AI showdown: info accuracy on protein quality content in foods from ChatGPT 3.5, ChatGPT 4, bard AI and bing chat

Hatice Merve Bayram and Arda Ozturkcan
Department of Nutrition and Dietetics, Istanbul Gelisim University, Istanbul, Turkey

Received 10 February 2024
Revised 11 April 2024
31 May 2024
Accepted 8 June 2024

Abstract

Purpose – This study aims to assess the effectiveness of different AI models in accurately aggregating information about the protein quality (PQ) content of food items using four artificial intelligence (AI) models – ChatGPT 3.5, ChatGPT 4, Bard AI and Bing Chat.

Design/methodology/approach – A total of 22 food items, curated from the Food and Agriculture Organisation (FAO) of the United Nations (UN) report, were input into each model. These items were characterised by their PQ content according to the Digestible Indispensable Amino Acid Score (DIAAS).

Findings – Bing Chat was the most accurate AI assistant with a mean accuracy rate of 63.6% for all analyses, followed by ChatGPT 4 with 60.6%. ChatGPT 4 (Cohen's kappa: 0.718, $p < 0.001$) and ChatGPT 3.5 (Cohen's kappa: 0.636, p: 0.002) showed substantial agreement between baseline and 2nd analysis, whereas they showed a moderate agreement between baseline and 3rd analysis (Cohen's kappa: 0.538, p: 0.011 for ChatGPT 4 and Cohen's kappa: 0.455, p: 0.030 for ChatGPT 3.5).

Originality/value – This study provides an initial insight into how emerging AI models assess and classify nutrient content pertinent to nutritional knowledge. Further research into the real-world implementation of AI for nutritional advice is essential as the technology develops.

Keywords Sustainable diet, Sustainability, Artificial intelligence, Al models, Food assessment, ChatGPT, Bard AI, Bing chat

Paper type Research paper

1. Introduction

The United Nations (UN) estimates that by 2050 the world population will reach 10.1 billion, contributing to almost 2.5 billion more mouths to feed (UN, 2019). Global wealth, particularly in emerging markets, has also increased over the past few decades. As a result, dietary habits are expected to change. Meat consumption has been a priority option for consumers when choosing what to eat, mainly because of its affordability and accessibility, as well as its nutritional value (Marques *et al.*, 2018). Indeed, research has found that, as living standards improve, so does the percentage of energy provided by animal-based products. The Food and Agriculture Organisation (FAO) of the UN concludes that meat production in the world has risen from 65 million tonnes to 279 million tonnes in the last 50 years, an increase of more than 400% (Vranken *et al.*, 2014). Additionally, according to the Organisation for Economic Co-operation and Development (OECD) and the FAO Agricultural Outlook 2023–2032, global and production growth is expected to increase by 12% by 2032, particularly in upper-middle-income countries (OECD and FAO, 2023). However, it is of concern because the consumption of livestock products is associated with a range of externalities, such as environmental impact, public health issues and food security.

The manuscript presents a research study that used AI chatbots for its investigations. In particular, it involved the use of ChatGPT versions GPT-3.5 and GPT-4.0, both developed by OpenAI. Additionally, this study used Bing Chat and Bard AI. All authors had access to the data and played a substantial role in writing the manuscript.

Funding: This research did not receive any particular grant from funding agencies in the public, commercial or not-for-profit sectors.

British Food Journal
Vol. 126 No. 9, 2024
pp. 3335-3346

0007-070X
DOI 10.1108/BFJ-02-2024-0158

The FAO defines sustainable diets as "Sustainable diets are those diets with low environmental impacts which contribute to food and nutrition security and a healthy life for present and future generations. Sustainable diets are protective and respectful of biodiversity and ecosystems, culturally acceptable, accessible, economically fair and affordable; nutritionally adequate, safe and healthy; while optimizing natural and human resources" (FAO, 2012). A sustainable and healthy diet is characterised energy intake in accordance with the needs of individuals who prefer plant-based foods instead of animal-based foods (Willett *et al.*, 2019). Animal-based diets were associated with higher greenhouse gas emissions (GHGE), whereas higher consumption of plant-based diets was associated with lower GHGE (Bayram and Ozturkcan, 2022; Clune *et al.*, 2017; Werner *et al.*, 2014). Additionally, animal-based products include harmful nutrients such as sodium, saturated fat and added sugars (Soedamah-Muthu and De Goede, 2018). Therefore, reduced consumption of animal-based products may result in higher sustainability, health advantages and a decrease in non-communicable disease risks (Bayram and Ozturkcan, 2022). Furthermore, switching from animal products to plant-based alternatives can have both benefits and drawbacks and whether they're "better" for health depends on various factors. In a recent review, salt was found to be the most commonly added ingredient in plant-based beverages (Fructuoso *et al.*, 2021). Similarly, high processing levels, a wide range of nutrients and high salt content were also found in plant-based meats and cheeses (Pointke and Pawelzik, 2022). It was also reported that plant-based meats have been suggested to have a 50% lower environmental impact than animal meat (Smetana *et al.*, 2023). However, there is a lack of robust, long-term evidence on the role of these plant-based products consumption in health (El Sadig and Wu, 2024).

Studies have reported a significant protein intake reduction between omnivores and vegans amongst adults (Allès *et al.*, 2017; Sobiecki *et al.*, 2016). The important point is protein quality (PQ) rather than protein quantity. Plant-based products on the market often have lower protein content and, more importantly, a lower PQ (Rojas Conzuelo *et al.*, 2022).

