Intelligent Health Systems – From Technology to Data and Knowledge
E. Andrikopoulou et al. (Eds.)
© 2025 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHTI250284

A Medical Decision Support System for Automatic Treatment Plan Generation Using Machine Learning Algorithms

Florian MAZURA[a,1], Marius GERDES[b], Rick PETZOLD[b], Wilhelm STORK[a,b], Markus SCHINLE[c], and Simon STOCK[b]
[a] *FZI Research Center for Information Technology*
[b] *Karlsruhe Institute of Technology KIT*
[c] *Offenburg University of Applied Sciences*
ORCiD ID: Florian Mazura https://orcid.org/0009-0008-0935-2703, Marius Gerdes https://orcid.org/0000-0002-6555-0199, Wilhelm Stork https://orcid.org/0000-0003-0579-4615, Markus Schinle https://orcid.org/0000-0002-3132-0280, Simon Stock https://orcid.org/0000-0001-8326-9758

Abstract. Due to demographic change, health economics is increasingly focused on the quality of life in advanced age and the associated cost aspects. Dementia is one of the key issues in this area and its efficient treatment will become increasingly relevant in the near future. Our system aims to automatically create treatment plans for dementia patients in a digital platform. To this end, three algorithms were implemented: a rule-based approach, an approach based on artificial neural networks, and an approach using large language models. A comprehensive synthetic dataset with fictitious patients, medical problems, and treatments was created for training and evaluation of the algorithms.

Keywords. artificial intelligence, dementia, digital health, machine learning, recommender system, treatment plan

1. Introduction

55 million people are living with dementia today and this number is forecast to rise to 139 million by 2050, posing a major challenge for the medical community [1]. The FINGER study explored the causes of dementia, revealing that a healthy diet, regular physical activity, cognitive training, and other interventions contribute decisively to preserving cognitive functions [2].

Addressing the challenges posed by dementia, the Metis research system consisting of a web platform and a smartphone app [3] empowers physicians to identify and assess risk factors, enabling personalized interventions for patients. Digital platforms such as the Metis system are not yet widely used by physicians, but there is a growing interest in them [4].

A core element of the Metis web platform is the treatment plan editor, allowing practitioners to create treatment plans graphically [5]. Figure 1 shows an example

[1] Corresponding Author: Florian Mazura; E-mail: mazura@fzi.de.

treatment plan addressing obesity (orange) with two interventions (red). The treatment plan is rounded off with challenges (blue) that motivate patients to achieve their personal health goals.

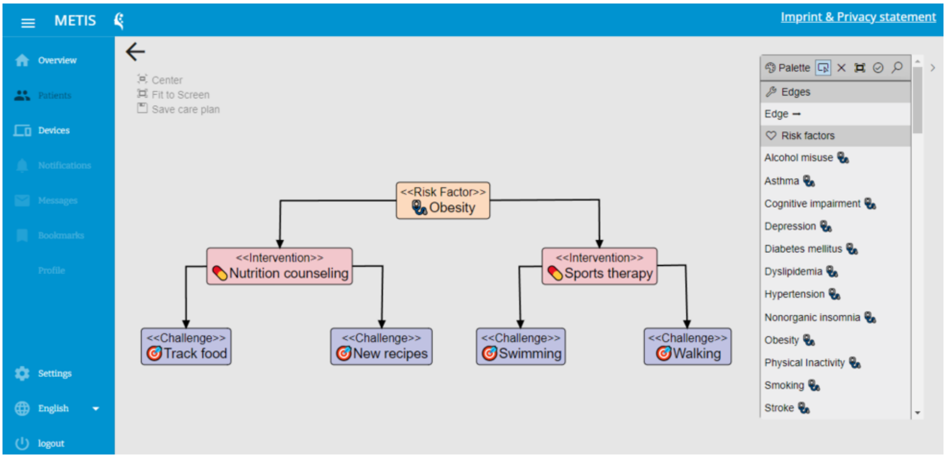

Figure 1. Treatment Planning in the Metis platform. Using the graphical editor, practitioners can create treatment plans for their patients.

With a growing shortage of physicians [6], exacerbated by demographic shifts, it is crucial to relieve physicians of time-consuming work such as creating individualized treatment plans, especially when similar to existing plans. Therefore, this paper discusses possibilities for the Metis platform to enable physicians to generate treatment plans automatically. A system designed to generate such plans automatically must: (1) select the appropriate treatment elements from all the available ones in the platform based on present risk factors and the patient's health data, and (2) relate the selected treatment elements correctly by drawing connections.

