

Article

Chatbots in Household Budget Management: A Comparative Study in the Context of Economic Vulnerability

Eulália Santos ^{1,*}, Corália Pimenta ^{2,3} and Fernando Oliveira Tavares ^{4,5}

¹ Centre for Organizational and Social Studies of Polytechnic of Porto (CEOS.PP), Higher Institute of Accounting and Administration of Coimbra (ISCAC), Polytechnic University of Coimbra, Rua da Misericórdia, Lagar dos Cortiços, S. Martinho do Bispo, 3045-093 Coimbra, Portugal

² Centre for Research and Innovation in Education (inED), School of Education, Polytechnic Institute of Porto, 4200-465 Porto, Portugal; coralia.pimenta@isec.pt

³ Coimbra Institute of Engineering (ISEC), Polytechnic University of Coimbra, Rua da Misericórdia, Lagar dos Cortiços, S. Martinho do Bispo, 3045-093 Coimbra, Portugal

⁴ Research on Economics, Management and Information Technologies, Department of Economics and Management, Universidade Portucalense, 4200-072 Porto, Portugal; ftavares@upt.pt

⁵ Instituto Superior Miguel Torga, Largo da Cruz de Celas No. 1, 3000-132 Coimbra, Portugal

* Correspondence: emsantos@iscac.pt

Abstract

Household budget management is a central element of financial literacy, recognized by the Organisation for Economic Co-operation and Development as a key determinant of financial inclusion, economic stability, and social participation. Despite efforts to promote financial education, significant weaknesses remain, reinforcing the need to resort to other accessible solutions in the current digital context, such as chatbots. Within this framework, this study analyses and compares the responses produced by different chatbots in solving a budget management task aimed at a family experiencing economic vulnerability. The results indicate that, overall, the models analysed demonstrated the ability to interpret the context and provide recommendations aligned with the principles of financial literacy, although differences were observed in the accuracy of calculations, clarity of responses, and suitability of the proposals to the family's specific circumstances. Overall, the findings suggest that chatbots, when properly mediated by the user and used critically, can serve as complementary tools to support budget management and the development of financial literacy, particularly relevant for families in situations of greater economic vulnerability, although the information and recommendations generated should be verified in light of the limitations identified in the study.

Keywords: budget management; financial literacy; digital literacy; chatbots; artificial intelligence



Academic Editor: Thanasis Stengos

Received: 31 July 2026

Revised: 18 August 2026

Accepted: 19 August 2026

Published: 1 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The financial market is in a state of permanent transformation, driven by technological innovation and increasing digitalization, offering a wide diversity of financial products that, although globally appealing and accessible, may entail heightened risks. This evolution has contributed to greater complexity within the financial system, requiring higher levels of knowledge and responsibility from ordinary citizens in decision-making, particularly in managing savings, essential for addressing unforeseen situations and future needs such as retirement (Tavares et al., 2020). This complexity has received particular attention from the Organisation for Economic Co-operation and Development (OECD), which has emphasized

the importance of financial literacy as a determining factor for financial inclusion, economic stability, and full participation in social life (OECD, 2014, 2017). Through instruments such as the Programme for International Student Assessment (PISA), the OECD has highlighted persistent weaknesses in financial literacy levels. The results of PISA 2018, for example, showed that many 15-year-olds were already using digital financial products, often without adequate preparation to deal with risks associated with security, privacy, and impulsive decision-making, with students displaying lower levels of financial literacy being more vulnerable to fraud. Subsequently, data from PISA 2022 revealed that around 18% of students failed to reach the basic level of proficiency in financial literacy, while higher levels of literacy were associated with more financially responsible behaviours, such as saving, comparing prices, and making informed choices (OECD, 2024).

Overall, the literature recognizes that financial literacy is a fundamental element for making more conscious and effective economic decision-making, with a direct impact on family well-being. The lack of basic financial knowledge may compromise the management of daily life and affect more acutely those groups in situations of greater social and economic vulnerability, such as individuals with low levels of education or reduced income (Lusardi, 2008), reinforcing the relevance of financial education in school contexts (Silva & Powell, 2015).