PQ is not only about the composition of the amino acids but also about their bioavailability (WHO/FAO/UNU Expert Consultation, 2007). Dietary PQ is primarily characterised by its indispensable amino acid (IAA) content. IAAs cannot be synthesised by the human body and must be obtained from the diet (Herreman *et al.*, 2020). FAO/WHO and the US Food and Drug Administration (FDA) first proposed a PQ evaluation method in 1991, the Protein Digestibility-Corrected Amino Acid Score (PDCAAS), and they made PDCAAS the official standard in 1993. But in 2013, FAO proposed shifting preference to a new evaluation method, the Digestible Indispensable Amino Acid Score (DIAAS) (FAO, 2013). The main limitation of PDCAAS is that it does not take into account anti-nutrient factors like phytic acid and trypsin inhibitors, which limit the absorption of protein amongst other nutrients, and its use of faecal digestibility, whereas in the DIAAS system, ileal digestibility of the essential amino acids is emphasised as a more accurate measure of protein absorption. For this reason, DIAAS is promoted as the superior method and preferable over the PDCAAS (FAO, 2013).

As the technological landscape rapidly evolves and integrates with healthcare, potential solutions to this dilemma have begun to emerge. Machine learning (ML) and large language models (LLMs) are both subsets of artificial intelligence (AI) but differ in their scope and application. ML involves algorithms that enable computers to learn from data and make predictions or decisions without being explicitly programmed. These algorithms can be trained on various tasks, such as image recognition, natural language processing (NLP) and predictive analytics (Hao and Li, 2022). On the other hand, LLMs like ChatGPT 3.5 and 4, Bard AI and Bing Chat are a specific type of ML model that focusses on understanding and generating human-like text. These models are trained on vast amounts of text data to learn the nuances of language, enabling them to generate coherent and contextually relevant responses (Hamaniuk, 2021). These AI models go beyond being just algorithms and represent the pinnacle of human creativity and technological progress. These models have the ability to analyse vast amounts of data with remarkable accuracy, learn complex patterns and produce consistent results (Tai, 2020). Whilst the integration of AI into various sectors such as finance and transport is well documented, its potential in healthcare is arguably the most revolutionary (Aung *et al.*, 2021; Qarajeh *et al.*, 2023).

It is essential that accurate but practical methods are available for the food industry, clinicians, public health practitioners, policymakers and food aid organisations to ensure that consumers can make the best food and dietary choices to meet their optimal nutritional needs. Consumers also need to be made more aware of the importance of high-quality protein in order to be able to choose the best food product from the mass options on the market. AI models, equipped with extensive repositories of information and advanced algorithms, possess the capability to aid in intricate tasks, including nutritional recommendations (Qarajeh *et al.*, 2023). Also, from the users' point of view, AI often has more biological and social characteristics (Poria *et al.*, 2017). At the same time, the positive effects of high AI competence on human–AI interactions have been emphasised (Jarrahi, 2018). This situation defines AI as a very powerful technique used to influence consumers' reactions and decisions (Wang *et al.*, 2022). The high accuracy of the PQ content of food items by AI assistants could encourage the public to practice sustainable nutrition more accurately. Therefore, this study aimed to assess the effectiveness of different AI models in accurately aggregating information about the PQ content of food items with four AI models – ChatGPT 3.5, ChatGPT 4, Bard AI and Bing Chat.

2. Method

2.1 Materials and study procedures

In this study, we evaluated the effectiveness of AI models in accurately aggregating information about the PQ content in food items, which is a vital aspect of the plant-based diet recommended as it positively affects both health and the environment. The AI models analysed include ChatGPT 3.5, ChatGPT 4, Bard AI and Bing Chat.

2.2 Selection of dietary items for protein quality

The FAO recommended DIAAS score was used to assess PQ. DIAAS scores >3 years of age (rest of the population) were used to analyse recommendations for adults (FAO, 2013).

To conduct this study, we carefully compiled a range of 22 foods (Supplementary Table 1). We selected the types of foods according to the literature, which analysed the DIAAS score of foods (Ertl *et al.*, 2016; Fanning *et al.*, 2015; FAO, 2013; Herreman *et al.*, 2020; Phillips, 2017; Rutherfurd *et al.*, 2015). Additionally, since food processing affects the DIAAS score, we searched for such processed foods separately in AI. For example, some foods lose their PQ when they are processed (e.g. cooked, ground into flour, etc.). However, some of them remain unchanged. For this reason, the nutrients that do not change are searched with a single name (e.g. soy, soy flour, soy protein isolate or soybean), which do not change in terms of PQ. Wheat, wheat flour, wheat bran, corn, corn grain, corn-based cereal and rice-cooked rice do not change PQ). However, pea has high PQ, whilst cooked pea has low PQ. Therefore, we searched for peas and cooked them separately.

2.3 Assessment of AI model performance and repeated analysis

Each of the selected AI models was tasked with the responsibility of categorising the curated food items based on their PQ. This categorisation procedure involved classifying the items into distinct categories: those possessing low (DIAAS score <75), high (DIAAS score 75–99) or excellent (DIAAS score ≥100) PQ content (FAO, 2013). The main accuracy rate means that the AI assistant correctly predicted the PQ of food items according to the DIAAS score. For the purpose of generating responses from the AI models, the following prompts were utilised:

> The prompt is that: Classify the following as low, high, or excellent protein-quality food: ___.
>
> (For example: Classify the following as low, high, or excellent protein-quality food: corn?)

