTj, and DIV_GPTj. These metrics are denoted in Table 5 as ASS, ACC, and DIV, respectively, and are calculated for GPT-3.5, GPT-4, and GPT-4o. Table 5 presents the ASS, ACC, and DIV metrics of average similarity between GPT and expert explanations of cardiovascular disease risk, which allow us to estimate the average performance of the GPT models in explaining CVD risk factors. As shown in Table 5, the GPT models showed approximately the same high semantic similarity according to the ASS metric and high accuracy according to the ACC metric, ranging from 71% for GPT-4 to 86% for the GPT-3.5 and GPT-4o models. The Diversity of explanations given by the GPT models compared to expert explanations for CVD risk factors is moderate, 50% to 57%, which follows from the DIV metric. Thus, the data presented in the Table 5 demonstrate the GPT models' effectiveness in generating high-quality explanations of CVD risk factors. A comparison of the GPT versions using the $Cosine\_similarity$ function is presented in Figure 5. The bar chart shows the similarity scores between GPT explanations and expert explanations of CVD risk factors grouped by factor number.

Table 5. Comparison of models' explanations of CVD risk factors with expert explanations (ground truth).

	Answer Semantic Similarity (ASS)%	Accuracy (ACC)%	Diversity (DIV)%
GPT-3.5	67.2	86	57
GPT-4	66.6	71	50
GPT-4o	66.9	86	57

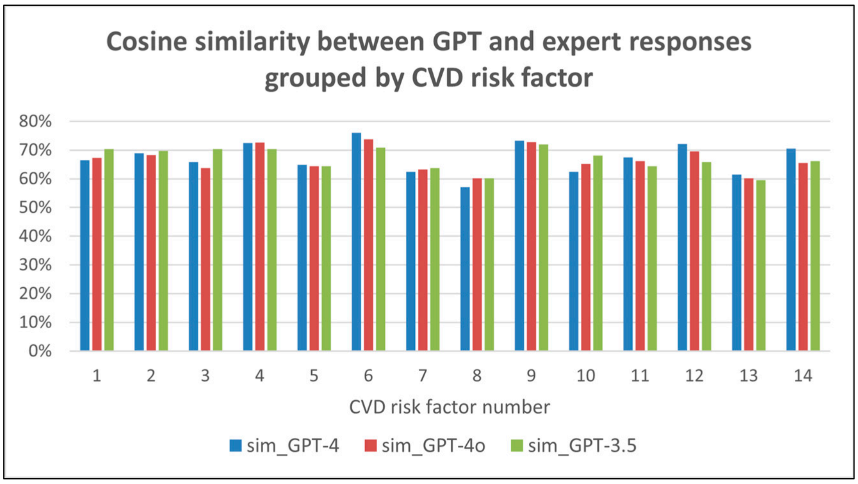

Figure 5. Comparison of GPT explanations with expert explanations (ground truth) by Cosine Similarity.

As one can see, the Cosine Similarity scores are relatively high, varying from 57% to 76%, indicating a good similarity. The highest value for Cosine Similarity is observed for the explanation of the sixth risk factor, which refers to obesity, while the lowest values for Cosine Similarity score were obtained for GPT explanations of the eighth risk factor, which is raised non-high-density lipoprotein.

To assess the GPT explanations according to the ACC metric, it is necessary to determine the threshold value, given in Formula (4). It is obvious that the higher the threshold, the lower the accuracy and vice versa. An illustration of the decline in the explainability accuracy for each GPT model, measured by ACC when the threshold increased, is shown in Figure 6. We found that only with a threshold equal to 61%, the accuracy of GPT-4 exceeds the accuracies of GPT-3.5 and GPT-4o. Moreover, for six threshold values, the accuracy of GPT-4 is lower than that of the compared models. As a result, using the threshold 63% for the ACC metric, we found the explanation accuracy of the GPT-3.5 and GPT-4o models is 86% versus 71% for the GPT-4 model.

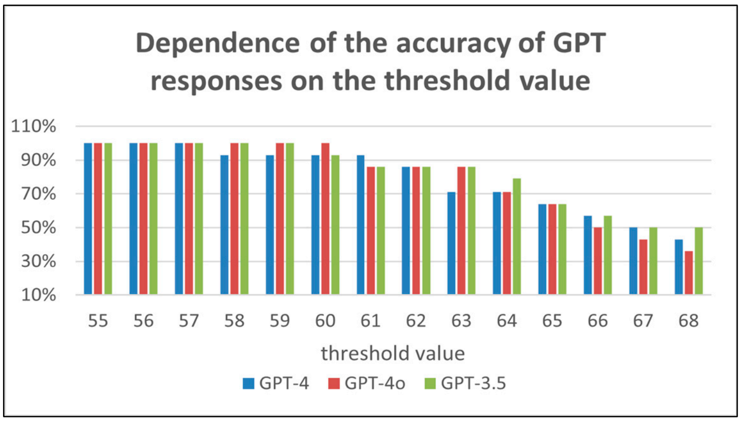

Figure 6. Dependence of ACC with increasing threshold value.

Figure 7 shows a comparison of the analyzed models by the variety of words in the text explaining CVD risk factors, grouped by the risk factor number.

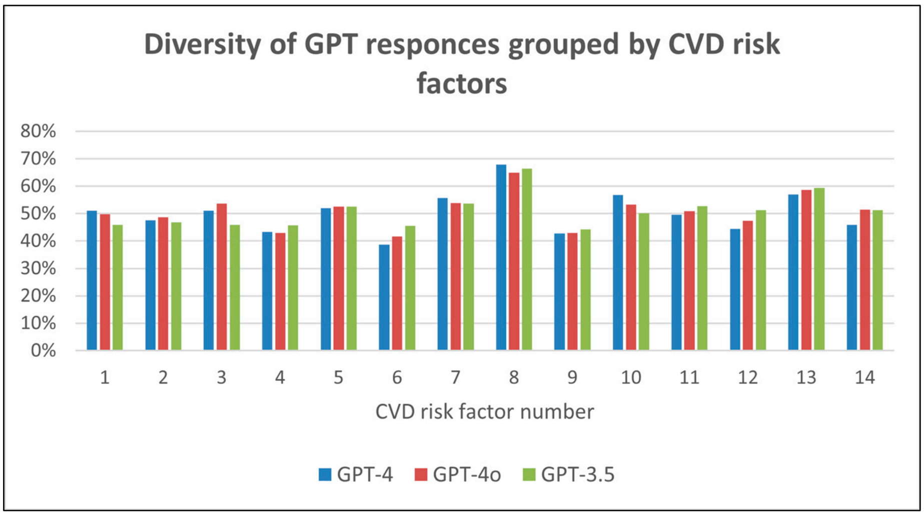

Figure 7. Comparison of the Diversity of GPT models.