2. Methods

Three different algorithms for automatic treatment planning of varying complexity were implemented, tested, and compared for the implementation of a decision support system.

The simplest, a rule-based approach, focuses on program stability and avoiding fundamentally incorrect treatment suggestions. It identifies common patterns and assigns the most frequently prescribed treatment for each medical issue based on existing treatment plans. In the implementation of the rule-based algorithm, all existing plans of all patients are considered, and it is counted for each risk factor how often which treatment was used against it. To generate a treatment plan for a new patient, the treatments with the highest score for this patient's risk factors are selected.

The balanced algorithm uses an artificial neural network in the form of a multi-layer perceptron (MLP). Input nodes represent risk factors available in the Metis platform, output nodes represent available interventions and challenges. The presence of a risk factor and the assignment of a treatment element are coded binary. In the resulting output possible treatment methods are either accepted (1) or rejected (0). Figure 2 shows a simplified MLP for treatment planning. Our implemented MLP has 13 input nodes and 28 output nodes.

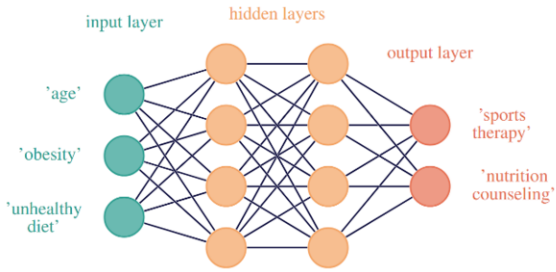

Figure 2. Simplified MLP architecture with one input layer, two hidden layers, and one output layer. Three input neurons are used to input information about a patient's diagnosis and his age. This network can be used to predict interventions for medical conditions.

The third algorithm is a large language model (LLM). The LLM is prompted to suggest suitable treatments for a patient's given risk factors. The desired structure of the response, i.e. the treatment plan, is also included in the prompt. It's based on a try-except structure to process responses, repeating up to ten times if errors occur. Using string extraction methods, the response can be translated into a treatment plan. Few-shot learning in the prompt stabilizes responses. The model used is GPT-3.5 Turbo from OpenAI. The LLM prediction process is visualized in Figure 3.

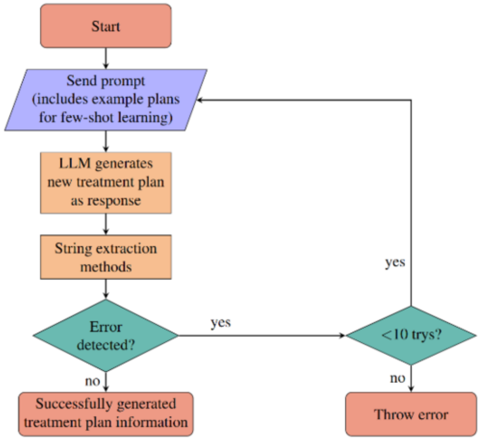

Figure 3. LLM Algorithm Flow Chart. The LLM responds to a prompt. The treatment plan information is extracted from this response, transformed, and checked for errors. If errors are found, the process is repeated up to 10 times.

3. Results

Data is required for training, testing, comparing, and evaluating the algorithms, but a challenge lies in the absence of a definitive ground truth. No publicly accessible dataset correlates (dementia) risk factors and interventions, and the highly personalized nature of treatment plans makes defining an ideal plan for an individual difficult.

For this reason, a synthetic dataset was created. It includes seven patients of different age groups (30-90), twelve of the most relevant dementia risk factors, 13 interventions, and 22 challenges. From this, 41 treatment plans were created, the meaningfulness of which is based on literature research. Table 1 shows an overview of which interventions were found to be effective against which risk factor, and the training dataset was created based on these proposals. For reasons of space, a more detailed insight into the development of the 41 ground truth treatment plans can be found on GitHub[2].

Table 1. Synthetic training data. The table shows valid treatment possibilities for the most relevant dementia risk factors. Treatments are only represented by interventions; specifying challenges have been omitted for reasons of space.