Each AI model was tasked with categorising the selected foods according to PQ on each date:

> ChatGPT 3.5, ChatGPT 4, Bard AI, and Bing Chat: first session on 15.10.2023, second session on 29.10.2023, third session on 14.11.2023. The same prompt was performed the second and third time for all AI models.

Due to the frequent updating of the AI models, a two-week period ensures consistency in the evaluation by minimising the risk of significant model changes. It also allows us to account for possible short-term fluctuations in AI responses due to transient technical factors (Qarajeh *et al.*, 2023). Therefore, in order to reduce the chance of results and analyse their consistency, also taking into account the temporal fluctuations of the AI models, the whole methodology was repeated three times, with a four-week interval between each instance.

2.4 Comparative analysis with established references

The database for DIAAS is not yet complete. Therefore, the results obtained by the AI models were then compared with the DIAAS scores obtained from the literature (Ertl *et al.*, 2016; Fanning *et al.*, 2015; FAO, 2013; Herreman *et al.*, 2020; Phillips, 2017; Rutherfurd *et al.*, 2015). The DIAAS scores were categorised according to the FAO: <75 (no quality claim), 75–99 (high-quality protein) and ≥100 (excellent quality protein) (FAO, 2013).

2.5 Quantitative and quality analysis

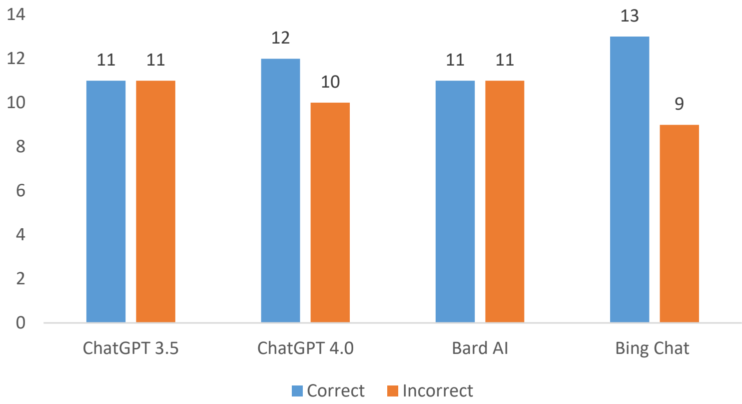
The data were analysed using SPSS 24.0. We used descriptive statistical techniques to summarise how the AI models performed in categorising the food items. Specifically, percentages and frequencies were calculated to quantify the accuracy of the categorisations for low, high and excellent PQ. Cohen's kappa statistic was used to measure the level of agreement between the different entities involved between Al models and DIAAS scores, according to the literature.

3. Results

The results of the first analysis revealed that Bing Chat accurately identified 59.1% of the food items (13 out of 22). Specifically, it correctly categorised 46.2% of foods as low PQ, 80.0% as high quality and 75.0% as excellent quality. ChatGPT 4 accurately identified 54.5% of food items (12 out of 22). It included 30.8% accuracy for low PQ, 100.0% for high PQ and 75.0% for excellent PQ (Supplementary Table 2). ChatGPT 3.5 and Bard AI showed similar results, correctly categorising 50.0% of food items (11 out of 22) (Figure 1).

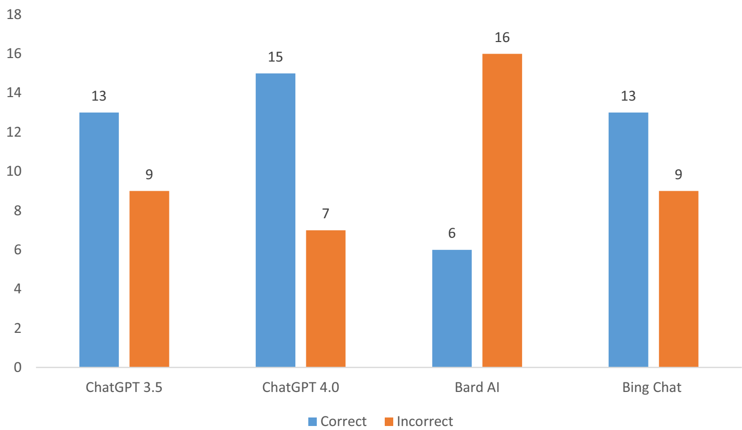
The results of the second analysis showed that ChatGPT 4 was the AI model that gave the most accurate results with 68.2% (15 out of 22), which correctly categorised 53.8% of foods as low PQ, 100.0% as high quality and 75.0% as excellent quality. ChatGPT 3.5 and Bing Chat accurately identified 59.1% of food items (13 out of 22). ChatGPT 3.5 included 61.5% accuracy for low PQ, 60.0% for high PQ and 50.0% for excellent PQ. These values were 61.5, 80.0 and 25.0%, respectively, according to Bing Chat (Supplementary Table 3). However, Bard Al had the lowest correct results rate with 27.3% (Figure 2).