The Diversity index calculated using Formula (5) ranges from 39% to 68%. Low Diversity values correspond to the risk factor numbers 4 (body mass index), 6 (diet), and 9 (stroke), while high Diversity values are obtained when explaining family history (factor 8). To calculate the DIV metric, a threshold of 50% was used. Then, the average Diversity of the GPT-4 model explanations of 50% is slightly lower than the average Diversity of the compared models of 57%. It appears that GPT-4 is a less suitable model for explaining CVD risk factors based on the above experimental results using the ASS, ACC, and DIV metrics presented in Table 5.

Taking into account that the expert explanation (ground truth) of CVD risk factors depends on the expert's vocabulary and experience, in addition to examining the GPT models using the metrics defined above, we conducted a secondary assessment of the models by the doctors.

Four independent physicians with 4 to 25 years of experience in the field of prevention and treatment of CVDs from three regions gave consent to a qualitative assessment of the GPT explanations from the point of view of evidence-based medicine. The physicians rated the correctness of 56 generated GPT explanations on three properties, namely safety for humans (safety), accuracy (E_accuracy), and compliance with official recommendations for the prevention of CVD (coherence). The score for each of the three properties was determined from 1 to 10, where 1 means that the explanation for a given CVD risk factor is completely inconsistent with the property, and 10 means that the explanation is completely consistent with the property. The physicians' ratings were then averaged across the analyzed properties and normalized by dividing by 10, the results of which are presented in Table 6 and Figure 8. The safety property of GPT-generated explanations received the highest score of almost 100% for each model. Physicians rated the coherence of the explanations with official CVD prevention guidelines quite highly, with the GPT-4 model receiving the highest average score of 92% and GPT-3.5 the lowest average score of 84%. As can be seen from Figure 8, the greatest variation is observed in E_accuracy: GPT-3.5 shows a minimum accuracy of 75%, and GPT-4 has the highest accuracy of 91%.

Table 6. Comparison of GPT models by physicians' estimates.

	Safety %	E_Accuracy %	Coherence %	Mean %
GPT-4	99	91	92	94
GPT-4o	99	85	90	92
GPT-3.5	99	75	84	86

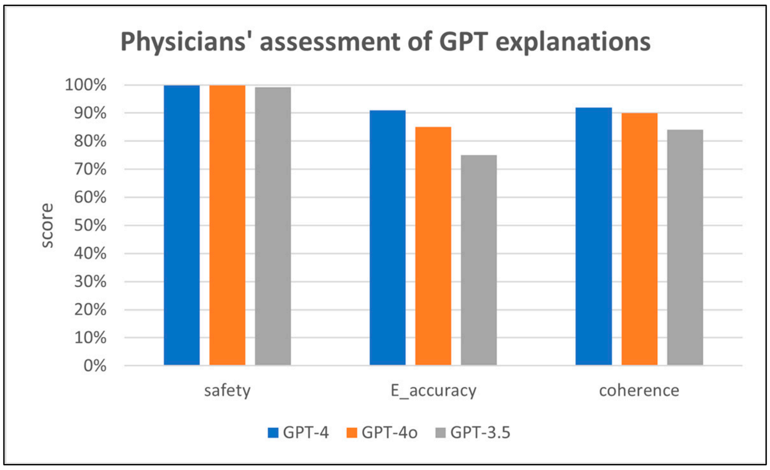

Figure 8. Physician estimates of GPT-based explanations.

We found that the accuracy estimates for the GPT-3.5 and GPT-4 models obtained using the ACC metric are not consistent with the estimates of E_accuracy given by the doctors, as shown in Figure 9. For the GPT-4o model only, the accuracy estimates given by the physicians and metrics are consistent and quite high.

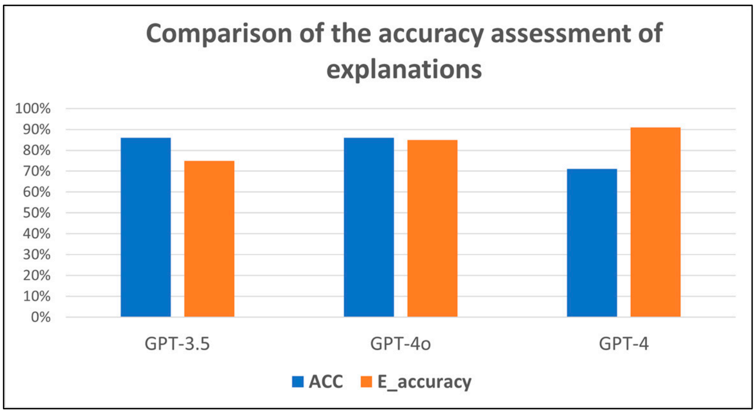

Figure 9. Comparison of accuracy by metric ACC with accuracy given by physicians.

For the first time, GPT models, namely ChatGPT-3.5, ChatGPT-4, and ChatGPT-4.o, were leveraged and studied to create explanatory items of the recommendations. In explaining recommendations, ChatGPT-3.5, ChatGPT-4, and ChatGPT-4.o demonstrate high accuracy, from 75% to 91%, and coherence with modern official guidelines, from 84% to 92%. Moreover, the safety property of ChatGPT-generated explanations estimated by doctors received the highest score of almost 100%. In [25], the authors found that a popular online artificial intelligence model (ChatGPT) provided largely adequate answers to simple questions about CVD prevention, as rated by preventive cardiologists. Unlike this study, in our study, the explanations for CVD risk factors of three ChatGPT models were assessed based on ASS, ACC, DIV metrics and on the expert knowledge of physicians by safety, E_accuracy, and coherence. Taking into account the stability and correctness of GPT-4.o responses, ChatGPT-4.o was used to explain why CVD risk factors should be managed.