Risk Factors	Alcohol misuse	Asthma	Cognitive impairment	Depression	Diabetes mellitus	Dyslipidemia	Hypertension	Obesity	Physical inactivity	Smoking	Stroke	Unhealthy diet
Art therapy				X								
Cognitive training			X									
Cognitive-behavioral therapy				X			X				X	X
Measurement (e.g. blood pressure)					X		X					
Medication		X			X	X	X				X	
Nicotine replacement therapy										X		
Nutrition counseling					X	X	X	X				X
Physical therapy					X	X						
Psychotherapy	X			X								X
Respiratory therapy		X										
Social therapy	X			X						X		
Sports therapy				X				X	X		X	

For testing, all three algorithms were applied to a large number of fictional patients. The automatically generated dataset includes all possible combinations of up to three risk factors. The age of the patients was randomly assigned between 20 and 100.

With the number of possible risk factors being twelve, this resulted in a total of $\binom{12}{1} + \binom{12}{2} + \binom{12}{3} = 298$ test patients.

All 3 algorithms were applied to each of the test patients. For all assignments of interventions and also challenges to risk factors, it was then evaluated whether these were correct according to Table 1.

A suitable comparison metric is the ratio of correct assignments of treatment elements to the total number of assignments made (see Eq. 1). The results for each of the algorithms can be found in Table 2.

$$'Correct'\ ratio = \frac{correct\ predictions}{total\ number\ of\ predictions} \tag{1}$$

Algorithm	Total Predictions	Correct Predictions	Correct Ratio
Rule-based	1608	1608	1
MLP	3593	3007	0.837
LLM	4145	2481	0.599

Table 2. Results of the Algorithms. The table summarizes the performance of the three methods on the test dataset. The rule-based method mirrors the training data, while the LLM shows the highest error rate. The MLP is right in the middle in terms of the correct ratio. All results should always be checked by a physician.

[2] https://github.com/f-mazura/automatic-treatment-plan-generation, 07.01.2025

4. Discussion

The rule-based algorithm offers simplicity and efficiency, leveraging established trends in medical practice to inform decision-making. However, it does not consider much detailed information about the patient, such as the patient's age or other risk factors.

The balanced algorithm does not deliver as rigid results as its rule-based counterpart but brings less creativity into the planning process than the creative algorithm. The balanced algorithm can detect/treat combinations of risk factors, and personal information about the patient (e.g. age, etc.) can be included in the treatment. This is the most notable aspect in which the rule-based algorithm lags behind the AI-based algorithms.

While the first two algorithms are based on existing treatment plans as a training set, the creative algorithm using an LLM is able to introduce new treatment ideas. Although entailing more risk, this approach could open new possibilities for automated treatment. The rapid development of LLMs in the recent past makes this algorithm particularly interesting. At present, however, the rule-based and MLP-based algorithms are the more consistent choice, as they show a lower deviation from the training dataset (see also Table 2).

5. Conclusions

The developed system can automatically create treatment plans for dementia patients within a digital dementia platform using one of three implemented algorithms of varying complexity. It provides a good and operational basis for further improvements. In particular, the application of LLMs is very promising for the future, as their rapid development is expected to further enhance the performance of the LLM-based algorithm. An important next step in enhancing the decision support system will be to expand the dataset in collaboration with medical experts.

In times of demographic change and increasing prevalence of dementia, this system has the potential to significantly reduce the time physicians spend on treatment planning, enabling them to focus more on patient care.

References

[1] Alzheimer's Disease International. Dementia statistics [Internet]. London: Alzheimer's Disease International; 2023 Aug 29 [cited 2024 Aug 29]. Available from: https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/

[2] Kivipelto M and et al. Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) [Internet]. Finland: Finnish Institute for Health and Welfare; 2009-2025 [cited 2024 Aug 29]. Available from: https://clinicaltrials.gov/study/NCT01041989

[3] Stock S, Schinle M, Gerdes M, Stork W. A data-driven approach towards an e-health platform for dementia intervention. Gerontechnology. 2024;23(2):6-6. doi:10.4017/gt.2024.23.s.862.6.sp

[4] Schinle M, Erler C, Kaliciak M, Milde C, Stock S, Gerdes M, Stork W. Digital Health Apps in the Context of Dementia: Questionnaire Study to Assess the Likelihood of Use Among Physicians. JMIR Form Res. 2022;6(6):e35961. doi:10.2196/35961

[5] Schinle M, Dietrich M, Stock S, Gerdes M, Stork W. Model-Driven Dementia Prevention and Intervention Platform. Stud Health Technol Inform. 2023 May;302:937-941. doi:10.3233/SHTI230313

[6] Lau T, Osterloh F. Ärztemangel: Schlechte Aussichten. Dtsch Arztebl [Internet]. 2022 [cited 2024 Aug 29];119(22-23):A1008-10. Available from: https://www.aerzteblatt.de/int/article.asp?id=225635