The results of the third analysis showed that Bing Chat was the AI model that gave the most accurate results with 72.7% (16 out of 22), which correctly categorised 84.6% of foods as low PQ, 80.0% as high quality and 25.0% as excellent quality. ChatGPT 3.5 and ChatGPT 4



Source(s): Authors' work

Figure 1. Number of correct and incorrect answers of AI models according to the first analysis



Source(s): Authors' work

Figure 2. Number of correct and incorrect answers of AI models according to the second analysis

accurately identified 59.1% of food items (13 out of 22). ChatGPT 3.5 included 61.5% accuracy for low PQ, 40.0% for high PQ and 75.0% for excellent PQ. These values were 53.8, 60.0 and 75.0%, respectively, according to ChatGPT 4 (Supplementary Table 4). However, Bard AI had the lowest correct results rate with 45.5% (Figure 3).

According to Cohen's kappa analysis, ChatGPT 4 (Cohen's kappa: 0.718, p < 0.001) and ChatGPT 3.5 (Cohen's kappa: 0.636, p: 0.002) showed substantial agreement between baseline and 2nd analysis, whereas there was no agreement in Bard AI and Bing Chat. However, there was substantial agreement for Bard AI (Cohen's kappa: 0.715, p: 0.001) between the baseline and 3rd analysis, and ChatGPT 4 (Cohen's kappa: 0.538, p: 0.011) and ChatGPT 3.5 (Cohen's kappa: 0.455, p: 0.030) showed a moderate agreement. There was no agreement in Bing Chat (Table 1).

4. Discussion

The use of AI assistants is rapidly increasing and permeates many aspects of people's daily lives in both personal and professional settings (Makridakis, 2017; Olhede and Wolfe, 2018).

Source(s): Authors' work

Figure 3. Number of correct and incorrect answers of AI models according to the third analysis

Table 1. Cohen's kappa values of Al models between analyses

	Cohen's kappa[1], p^1-value	Cohen's kappa[2], p^2-value
ChatGPT 3.5	0.636, *0.002**	0.455, *0.030**
ChatGPT 4	0.718, < *0.001***	0.538, *0.011**
Bard AI	0.364, 0.056	0.727, *0.001**
Bing Chat	0.248, 0.245	0.108, 0.595

Note(s): $^*p < 0.05$, $^{**}p < 0.001$; [1]Differences between baseline and 2nd analysis; [2]Differences between baseline and 3rd analysis
Source(s): Authors' work

The development of AI assistants also offers new opportunities for research on nutrients and medical sensing technology, which are important for nutritional/health implication for individuals (Sak and Suchodolska, 2021). Today, AI-driven assistants are becoming increasingly prevalent understanding their strengths and limitations in categorising the nutrient content of food items is crucial for their responsible integration into daily life as well as clinical care. This study explored, for the first time, the capabilities of four well-known AI assistants, including ChatGPT 3.5, ChatGPT 4, Bard AI and Bing Chat and provided insights into their potential to enhance nutritional knowledge. Generative AI (particularly Bing Chat) was moderately effective in providing essential amino acid bioavailability information. Bing Chat was the most accurate AI assistant, with a correct response rate of 59.1% in the first analysis and 72.7% when the same questions were asked 4 weeks later, according to the results of 4 AI assistants classifying the PQ of 22 food items. In all analyses, Bard AI was the AI assistant with the least correct answers. ChatGPT 4 and ChatGPT 3.5 showed substantial agreement between baseline and 2nd analysis. In contrast, there was a moderate agreement for ChatGPT 4 and ChatGPT 3.5 between baseline and 3rd analysis, whereas Bard AI showed substantial agreement. There was no agreement in Bing Chat.

In recent years, new protein sources have emerged to support the transition to more sustainable food production for human consumption. Animal-based proteins, such as those derived from milk and meat, have been shown to have a significant impact on GHGE and are known to contribute to the depletion of natural sources (Aiking, 2014; Godfray *et al.*, 2018) and hazards for human health related to non-communicable diseases (WHO, 2021). These proteins are widely available in Western diets, and their consumption is on the rise in developing countries (Godfray *et al.*, 2018). It has been recommended that global protein consumption shift towards plant-based proteins (WHO, 2021; Willett *et al.*, 2019). However, attention should be paid to the nutritional quality of alternative protein sources (Herreman *et al.*, 2020). Therefore, it is important to consider protein quantity and PQ in sustainable diets. The FAO has recently recommended the newer DIAAS to supersede PDCAAS, which is promoted as the superior method and preferable over the PDCAAS (FAO, 2013). In this study, Bing Chat was the most accurate AI assistants with an average accuracy rate of 63.6% of PQ in food items, followed by ChatGPT 4 with 60.6%. Despite the notable progress seen in models such as ChatGPT 4 and Bing Chat, errors still occur in almost at least one out of every two foods or more. Hurrying to integrate AI into nutritional advice could potentially cause harm by making inappropriate recommendations (Qarajeh *et al.*, 2023).