7.3. Study of the Effectiveness of a Recommendation Algorithm

To study of the effectiveness of the CaRiFaM algorithm, a prototype web service implementing a recommendation algorithm was designed using a stack for modern web application development. Kotlin 1.6.10 and Spring Boot 2.7.17 were used as powerful and productive backend development tools, which offer a vast array of features enabling rapid, accessible microservice development and deployment. For building user interfaces, ReactJS 18.2.0 was applied, since it provides a robust frontend framework for building interactive user interfaces. Object-relational database system PostgreSQL 12 was chosen and implemented as a reliable and feature-rich solution for recommendation items and users' profile storage. In the study, Google Collaboration and Python 3.10, with the necessary libraries (numpy, scikit-learn, pandas, keras, tensorflow, and autogluon.tabular), were used.

To evaluate the proposed CaRiFaM algorithm from the user's point of view, a pilot study was conducted with the participation of 15 independent individuals. The age of the participants varied from 34 to 70 years; among them, 53% were men, and 47% were women. The purpose of the study was to find out the degree of user satisfaction with the content of the recommendations received and the degree of agreement of their opinions. To analyze the opinions of users, a special questionnaire was developed, the questions of which meet the requirements defined in Section 3 and are presented in Table 7. The Likert scale was used to create a survey of participant satisfaction with the content of the recommendations received. The users assessed the individual recommendations according to seven criteria, using a score from 1 to 5, where score 1 corresponds to completely disagree and 5 corresponds to completely agree with the statement given in the questionnaire. The results of the user survey based on developed questionnaire can be seen in Table 7. A degree of user satisfaction (DUS) is calculated by dividing the mean by the maximum score, i.e., 5:

$$DUS_l = \frac{100\%}{5k} \sum_{v=1}^{k} y_{lv}$$

Table 7. Questionnaire and assessment of user satisfaction. Here, SD designates standard deviations, and DUS denotes degree of users' satisfaction.

No.	Feature of the Individual Recommendations	Mean	SD	DUS%
1	Completeness. The resulting recommendations comprehensively cover your cardiovascular health concerns and inform you about cardiovascular risk factors	4.4	0.40	88
2	Explainability. Recommendations explain the impact of risk factors on cardiovascular health	4.4	0.40	88
3	Timeliness. The recommendations received are relevant for seeking medical help	4.3	0.52	87
4	Personalization. The recommendations received are specific to you and presented in understandable language	4.2	0.46	84
5	Planning. A recommended plan of steps you need to take to help you achieve your cardiovascular disease prevention goals	4.6	0.40	92
6	Motivation. The recommendations you receive will motivate you to manage your CV health	4.2	0.31	84
7	Usefulness. The recommendations received are useful for you	4.6	0.40	92
	Mean	4.4	0.44	88
	Cronbach's alpha = 0.86			

Here, y_{lv} denotes the score of the v-th participant on the l-th question of the questionnaire, and k is the number of participants.

To determine the consistency of the survey results, Cronbach's alpha statistics were calculated using MS Excel. A Cronbach's alpha value of 0.86 indicates a good user agreement on the quality of recommendations received. The data in Table 7 show that the participants agree that individual recommendations meet all quality criteria.

Moreover, some persons noted that they had previously received the exact such recommendations from their attending physicians. An analysis of the average ratings of the features of recommendations shows that the usefulness property received the highest rating of 92%, while properties such as personalization and motivation received the lowest, but at the same time, a quite high rating of 84%. Interestingly, the explainability of individual recommendations, for the generation of which ChatGPT-4o was used, received a high score of DUS = 88%. The difference in the degree of satisfaction with recommendations for the management of CVD risks in men and women is shown in Figure 10.

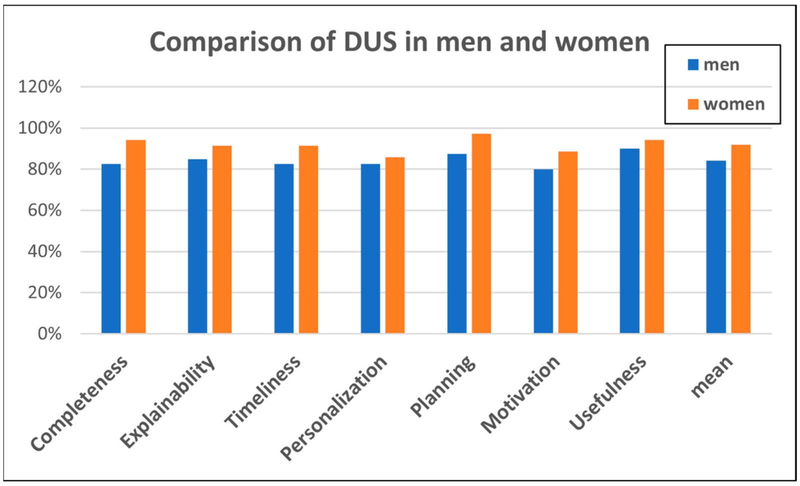

Figure 10. Level of recommendation satisfaction in men and women.

Through meticulous analysis, it turned out that for all characteristics of the proposed recommendations, the DUS for men is lower than for women and ranges from 80% for motivation to 90% for usefulness (Figure 10). This result could be explained by the fact that men are less adherent and perhaps more critical in their thinking. It can be seen in Figure 11 that across most features of recommendations, the highest DUS was demonstrated by participants in the 46–56 age group, and the lowest DUS was expressed by participants in the 34–45 age group. It is known that CVD and risk factors are significantly less common in people aged 34–45 years than in the age group 46–79 years, which may be why younger people are expected to receive more detailed advice on physical activity or nutrition, while for such people, the proposed algorithm generates fairly general recommendations based on official guidelines. We also found that as the age of survey participants increases, their satisfaction with the following properties of received recommendations increases: usefulness, motivation, and completeness.