The development and dissemination of systems based on Artificial Intelligence (AI) have been widely debated in both public and academic spheres, and are being integrated into various sectors of society, given their relevance in terms of competitiveness, as highlighted by different professionals in the technology field (Zhu et al., 2026). In parallel, the literature has emphasized the potential of AI-based technologies to support economic decision-making, particularly in the domain of financial management, suggesting that these tools may contribute to financial planning and management through recommendations, whose effectiveness appears to depend on users' levels of financial literacy and digital skills (Chumakova et al., 2019). The CashSave system constitutes an example of such an application, having shown positive results in supporting financial management, although limitations associated with authentication and accessibility remain (Wong & Salleh, 2023). This evidence has led several authors to advocate for the need to articulate financial education with the strategic use of digital technologies, as a means of promoting individual autonomy and informed economic decision-making, especially in contexts of greater social and economic vulnerability (Lusardi, 2008; Chatterjee et al., 2017). Within this framework, effective budget management emerges as a central dimension of economic and social well-being, particularly for families experiencing economic vulnerability, and as such, the present article analyses and compares the responses produced by different chatbots to a household budgeting task, seeking to assess their suitability for analysing the family's financial situation and planning the budget of a family experiencing economic vulnerability.

The article is organised as follows: after this introduction, a literature review on financial literacy in the digital age is presented, addressing its concepts and evolution, as well as the impact of financial literacy on decision-making in vulnerable contexts. Subsequently, the potential of digital technologies in budget management is discussed, along with the specific role of chatbots as tools to support financial literacy. Thereafter, the methodology adopted in the study is described, the results obtained are presented and discussed, and finally, the main conclusions and implications of the study are outlined.

2. Financial Literacy in the Digital Age

2.1. Financial Literacy as an Evolving Concept

The concept of financial literacy has evolved within the OECD since the early 2000s. The first official definition emerged in 2005, when it was understood as a process through

which consumers and investors improve their understanding of financial concepts and products, developing the skills and confidence needed to make informed decisions, identify risks, and improve their financial well-being (Silva & Powell, 2015). At this stage, the definition mainly focused on the educational process and on the implementation of strategies and programs aimed at building the capacity of citizens. Subsequently, with the development of public policies and the need to assess the impact of financial education, the OECD and the International Network on Financial Education (INFE) began to use the term financial literacy as the outcome of that educational process.

In the case of adults, it was defined as the combination of awareness, knowledge, attitudes, and behaviours necessary to make informed and sustainable financial decisions, and in the case of young people, it includes not only knowledge and understanding of financial concepts and risks, but also the motivation and confidence required to apply this knowledge in different contexts, aiming at individual and social well-being (Silva & Powell, 2015), based on the results of PISA 2012. It should be noted that both concepts coexist, with financial education being the primary pathway for promoting financial literacy. In this study, financial literacy is understood as the set of knowledge, skills, and attitudes necessary to make appropriate decisions in the management of personal and family resources.

2.2. Impact of Financial Literacy on Decision-Making in Vulnerable Contexts

Financial literacy plays a central role in how individuals make decisions in their daily economic lives, influencing choices that have a direct impact on their security and well-being throughout life. Some studies indicate that mastery of concepts such as compound interest and risk management is essential for conscious financial administration, enabling not only better preparation for retirement but also the gradual accumulation of wealth (Lusardi, 2008). In this sense, financial literacy goes beyond mere technical knowledge, constituting a practical tool that guides more informed and responsible decisions, such as building savings for unexpected situations or for long-term goals (Chatterjee et al., 2017; Santos et al., 2022). However, this capability is not widespread, and significant differences in levels of financial literacy can be observed, conditioned by social and educational factors. Individuals with higher levels of education and income tend to demonstrate more developed financial skills, which is reflected in more prudent economic behaviours.

Conversely, social inequality and high levels of household indebtedness are often associated with low levels of financial literacy, limiting the capacity to manage the household budget and to assess financial risk (Nogueira et al., 2025). For more vulnerable groups, these limitations translate into less effective financial decisions and a reduced capacity to accumulate wealth, perpetuating situations of economic fragility (Lusardi, 2008). In this context, the acquisition of financial literacy emerges as a factor of change with a direct impact on individuals' daily lives, particularly for those facing greater economic constraints. The development of these skills leads to more conscious financial choices, reflected in the adoption of more regular saving practices, the creation of emergency funds, and more structured retirement planning (Chatterjee et al., 2017). Beyond these behavioural changes, improvements in financial literacy have been associated with reductions in excessive indebtedness and increases in economic well-being, also contributing to the strengthening of social inclusion. Thus, more than an educational purpose, financial literacy emerges as a strategic instrument of individual capacity building, capable of reducing vulnerabilities, promoting autonomy in financial decision-making, and contributing to more sustainable and equitable development (Santos et al., 2022; Gomes et al., 2024).

Recognising the importance of financial literacy for all citizens, the present study aims to analyse how chatbots can help families, especially the most vulnerable, to manage their household budgets.