The AI company OpenAI developed ChatGPT, a conversational interface based on GPT 3.5 (GPT 4 as of March 14, 2023, only for paid users). ChatGPT can generate texts on a wide variety of topics. In terms of nutritional advice, ChatGPT can inform users on how to eat healthy, suggest diets, plan recipes, etc. (Chatelan *et al.*, 2023). A study reported that ChatGPT 4 and Bing Chat demonstrated the highest accuracy rates, exceeding 80%, in effectively categorising potassium content compared to ChatGPT 3.5 and Bard AI (Qarajeh *et al.*, 2023). In another study, where 33 physicians subjectively classified 284 medical questions from 17 specialities as easy, medium or difficult with binary (yes/no) or descriptive answers, ChatGPT was found to be largely accurate (Johnson *et al.*, 2023). In this study, in the first analysis, Chat GPT 4 and ChatGPT 3.5 answered 54.5% and 50% of the foods correctly in terms of PQ compared to the DIAAS scores. In the second analysis, these values increased to 68.2 and 59.1%, respectively; however, in the third analysis, the accuracy of ChatGPT 4 decreased to 59.1%, and ChatGPT 3.5 showed a similar result with 59.1%. Additionally, ChatGPT 4 had the highest accuracy rate in the second analysis. According to Cohen's kappa analysis, ChatGPT 4 and ChatGPT 3.5 showed substantial agreement between baseline and 2nd analysis, whereas after 4-week later, they showed moderate agreement. Thus, the increase in the number of correct answers from both AI assistants can be seen as promising in terms of nutritional recommendations.

However, the limitations of ChatGPT are numerous. First, ChatGPT can "lie", i.e. it can produce reasonable-sounding but incorrect answers, https://platform.openai.com/docs/models as seen above. In this study, ChatGPT 4 gave 45.5% and ChatGPT 3.5 gave 50% wrong answers in the first analysis for PQ prediction. In the first analysis, ChatGPT 3.5 and 4 classified 69.2% of the food items with low PQ as high PQ. In the second and third analyses, these values were 38.5% and 46.2%, respectively. Secondly, ChatGPT makes up or distorts facts. For example, the made-up reference is known as the hallucination problem in AI. Thirdly, ChatGPT does not have an evidence-based hierarchy of sources. In fact, for nutrition topics where there are many non-evidence-based viewpoints expressed on the web, ChatGPT could spread widespread falsehoods and unscientific opinions (Chatelan *et al.*, 2023). Fourthly, ChatGPT has difficulty establishing causal relationships, for example, in clinical reasoning (Cascella *et al.*, 2023).

Bing Chat, developed by Microsoft, has been optimised for web-based interactions and tends to generate concise and direct responses. Bing Chat's conversational interface style seems to mimic that of ChatGPT, and unlike many other chatbots, it can also specify its sources (Xuan-Quy *et al.*, 2023). A study reported that Bing Chat showed the highest accuracy rates in categorising high phosphorus and potassium content (Qarajeh *et al.*, 2023). In this study, Bing Chat gave the highest number of correct answers at 59.1% in the first analysis, gave the same number of correct answers at 59.1% in the second analysis, and the accuracy rate increased to 72.7% for PQ of food items in the third analysis. When the average of the results of the analyses performed at two-week intervals was examined, Bing Chat gave almost a similar number of correct answers as ChatGPT 4 (the accuracy rate was 63.6 and 60.6%, respectively). This model, with its swift processing capabilities, could be particularly beneficial in scenarios where rapid information retrieval is essential.

Bard is Google's experimental, conversational AI chat service. It is meant to function similarly to ChatGPT, with the most significant difference being that BARD pulls its information from the web (Tisman and Seetharam, 2023). Bard Al was found to have the highest accuracy in categorising high phosphorus content and the lowest accuracy in categorising potassium content (Qarajeh *et al.*, 2023). It was observed that Bard consistently provided inaccurate information even after multiple tests, and Bard could not achieve 100% accuracy in its responses (TS2 Space Sp, 2023). In this study, although Bard AI gave a 50% accuracy rate of PQ of food items in the first analysis, it showed a serious decrease in the number of correct answers by giving 27.3% correct answers in the second analysis. However, the accuracy rate increased after 4 weeks with 45.5%. Additionally, there was a substantial agreement for Bard AI between the baseline and 3rd analysis. Google Bard fell short in accuracy and knowledge compared to ChatGPT and Bing Chat. When Bard AI was asked to classify PQ, it was based on the DIAAS scores, which may explain the low accuracy. Future studies could delve into the underlying factors contributing to the models' successes and limitations, offering a more comprehensive understanding of their functioning.

Two factors are crucial in the development of AI assistants: ML and NLP. ML enables the assistant to learn and improve its performance over time, whilst NLP enables it to understand and process human language. ML algorithms allow AI assistants to analyse data, identify patterns and make predictions or decisions based on the information they have learnt (González García *et al.*, 2019; Sak and Suchodolska, 2021). Transductive Support Vector Machine, one of the ML algorithms, predicts the categories of future instances in AIs based on input instances. In this way, this learning attempts to create groups by labelling the elements of the received inputs (González García *et al.*, 2019). Therefore, AI assistants may have misclassified foods for which no input data was available, such as faba beans, roasted peanuts and almonds.

Additionally, PQ predictions for some foods, such as potatoes, were significantly less accurate than others. The lower accuracy of AI assistants in evaluating the PQ content of foods like potatoes can be attributed to the intricate composition of potato proteins, the lack of comprehensive data on novel protein sources and the inherent complexity involved in assessing the nutritional quality of certain food components (Fritsch *et al.*, 2017). Further research and development in AI technology tailored to handle the complexities of diverse food compositions are essential to improve the accuracy of PQ assessments for various food types.