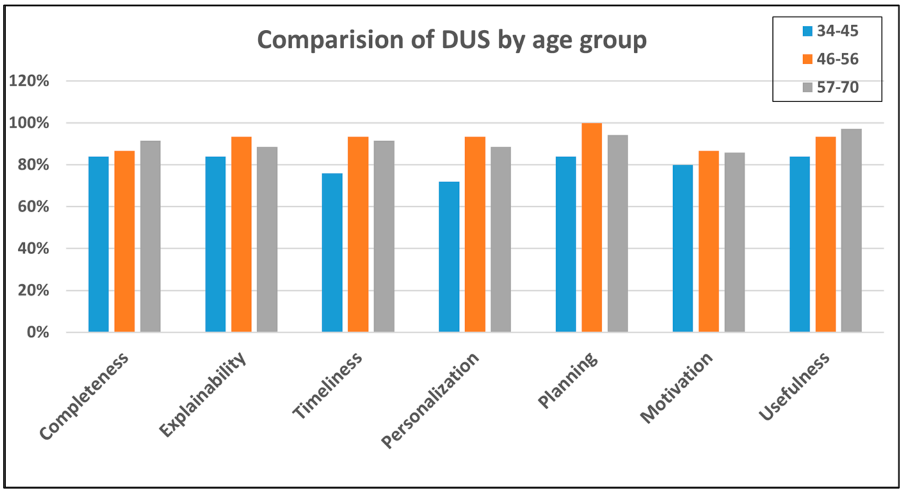

Figure 11. Comparison of recommendation satisfaction by age of users.

Two independent physicians specializing in the prevention and treatment of CVD gave their consent for a qualitative assessment of the proposed CaRiFaM algorithm from the point of view of evidence-based medicine.

The first expert is a Doctor of Medical Sciences and professor in the medical faculty of the Institute of Medicine, Ecology and Physical Culture of Ulyanovsk State University, Russia; the second one is a cardiologist and therapist working at the Moscow clinic GLOBAL MEDICAL SYSTEMS (GMS LLC), Russia. Both physicians have extensive experience: more than 20 years of practice in the treatment of persons with CVD as well as experience in preventive counseling. The description of the algorithm and the texts of recommendations received by the participants were independently analyzed by the physicians. The focus was on assessing the validity of the recommendations generated by the CaRiFaM algorithm and their relevance to the persons' CVD risk factors and official guidelines [1].

Both experts expressed the opinion that the proposed algorithm generates recommendations that are consistent with official guidelines and clinical practice, so it can be used by adults in the self-monitoring of CVD risk factors. They also noted the usefulness and safety of the recommendations and gave a positive assessment of the completeness of the set of CVD risk factors and their explanations obtained using ChatGPT. The physicians agree that the proposed algorithm is also promising in clinical practice, since it can reduce the doctor's time when preparing recommendations for the patient. At the same time, the physicians made some useful comments. In particular, to support adults with cardiovascular disease at home, the content of recommendations should be more specific and could be expanded using electronic health record data. For example, it would be useful for such patients to add recommendations for taking prescribed medications or a description of the dynamics of CVD health indicators.

7.4. Assessment of the Potential Benefit of Using the Proposed Algorithm

The main user of the CaRiFaM algorithm is a person interested in consciously managing their health to increase the duration of an active, full life. Considering that such individuals may have various CVD risk factors, including CVDs, in traditional prevention scenarios, they are regularly forced to participate in preventive programs and monitor the basic CV health indicators.

Barriers and limitations in the Implementation of Prevention Strategies for Chronic Disease Patients at the individual level are considered in the systematic review [38]. Based on the Action Plan for the Prevention and Control of Non-Communicable Diseases in the WHO European Region, this review highlighted the priorities at the individual level, in particular cardio-metabolic risk assessment and management as well the early detection of CVD. In a scenario of utilizing the proposed recommender algorithm, a person can perform most of the above activities at home and at any time with the evidence-based resources. This not only saves a person's time but also raises their engagement in the process of prevention and control of CVD, which consequently increases its effectiveness. Moreover, the activities aimed at prevention and self-managing CVD risk factors can eliminate or mitigate the need for future expensive treatments and rehabilitation.

Increasing the availability of preventive care for the population 24/7 and increasing the proportion of the population covered for the prevention of CVD based on information technology are the obvious advantages of using the proposed algorithm of recommendations at the population level.

The potential benefit of incorporating the proposed recommendation algorithm into a population-based CVD prevention strategy will be estimated as the level of reduction in CVD mortality resulting from following the recommendations generated by the proposed algorithm. For this purpose, we used the expected mortality proportion of the population caused by CVD of 36% for European middle-income countries [39].

Since following the recommendations for managing CVD risk requires a person to make systematic efforts to modify their lifestyle, first, we will need to determine the proportion of people who are committed to a healthy lifestyle. Given that adherence to a healthy lifestyle may vary for ethnic groups and individuals with different income levels, we will use an average proportion of the population adhering to a healthy lifestyle of 25%, following the results presented in previous articles [40–42].

To assess the potential benefit from following the recommendations for managing CVD risk factors generated by the CaRiFaM algorithm, the Cleveland Clinic Heart Disease Dataset was explored in [43] and described in detail in [44]. This dataset contains data from an experimental study of participants with suspected CVDs who do not have a history of them before, and, as can be assumed, did not control their CVD risk factors. As a result of the experiment, which consisted in a thorough clinical examination, the participants were divided into two groups, namely those for whom CVD was diagnosed and those for whom CVD was not detected.

It is natural to assume that if the participants for whom CVD was established had received and followed the proposed recommendations for self-control of their CVD risk factors, it is potentially possible that in this experiment, they would have been assigned to the group without CVD. The Cleveland Clinic Heart Disease Dataset comprises 303 observations, 13 descriptive attributes, and 1 target attribute, but only 7 are used in our study: age, gender, blood pressure (trestbps), cholesterol level (chol), fasting blood sugar (fbs), exercise induced angina (exang), and diagnosis of heart disease (target). These CVD indicators were chosen, as they are consistent with the CVD risk factors considered in the paper.