The present study has some limitations. Firstly, the evaluation of AI models was conducted within a controlled experimental environment using a literature search of the PQ of food items. The real-world diversity of foods and variations in nutrient content may not have been fully represented, potentially impacting the models' performance in practical scenarios. Second, ChatGPT, Bard AI, and Bing Chat showed varying degrees of accuracy. The reasons for these discrepancies have not been studied in detail. Third, this study primarily focussed on the models' accuracy in nutrient PQ classification, leaving aside potential considerations such as user experience, usability and the integration of AI recommendations. These limitations underscore the need for ongoing research and refinement in the application of AI models.

5. Conclusion

This study provides a first insight into how accurately emerging AI assistants assess PQ content in food items. Although ChatGPT 4 showed the most improvement in the analysis with a two-week interval, the average number of correct answers in both analyses was similar to Bing Chat. Moreover, only ChatGPT 3.5 and 4 showed substantial agreement between the two analyses. Google Bard fell short in terms of accuracy and information compared to ChatGPT and Bing Chat. Whilst the results demonstrate the potential of AI, they also emphasise the need for human oversight. However, provided that transparency, accuracy, consistency and validation against clinical standards continue to improve, AI can help as a complementary tool to enhance nutrition education and counselling whilst easing the workload on healthcare teams rather than replacing healthcare professionals. Further research into the real-world implementation of AI for nutritional advice is essential as the technology develops.

References

Aiking, H. (2014), "Protein production: planet, profit, plus people?", *The American Journal of Clinical Nutrition*, Vol. 100 No. suppl_1, pp. 483S-489S, doi: 10.3945/ajcn.113.071209.

Allès, B., Baudry, J., Méjean, C., Touvier, M., Péneau, S., Hercberg, S. and Kesse-Guyot, E. (2017), "Comparison of sociodemographic and nutritional characteristics between self-reported vegetarians, vegans, and meat-eaters from the NutriNet-Santé study", *Nutrients*, Vol. 9 No. 9, p. 1023, doi: 10.3390/nu9091023.

Aung, Y.Y., Wong, D.C. and Ting, D.S. (2021), "The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare", *British Medical Bulletin*, Vol. 139 No. 1, pp. 4-15, doi: 10.1093/bmb/ldab016.

Bayram, H.M. and Ozturkcan, S.A. (2022), "Greenhouse gas emissions in the food system: current and alternative dietary scenarios", *Mediterranean Journal of Nutrition and Metabolism*, Vol. 15 No. 4, pp. 463-477, doi: 10.3233/mnm-220006.

Cascella, M., Montomoli, J., Bellini, V. and Bignami, E. (2023), "Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios", *Journal of Medical Systems*, Vol. 47 No. 1, p. 33, doi: 10.1007/s10916-023-01925-4.

Chatelan, A., Clerc, A. and Fonta, P.A. (2023), "ChatGPT and future artificial intelligence chatbots: what may be the influence on credentialed nutrition and dietetics practitioners?", *Journal of the Academy of Nutrition and Dietetics*, Vol. 123 No. 11, pp. 1525-1529, doi: 10.1016/j.jand.2023.08.001.

Clune, S., Crossin, E. and Verghese, K. (2017), "Systematic review of greenhouse gas emissions for different fresh food categories", *Journal of Cleaner Production*, Vol. 140, pp. 766-783, doi: 10.1016/j.jclepro.2016.04.082.

El Sadig, R. and Wu, J. (2024), "Are novel plant-based meat alternatives the healthier choice?", *Food Research International*, Vol. 183, 114184, doi: 10.1016/j.foodres.2024.114184.

Ertl, P., Knaus, W. and Zollitsch, W. (2016), "An approach to including protein quality when assessing the net contribution of livestock to human food supply", *Animal*, Vol. 10 No. 11, pp. 1883-1889, doi: 10.1017/s1751731116000902.

Fanning, A.C., Miller, B.J. and Moughan, P.J. (2015), "Protein digestibility-corrected amino acid scores and digestible indispensable amino acid scores differentially describe protein quality in growing male rats", *The Journal of Nutrition*, Vol. 145 No. 2, pp. 372-379, doi: 10.3945/jn.114.195438.

FAO (2012), "Sustainable Diets and Biodiversity. Directions and solutions for policy, research and action", available at: https://www.fao.org/3/i3004e/i3004e00.htm/ (accessed 1 November 2023).

FAO (2013), "Report of an FAO Expert Consultation: dietary protein quality evaluation in human nutrition", available at: https://www.fao.org/ag/humannutrition/35978-02317b979a686a57aa4593304ffc17f06.pdf/ (accessed 2 November 2023).

Fritsch, C., Staebler, A., Happel, A., Cubero Márquez, M.A., Aguiló-Aguayo, I., Abadias, M., Gallur, M., Cigognini, I.M., Montanari, A., Lopez, M.J., Suarez-Estrella, F., Brunton, N., Luengo, E., Sisti, L., Ferri, M. and Belotti, G. (2017), "Processing, valorization and application of bio-waste derived compounds from potato, tomato, olive and cereals: a review", *Sustainability*, Vol. 9 No. 8, p. 1492, doi: 10.3390/su9081492.

Fructuoso, I., Romão, B., Han, H., Raposo, A., Ariza-Montes, A., Araya-Castillo, L. and Zandonadi, R.P. (2021), "An overview on nutritional aspects of plant-based beverages used as substitutes for cow's milk", *Nutrients*, Vol. 13 No. 8, p. 2650, doi: 10.3390/nu13082650.