We analyzed data from all participants with CVD, including 82% men and 18% women. The prevalence of the identified CVD risk factors in these 137 participants is shown in Figure 12. It can be seen that biological CVD risk factors are quite common for participants of working age. In this subgroup, the percentage of participants with major CVD risk factors varies, with cholesterol ranging from 68% to 90% and blood pressure ranging from 36% to 83%. Moreover, for participants of working age, angina pectoris, which is a significant symptom of CVD that may indicate the onset of an event leading to the end-point outcome, is noted for 59% of participants in the age group 50–59 years and 48% for the 30–49 and 60–69 age groups.

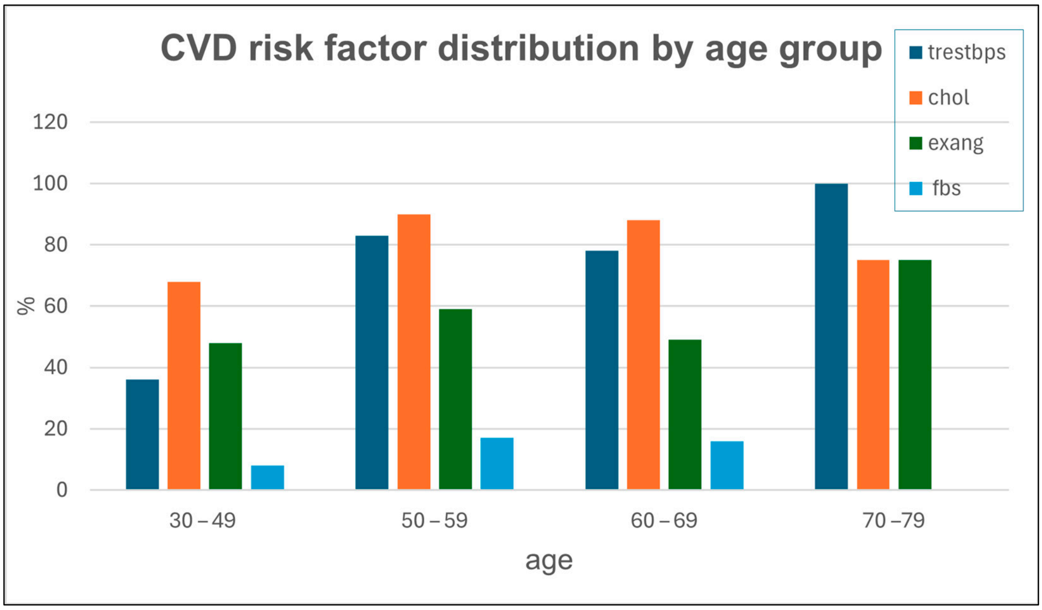

Figure 12. Comparison of age groups with CVD risk factors.

We also found that most individuals with established CVD had more than two risk factors for CVD, indicating low adherence to a healthy lifestyle and uncontrollability of the risk factors. Using an expected mortality proportion of the population caused by CVD of 36%, the expected mortality for the Cleveland Dataset participants was determined for 49 persons. Assuming that 100% of the participants received and 25% of them, namely 35 people, adhered to the recommendations generated by the proposed algorithm, the expected mortality for the remaining participants would be 37 people. Thus, for Cleveland participants with established CVD, the reduction in premature mortality will be 12 people (9%), who are mainly at the working age. A study of an open dataset shows that the potential benefit of including the proposed recommendation algorithm in the strategy for CVD prevention at the population level can be estimated at an approximately 9% reduction in premature mortality from CVD. Let us generalize the results of the considered example so that they can be used for any number of people with risk factors leading to CVD.

Let k be the number of people with CVD risk factors corresponding to the multifactorial CVD model introduced in the article, among which there are those with established or predicted CVD. Let $\beta_m \in [0,1]$ be the proportion of expected mortality caused by CVD, and $\beta_a \in [0,1]$ be the proportion of people who adhere to following recommendations for managing their CVD risk. To assess adherence to recommendations for CVD risk management, we will look at adherence to a healthy lifestyle.

Statement: The potential benefit Y from using the proposed recommendation algorithm is estimated by the level of reduction in the expected mortality in the studied population and is determined using the formula

$$Y = \beta_m \cdot \beta_a, Y \in [0,1].$$

Let us calculate the expected mortality in the study population in the form

$$E_m = \beta_m \cdot k$$

and potential adherence in the form

$$P_a = \beta_a \cdot k.$$

Then, the expected mortality in the study population, excluding individuals committed to following recommendations for managing their CVD risk, will be determined using the equation

$$E_{m1} = k \cdot \beta_m \cdot (1 - \beta_a).$$

The number of people who have reduced the risk of premature mortality from CVD was calculated as

$$E_{reduction} = E_m - E_{m1} = k \cdot \beta_m \cdot \beta_a.$$

As a result, the level of reduction in the expected mortality in the population is determined using the following expression:

$$Y = \beta_m \cdot \beta_a.$$

The following conclusions can be drawn from the proven statement:

Firstly, with $\beta_m = const$ and with an increase in the proportion of adherents β_a, the potential benefit of using the recommendation algorithm increases by preserving the working-age population and reducing healthcare costs. Secondly, when $\beta_m = const$ and $\beta_a = const$, the potential benefit of using a recommendation algorithm at the population level does not depend on the number of people in the population and is constant. For example, with β_m =0.36 [40] and β_a =0.25 [42], the potential benefit is 0.09, while, for example, for 100,000 subjects, the number of subjects who reduced the expected premature mortality will be 9000 people. The above findings allow us to highly evaluate the potential benefits of using the proposed algorithm at the population level.

8. Discussion

The proposed CaRiFaM algorithm has several key advantages as a knowledge-based recommendation algorithm. The first advantage is that it overcomes the data shortage obstacles that plague collaborative and content filtering recommendation algorithms. Compared to collaborative algorithms and content filtering, the proposed algorithm does not require many users and does not require users' ratings [5]. The novelty of the proposed algorithm from knowledge-based recommendation algorithms lies in its multidimensional recommendations, user profiles including predictive assessments of CVD risk, and generation of recommendations combining expert knowledge based on official guidelines with

ChatGPT knowledge. To highlight the opportunity of CaRiFaM algorithm, a comparison with similar knowledge-based recommendation algorithms is presented in Table 8.

Table 8. The comparison of health knowledge-based recommendation algorithms.

Criteria of Comparison		Spoladore [45]	Lopez-Barreiro [12]	Wang [11]	Proposed
Perspective of Comparison	**Indicator of Comparison**				
Number of risk factors in user profile corresponding to CVD risk factors	Lifestyle	2	9	4	3
	Biological	-	-	4	5
	Non-modifiable factors	-	-	4	5
	Descriptive estimates	2	9	12	13
	Predictive estimates	-	-	-	2
Recommendation associated with self-management of CVD risk factors	User focus	One class: older adults with chronic pathologies	One class: persons without medical conditions	One class: chronic disease patients, including CVD	Adults with or without any CVD risk factors as well symptoms of angina
	Goal	Support healthy diet and physical activity	Support physical, mental, and nutritional health	Support selection of educational materials	Support self-management of CVD risk factors
	Results	32 dishes, with diet plan for each day	Ranked list of challenges, selected from 30 health challenges	Ranked list of educational materials	Goal, information, and explanation for each person's risk factor
	Knowledge models and algorithms	Ontology and expert knowledge	Expert ranking model	Rules, ontology, and natural language processing	Rules, ChatGPT, and LSTM
	Guidelines or other official materials used in clinic practice	ACSM guideline and clinical literature	SF-36, sHEI-15, and OSC questionnaires	Corpus of educational materials	2021 ESC Guidelines on CVD prevention in clinical practice, SCORE algorithms
	Evaluation	Two use cases	Four experts, 30 uses	MAE on key words for corpus of educational materials, marked by two experts	Testing on open dataset, ASS, ACC, DIV, Safety, E_accuracy, and Coherence; four physicians, 15 users, and potential benefit assessment

Table 8 shows that the proposed algorithm evaluates a larger number of CVD risk factors in the user profile to generate recommendations, although it is inferior to other algorithms in the number of lifestyle factors. Moreover, only the proposed algorithm calculates predictive assessments of CVD risk using LSTM and SCORE models, while in the compared algorithms, the user profile contains only descriptive assessments. It follows that in the proposed algorithm, the user profile can describe more different health conditions of the human cardiovascular system. Consequently, this provides a greater variety of recommendations generated by the proposed algorithm than the compared ones. At the same time, we note that the proposed user profile does not include all CVD factors listed in the 2021 European guidelines [1]; in particular, the user profile does not include excessive alcohol consumption, regular stress, or poor-quality sleep. In [11,45], the recommendations are intended for one class of users, namely adults with chronic pathologies, which is narrower than in the proposed algorithm, since in addition to such users, recommendations are intended for different classes of users, including adults with or without any CVD risk factors as well symptoms of angina. Regarding the content of the recommendations, in contrast to one-dimensional recommendations presented as a list of dishes, exercises [45], challenges [12], or educational materials [11], the proposed recommendations are multidimensional, the contents of which are revealed in three aspects for each person's risk factor: targeted, informational, and explanatory. Information and semantic capacity are what distinguishes the content of the proposed multidimensional model of recommendation from those compared.

The goal of generating recommendations to support adults in the management of CVD risk factors at home was formulated only in the proposed algorithm. Moreover, to the best of our knowledge, this goal and its corresponding results have not been reported in research on HRSs. In our opinion, the reason is that the field of disease prevention, including CVD, is a new area for HRS with its own challenges and opportunities.

The conducted study of satisfaction analysis both from the point of view of end users and from the point of view of doctors showed that the recommendations for the self-management of CVD risk factors received high ratings. In summary, the proposed algorithm, CaRiFaM, integrates several key enhancements of health recommender algorithms, including a modified structure of recommendation and extended user profile, leveraging LLMs and deep ML models.

A limitation of the proposed CaRiFaM algorithm is the lack of recommendations for the use of medications, such as statins or aspirin. It is assumed that the appointment of such an intervention is the task of the attending physician. We also note that in the proposed algorithm, data input and recommendation output are implemented only in text form. The lack of analysis of tendencies in the patient's profile and his/her CV risk factors is another limitation of the proposed recommendation algorithm. In our opinion, a common limitation of knowledge-based recommendation algorithms is the use of subjective quality assessments. We recognize that larger studies involving medical professionals and different types of users are needed to evaluate its effectiveness in preventing CVD. Despite the fact that the proposed algorithm used most of the CVD factors listed in the 2021 European guidelines, psychological factors and air pollution were not considered. Also, the limitation of this study is its focus on only the European population.

9. Conclusions

The study proposes and investigates a new CaRiFaM algorithm for recommendations for the self-management of CVD risk factors. Before developing the algorithm, the basic requirements for the recommendation were defined, which determined its structure with explanatory items.

The content of the generated recommendations is based on the proven medical knowledge provided in the European Clinical Guidelines of 2021 [1], on the knowledge about user's CVD risks, and on ChatGPT knowledge. Preliminary and thorough studies of ChatGPT-3.5, ChatGPT-4, and ChatGPT-4o showed the effectiveness of these LLMs in generating explainable recommendations within the proposed algorithm, among which ChatGPT-4o responses were estimated as the most acceptable on average (accuracy 85%, safety 99%, and coherence 90%). Thus, these LLMs have big potential for HRS development by providing recommendation explanations. Leveraging ChatGPT-4o for the explanations of recommendation determines the novelty of the proposed algorithm. A distinctive feature of the user profile proposed in the article is the expansion of CVD risk factors by including angina symptoms and predictive assessments of CV health. To obtain predictive assessments of the user's CVD health, in addition to the SCORE2/SCORE2-OP models recommended in clinical practice, the LSTM model was used. The recommendations for the self-management of CVD risk factors received high ratings from both end users and physicians. End-user satisfaction with the recommendations was assessed using seven criteria, among which explainability and usefulness received high ratings, 88% and 92%, respectively. We emphasize that the global purpose of developing the CaRiFaM algorithm is to reduce mortality from CVD, which is often caused by a person's ignorance of their CVD risk factors, lack of knowledge on how to control them, and what conditions of their CV system require medical care, including urgent care. We hope that future development and dissemination of the proposed recommender algorithm will improve population-level CVD prevention strategies and reduce premature mortality.