Godfray, H.C.J., Aveyard, P., Garnett, T., Hall, J.W., Key, T.J., Lorimer, J., Pierrehumbert, R.T., Scarborough, P., Springmann, M. and Jebb, S.A. (2018), "Meat consumption, health, and the environment", *Science*, Vol. 361 No. 6399, eaam5324, doi: 10.1126/science.aam5324.

González García, C., Núñez Valdéz, E.R., García Díaz, V., Pelayo García-Bustelo, B.C. and Cueva Lovelle, J.M. (2019), "A review of artificial intelligence in the internet of things", *International Journal Of Interactive Multimedia and Artificial Intelligence*, Vol. 5 No. 4, pp. 9-20, doi: 10.9781/ijimai.2018.03.004.

Hamaniuk, V.A. (2021), "The potential of Large Language Models in language education", *Educational Dimension*, Vol. 5, pp. 208-210, doi: 10.31812/ed.650.

Hao, C. and Li, Y. (2022), "Simulation of English speech recognition based on improved extreme random forest classification", *Computational Intelligence and Neuroscience*, Vol. 2022, pp. 1948159-10, doi: 10.1155/2022/1948159.

Herreman, L., Nommensen, P., Pennings, B. and Laus, M.C. (2020), "Comprehensive overview of the quality of plant-And animal-sourced proteins based on the digestible indispensable amino acid score", *Food Science and Nutrition*, Vol. 8 No. 10, pp. 5379-5391, doi: 10.1002/fsn3.1809.

Jarrahi, M.H. (2018), "Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making", *Business Horizons*, Vol. 61 No. 4, pp. 577-586, doi: 10.1016/j.bushor.2018.03.007.

Johnson, D., Goodman, R., Patrinely, J., Stone, C., Zimmerman, E., Donald, R., Chang, S., Berkowitz, S., Finn, A., Jahangir, E., Scoville, E., Reese, T., Friedman, D., Bastarache, J., van der Heijden, Y., Wright, J., Carter, N., Alexander, M., Choe, J., Chastain, C., Zic, J., Horst, S., Turker, I., Agarwal, R., Osmundson, E., Idrees, K., Kieman, C., Padmanabhan, C., Bailey, C., Schlegel, C., Chambless, L., Gibson, M., Osterman, T. and Wheless, L. (2023), "Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model", *Research Square*, Preprint, doi: 10.21203/rs.3.rs-2566942/v1.

Makridakis, S. (2017), "The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms", *Futures*, Vol. 90, pp. 46-60, doi: 10.1016/j.futures.2017.03.006.

Marques, A.C., Fuinhas, J.A. and Pais, D.F. (2018), "Economic growth, sustainable development and food consumption: evidence across different income groups of countries", *Journal of Cleaner Production*, Vol. 196, pp. 245-258, doi: 10.1016/j.jclepro.2018.06.011.

OECD and FAO (2023), "OECD-FAO agricultural Outlook 2023-2032", available at: https://www.fao.org/documents/card/en/c/cc6361en/ (accessed 2 November 2023).

Olhede, S.C. and Wolfe, P.J. (2018), "The growing ubiquity of algorithms in society: implications, impacts and innovations", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 376 No. 2128, 20170364, doi: 10.1098/rsta.2017.0364.

Phillips, S.M. (2017), "Current concepts and unresolved questions in dietary protein requirements and supplements in adults", *Frontiers in Nutrition*, Vol. 4, p. 13, doi: 10.3389/fnut.2017.00013.

Pointke, M. and Pawelzik, E. (2022), "Plant-based alternative products: are they healthy alternatives? Micro- and macronutrients and nutritional scoring", *Nutrients*, Vol. 14 No. 3, p. 601, doi: 10.3390/nu14030601.

Poria, S., Cambria, E., Bajpai, R. and Hussain, A. (2017), "A review of affective computing: from unimodal analysis to multimodal fusion", *Information Fusion*, Vol. 37 No. 98, p. 125, doi: 10.1016/j.inffus.2017.02.003.

Qarajeh, A., Tangpanithandee, S., Thongprayoon, C., Suppadungsuk, S., Krisanapan, P., Aiumtrakul, N., Valencia, O.A.G., Miao, J., Qureshi, F. and Cheungpasitporn, W. (2023), "AI-powered renal diet support: performance of ChatGPT, bard AI, and bing chat", *Clinics and Practice*, Vol. 13 No. 5, pp. 1160-1172, doi: 10.3390/clinpract13050104.

Rojas Conzuelo, Z., Robyr, R. and Kopf-Bolanz, K.A. (2022), "Optimization of protein quality of plant-based foods through digitalized product development", *Frontiers in Nutrition*, Vol. 9, 902565, doi: 10.3389/fnut.2022.902565.

Rutherfurd, S.M., Fanning, A.C., Miller, B.J. and Moughan, P.J. (2015), "Protein digestibility-corrected amino acid scores and digestible indispensable amino acid scores differentially describe protein quality in growing male rats", *The Journal of Nutrition*, Vol. 145 No. 2, pp. 372-379, doi: 10.3945/jn.114.195438.

Sak, J. and Suchodolska, M. (2021), "Artificial intelligence in nutrients science research: a review", *Nutrient*, Vol. 13 No. 2, p. 322, doi: 10.3390/nu13020322.