Future research will focus on two directions: (1) assessing the capabilities of fuzzy logic rules for assessing CVD risk and ranking recommendation items and (2) analyzing the distributions of risk factors in different ethnic populations.

Author Contributions: Methodology, conceptualization, original draft preparation, visualization, and formal analysis, T.V.A.; investigation, data curation, formal analysis, and software, P.V.P.; writing, editing, and visualization, A.V.K. (Andrey V. Komolov); investigation, data analysis, and software, A.V.K. (Andrey V. Kuzlyakin). All authors contributed equally. All authors have read and agreed to the published version of the manuscript.

Funding: This research was performed in the framework of the state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation, project "Models, methods, and algorithms of artificial intelligence in the problems of economics for the analysis and style transfer of multidimensional datasets, time series forecasting, and recommendation systems design", grant no. FSSW-2023-0004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The data will be made available by the authors on request.

Acknowledgments: The authors express their gratitude to the experts for their careful analysis and valuable comments, which undoubtedly improved the content of this article: Alexander Mikhailovich Shutov, cardiologist at Ulyanovsk Regional Clinical Hospital, Ulyanovsk, Russia, and Yuri Sergeyevich Zamashkin, cardiologist at GLOBAL MEDICAL SYSTEMS Limited Liability Company (LLC GMS), Moscow, Russia.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Visseren, F.L.J.; Mach, F. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur. Heart J.* **2021**, *42*, 3227–3337. [CrossRef] [PubMed]
2. *Global Strategy on Digital Health 2020–2025*; World Health Organization: Geneva, Switzerland; Available online: https://www.who.int/publications/i/item/9789240020924 (accessed on 15 July 2024).
3. Cai, Y.; Yu, F. Health Recommender Systems Development, Usage, and Evaluation from 2010 to 2022: A Scoping Review. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15115. [CrossRef]
4. Uta, M.; Felfernig, A.; Le, V.-M.; Burgstaller, T. Knowledge-based recommender systems: Overview and research directions. *Front. Big Data* **2024**, *7*, 1304439. [CrossRef]
5. Burke, R. Knowledge-based recommender systems. In *Encyclopedia of Library and Information Science*; CRC Press: Boca Raton, FL, USA, 2000; Volume 69, pp. 180–200.
6. Aggarwal, C.C. *Recommender Systems: The Textbook*; Springer International Publishing: Cham, Switzerland, 2016.
7. SCORE2 working group and ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: New models to estimate 10-year risk of cardiovascular disease in Europe. *Eur. Heart J.* **2021**, *42*, 2439–2454. [CrossRef]
8. Tran, T.N.T.; Felfernig, A.; Trattner, C.; Holzinger, A. Recommender systems in the healthcare domain: State-of-the-art and research issues. *J. Intell. Inf. Syst.* **2020**, *57*, 171–201. [CrossRef]
9. Etemadi, M.; Abkenar, S. A systematic review of healthcare recommender systems: Open issues, challenges, and techniques. *Expert Syst. Appl.* **2023**, *213*, 118823. [CrossRef]
10. De Croon, R.; Van Houdt, L. Health Recommender Systems: Systematic Review. *J. Med. Internet Res.* **2021**, *23*, e18035. [CrossRef] [PubMed]
11. Wang, Z.; Huang, H.; Cui, L.; Chen, J.; An, J.; Duan, H.; Ge, H.; Deng, N. Using Natural Language Processing Techniques to Provide Personalized Educational Materials for Chronic Disease Patients in China: Development and Assessment of a Knowledge-Based Health Recommender System. *Med. Inf. J.* **2020**, *8*, e17642. [CrossRef]
12. Lopez-Barreiro, J.; Garcia-Soidan, J.L.; Alvarez-Sabucedo, L.; Santos-Gago, J.M. Practical Approach to Designing and Implementing a Recommendation System for Healthy Challenges. *Appl. Sci. J.* **2023**, *13*, 9782. [CrossRef]
13. Dogan, A.; Li, Y.; Odo, C.P.; Sonawane, K.; Lin, Y.; Liu, C. A utility-based machine learning-driven personalized lifestyle recommendation for cardiovascular disease prevention. *J. Biomed. Inform.* **2023**, *141*, 104342. [CrossRef]

14. Honka, A.M.; Nieminen, H.; Similä, H.; Kaartinen, J.; Gils, M.J. A Comprehensive User Modeling Framework and a Recommender System for Personalizing Well-Being Related Behavior Change Interventions: Development and Evaluation. *IEEE Access J.* **2022**, *10*, 116766–116783. [CrossRef]
15. Kobrinskii, B.A.; Grigoriev, O.G.; Molodchenkov, A.I.; Smirnov, I.V.; Blagosklonov, N.A. Artificial Intelligence Technologies Application for Personal Health Management. *IFAC-Pap.* **2019**, *52*, 70–74. [CrossRef]
16. Kung, T.H.; Cheatham, M.; Medenilla, A.; Sillos, C.; De Leon, L.; Elepaño, C.; Madriaga, M.; Aggabao, R.; Diaz-Candido, G.; Maningo, J.; et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLoS Digit. Health* **2023**, *2*, e0000198. [CrossRef]
17. Tran, T.N.T.; Felfernig, A.; Man Le, V.; Chau, T.M.N.; Mai, T.G. User Needs for Explanations of Recommendations: In-depth Analyses of the Role of Item Domain and Personal Characteristics. In Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23), Limassol, Cyprus, 26–29 June 2023; Association for Computing Machinery: New York, NY, USA, 2023; pp. 54–65. [CrossRef]
18. Li, Z.; Chen, Y.; Zhang, X.; Liang, X. BookGPT: A General Framework for Book Recommendation Empowered by Large Language Model. *Electronics* **2023**, *12*, 4654. [CrossRef]
19. Ray, P.P. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 121–154. [CrossRef]
20. Shahab, O.; El Kurdi, B.; Shaukat, A.; Nadkarni, G.; Soroush, A. Large language models: A primer and gastroenterology applications. *Ther. Adv. Gastroenterol.* **2024**, *17*, 1–15. [CrossRef]
21. Gala, D.; Makaryus, A.N. The Utility of Language Models in Cardiology: A Narrative Review of the Benefits and Concerns of ChatGPT-4. *Int. J. Environ. Res. Public Health* **2023**, *20*, 6438. [CrossRef]
22. Rao, A.; Pang, M.; Kim, J.; Kamineni, M.; Lie, W.; Prasad, A.K.; Landman, A.; Dreyer, K.; Succi, M.D. Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study. *J. Med. Internet Res.* **2023**, *25*, e48659. [CrossRef]
23. Dave, T.; Athaluri, S.A.; Singh, S. ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front. Artif. Intell.* **2023**, *6*, 1169595. [CrossRef]
24. Groza, T.; Caufield, H.; Gration, D.; Baynam, G.; Haendel, M.A.; Robinson, P.N.; Mungall, C.J.; Reese, J.T. An evaluation of GPT models for phenotype concept recognition. *BMC Med. Inform. Decis. Mak.* **2024**, *24*, 30. [CrossRef]
25. Sarraju, A.; Bruemmer, D.; Van Iterson, E.; Cho, L.; Rodriguez, F.; Laffin, L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. *JAMA* **2023**, *329*, 842–844. [CrossRef] [PubMed]
26. Valery, L.; Feigin; Krishnamurthi, R. Digital solutions for primary stroke and cardiovascular disease prevention: A mass individual and public health approach. *Lancet Reg. Health West. Pac.* **2022**, *29*, 100511. [CrossRef]
27. Rose, G.A. The diagnosis of ischemic heart pain and intermittent claudication in field surveys. *Bull. World Health Organ.* **1962**, *27*, 645–658. [PubMed]
28. ChatGPT: Optimizing Language Models for Dialogue. Open AI. Available online: https://openai.com/blog/chatgpt (accessed on 15 July 2024).
29. Baktash, A.J.; Dawodi, M.; Chat GPT. GPT-4: A Review on Advancements and Opportunities in Natural Language Processing. *J. Electr. Electron. Eng.* **2023**, *2*, 548–549. [CrossRef]
30. Erickson, N.; Mueller, J.; Shirkov, A.; Zhang, H.; Larroy, P.; Li, M.; Smola, A. AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. *arXiv* **2020**, arXiv:2003.06505v1.
31. Cardiovascular Disease Dataset. Kaggle. Available online: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset (accessed on 15 July 2024).
32. Naseer, Z.A.N.; Nazeer, H. Cardiovascular Disease Detection Using Multiple Machine Learning Algorithms and their Performance Analysis. In Proceedings of the 2022 International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE), Lahore, Pakistan, 2–4 December 2022; pp. 1–7. [CrossRef]
33. Fayez, M.; Kurnaz, S. Novel method for diagnosis diseases using advanced high-performance machine learning system. *Appl. Nanosci.* **2021**, *13*, 1787. [CrossRef]
34. Maiga, J.; Hungilo, G.G.; Pranowo, P. Comparison of Machine Learning Models in Prediction of Cardiovascular Disease Using Health Record Data. In Proceedings of the 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, 24–25 October 2019; pp. 45–48. [CrossRef]
35. Zero-Shot Prompting. Prompt Engineering Guide. Available online: https://www.promptingguide.ai/techniques/zeroshot (accessed on 15 July 2024).
36. Singhal, A. Modern Information Retrieval: A Brief Overview. *IEEE Data Eng. Bull.* **2001**, *24*, 35–43.
37. Wu, S.; Liu, F.; Zhang, K. Short text similarity calculation based on jaccard and semantic mixture in Bio-Inspired Computing: Theories and Applications. In *Bio-Inspired Computing: Theories and Applications*; Springer: Singapore, 2021; pp. 37–45.
38. Duda-Sikuła, M.; Kurpas, D. Barriers and Facilitators in the Implementation of Prevention Strategies for Chronic Disease Patients-Best Practice GuideLines and Policies' Systematic Review. *J. Pers. Med.* **2023**, *13*, 288. [CrossRef]
39. Timmis, A.; Vardas, P.; Townsend, N.; Torbica, A.; Katus, H.; De Smedt, D.; Gale, C.P.; Maggioni, A.P.; Petersen, S.E.; Huculeci, R.; et al. European Society of Cardiology: Cardiovascular disease statistics 2021. *Eur. Heart J.* **2022**, *43*, 716–799. [CrossRef]

40. Shalnova, S.A.; Maksimov, S.A.; Balanova, Y.A.; Evstifeeva, S.E.; Imaeva, A.E.; Kapustina, A.V.; Karamnova, N.S.; Muromtseva, G.A.; Viktorova, I.A.; Prishchepa, N.N.; et al. Adherence to a healthy lifestyle of the Russian population depending on the socio-demographics. *Cardiovasc. Ther. Prev.* **2020**, *19*, 2452. (In Russian) [CrossRef]
41. Da Costa, F.A.; Pedro, A.R.; Teixeira, I.; Bragança, F.; da Silva, J.A.; Cabrita, J. Primary non-adherence in Portugal: Findings and implications. *Int. J. Clin. Pharm.* **2015**, *37*, 626–635. [CrossRef] [PubMed]
42. Andualem, A.; Gelaye, H.; Damtie, Y. Adherence to Lifestyle Modifications and Associated Factors Among Adult Hypertensive Patients Attending Chronic Follow-Up Units of Dessie Referral Hospital, North East Ethiopia, 2020. *Integr. Blood Press. Control* **2020**, *13*, 145–156. [CrossRef] [PubMed] [PubMed Central]
43. Cleveland Clinic Heart Disease Dataset. Kaggle. Available online: https://www.kaggle.com/datasets/aavigan/cleveland-clinic-heart-disease-dataset (accessed on 15 July 2024).
44. Marateb, H.R.; Goudarzi, S. A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system. *J. Res. Med. Sci.* **2015**, *20*, 214–223. [PubMed] [PubMed Central]
45. Spoladore, D.; Colombo, V.; Arlati, S.; Mahroo, A.; Trombetta, A.; Sacco, M. An Ontology-Based Framework for a Telehealthcare System to Foster Healthy Nutrition and Active Life-style in Older Adults. *Electronics* **2021**, *10*, 2129. [CrossRef]