Smetana, S., Ristic, D., Pleissner, D., Tuomisto, H.L., Parniakov, O. and Heinz, V. (2023), "Meat substitutes: resource demands and environmental footprints", *Resources, Conservation and Recycling*, Vol. 190, 106831, doi: 10.1016/j.resconrec.2022.106831.

Sobiecki, J.G., Appleby, P.N., Bradbury, K.E. and Key, T.J. (2016), "High compliance with dietary recommendations in a cohort of meat eaters, fish eaters, vegetarians, and vegans: results from the European Prospective Investigation into Cancer and Nutrition–Oxford study", *Nutrition Research*, Vol. 36 No. 5, pp. 464-477, doi: 10.1016/j.nutres.2015.12.016.

Soedamah-Muthu, S.S. and De Goede, J. (2018), "Dairy consumption and cardiometabolic diseases: systematic review and updated meta-analyses of prospective cohort studies", *Current Nutrition Reports*, Vol. 7 No. 4, pp. 171-182, doi: 10.1007/s13668-018-0253-y.

Tai, M.C.T. (2020), "The impact of artificial intelligence on human society and bioethics", *Tzu-Chi Medical Journal*, Vol. 32 No. 4, p. 339, doi: 10.4103/tcmj.tcmj_71_20.

Tisman, G. and Seetharam, R. (2023), "OpenAi's ChatGPT-4, BARD and YOU. Com (AI) and the cancer patient, for now, caveat emptor, but stay tuned", *Digital Medicine and Healthcare Technology*, Vol. 2 No. 1, pp. 1-27, doi: 10.5772/dmht.19.

TS2 Space Sp (2023), "Google bard falls short in accuracy and knowledge compared to ChatGPT and bing chat", available at: https://ts2.space/en/google-bard-falls-short-in-accuracy-and-knowledge-compared-to-chatgpt-and-bing-chat/#:~:text=Even%20after%20multiple%20tests%2C%20Bard,the%20capabilities%20of%20certain%20smartphones/ (accessed 2 November 2023).

United Nations (2019), "World population prospects 2019", available at: https://population.un.org/wpp/Publications/Files/WPP2019_DataBooklet.pdf/ (accessed 2 November 2023).

Vranken, L., Avermaete, T., Petalios, D. and Mathijs, E. (2014), "Curbing global meat consumption: emerging evidence of a second nutrition transition", *Environmental Science and Policy*, Vol. 39, pp. 95-106, doi: 10.1016/j.envsci.2014.02.009.

Wang, B., Rau, P.-L.P. and Yuan, T. (2022), "Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale", *Behaviour and Information Technology*, Vol. 42 No. 9, pp. 1324-1337, doi: 10.1080/0144929x.2022.2072768.

Werner, L.B., Flysjö, A. and Tholstrup, T. (2014), "Greenhouse gas emissions of realistic dietary choices in Denmark: the carbon footprint and nutritional value of dairy products", *Food and Nutrition Research*, Vol. 58, pp. 1-16, doi: 10.3402/fnr.v58.20687.

WHO (2021), "Plant-based diets and their impact on health, sustainability and the environment: a review of the evidence: WHO European Office for the Prevention and Control of Noncommunicable Diseases", available at: https://www.who.int/europe/publications/i/item/WHO-EURO-2021-4007-43766-61591 (accessed 2 November 2023).

WHO/FAO/UNU Expert Consultation, Bend, J., Bolger, M., Knaap, A.G.A.C., Kuznesof, P.M., Larsen, J.C., Mattia, A., Meylan, I., Pitt, J.I., Resnik, S., Schlatter, J., Vavasour, E., Rao, M.V., Verger, P., Walker, R., Wallin, H., Whitehouse, B., Abbott, P.J., Adegoke, G., Baan, R., Baines, J., Barlow, S., Benford, D., Bruno, A., Charrondiere, R., Chen, J., Choi, M., DiNovi, M., Fisher, C.E., Iseki, N., Kawamura, Y., Konishi, Y., Lawrie, S., Leblanc, J.C., Leclercq, C., Lee, H.M., Moy, G., Munro, I.C., Nishikawa, A., Olempska-Beer, Z., de Peuter, G., Pronk, M.E.J., Renwick, A.G., Sheffer, M., Sipes, I.G., Tritscher, A., Soares, L.V., Wennberg, A. and Williams, G.M. (2007), "Protein and amino acid requirements in human nutrition", *World Health Organization Technical Report Series*, Vol. 935 No. 947, pp. 1-265.

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L.J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J.A., De Vries, W., Majele Sibanda, L., Afshin, A., Chaudhary, A., Herrero, M., Agustina, R., Branca, F., Lartey, A., Fan, S., Crona, B., Fox, E., Bignet, V., Troell, M., Lindahl, T., Singh, S., Cornell, S.E., Srinath Reddy, K., Narain, S., Nishtar, S. and Murray, C.J.L. (2019), "Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems", *The Lancet*, Vol. 393, pp. 447-492, 10170, doi: 10.1016/s0140-6736(18)31788-4.

Xuan-Quy, D., Ngoc-Bich, L., Bac-Bien, N. and Xuan-Dung, P. (2023), "LLMs' capabilities at the high school level in chemistry: cases of ChatGPT and Microsoft bing chat", *Journal of Chemical Education*, Preprint.

Supplementary Material
The supplementary material for this article can be found online.

Corresponding author
Arda Ozturkcan can be contacted at: sozturkcan@gelisim.edu.tr