2.3. Digital Technologies in Budget Management

Household budget management can be understood as a process that involves planning, organizing, and controlling the financial resources of a household in order to ensure a balance between income, expenses, and short- and long-term economic goals. However, it is not merely a technical process, as it also includes behavioural dimensions such as discipline and self-control (Brounen et al., 2016; Chumakova et al., 2019; Gomes et al., 2024). Within this context, models have been developed that seek to support more effective financial management, particularly the 50/30/20 rule, which proposes allocating monthly net income into three categories: 50% for essential needs such as housing, food, transportation, and health; 30% for discretionary expenses associated with leisure, culture, and non-essential consumption; and 20% for savings or debt repayment, contributing to financial balance and the creation of safety margins for unexpected situations.

Digital technologies play an important role in household budget management, and there are currently several applications that guide users toward saving and controlling expenses (Wong & Salleh, 2023; Chumakova et al., 2019). These tools can be particularly relevant for low-income families, who often face a disproportion between income and expenses and have little room to adjust spending. This situation frequently leads to the reduction in essential expenses and the inability to build emergency funds or plan for retirement (Lusardi, 2008; Chatterjee et al., 2017; Martins et al., 2021; Moura et al., 2024; Gomes et al., 2024).

The development of financial literacy can be promoted through inclusion policies and the use of digital technologies, which may help strengthen financial autonomy, reduce social vulnerabilities, and expand individuals' capacity for financial planning (Lusardi, 2008; Brounen et al., 2016; Chatterjee et al., 2017; Gomes et al., 2024).

2.4. Chatbots as Tools to Support Financial Literacy

The development of digital literacy has been assuming an increasingly relevant role in mitigating the risks associated with fraud and exclusion, while artificial intelligence tools contribute to the dissemination of knowledge, articulating with financial literacy in the promotion of inclusion and individual autonomy (Vieira et al., 2024). The advancement of artificial intelligence, particularly generative AI and language models, as well as the development of chatbots, has facilitated interaction between people and computers, bringing benefits to various areas, with particular emphasis on education and business.

In this context, chatbots stand out for their potential to support learning processes, analysis, and decision-making in the financial domain. Solutions such as ChatGPT, Gemini, Microsoft Copilot, Claude, DeepSeek-Chat, and Amália illustrate different approaches to the use of these technologies, ranging from pedagogical support and the contextualization of financial information to data analysis, budget organisation, and the promotion of digital inclusion. Despite their individual specificities, these tools share the common objective of facilitating access to financial knowledge in an interactive and personalized way, contributing to the development of financial literacy and supporting household budget management, especially in contexts of greater vulnerability. Through these solutions, users can understand economic concepts in an accessible way, explore financial management scenarios, receive savings recommendations tailored to their needs, and define short- and long-term financial goals.

Chatbots possess a set of skills relevant for solving tasks, particularly in terms of linguistic fluency, pattern recognition, and the generation of responses in natural language. They have the potential to reformulate questions and infer semantic relationships, producing quick and intuitive responses. However, they do not possess true knowledge in the traditional sense, as the information they provide is predominantly based on statis-

tical associations, which may compromise the reliability of the content they present and autonomous decision-making.

Additionally, recent studies have highlighted structural limitations of these models in managing extensive contexts. [Liu et al. \(2023\)](#) demonstrate that language models tend to perform less effectively when the information needed to generate an answer is distributed throughout a text, particularly when it appears in the middle of the context, as opposed to the beginning or the end. This limitation may be particularly relevant in tasks applied to financial management, such as budget simulations, in which it is necessary to integrate multiple data points and maintain internal coherence across several informational elements.

Some studies suggest that integrating chatbots with structured external knowledge and semantically aware reasoning mechanisms (Semantic-Aware Reasoning, SAR) can mitigate some of these limitations, significantly improving the accuracy and relevance of responses. In this framework, chatbots are more effective when integrated into hybrid systems complemented by semantic-aware reasoning and explicit external knowledge and explicit factual knowledge ([Sornlertlamvanich, 2025](#)).

Considering the characteristics of the present study, where the development of financial literacy is articulated with the application of mathematical knowledge, it is also important to highlight that the literature has emphasized the relevant role of artificial intelligence in this domain. In particular, the possibility of users learning at their own pace and benefiting from learning experiences tailored to individual needs has been associated with a better understanding of mathematical concepts, as well as an increase in students' confidence and motivation ([Opesemowo & Adewuyi, 2024](#)).

3. Methodology

3.1. Task Design and Development

Based on the literature review on budgetary financial literacy and conversational artificial intelligence, as well as on institutional sources such as the Social Security system, the platform Simula ([Simula.pt, n.d.](#)), and the Porta 65 program, incorporating specific aspects of legislation and support measures available in Portugal for families in situations of socioeconomic vulnerability, an exploratory task was developed focusing on the application of financial literacy knowledge to household budget management. This task focuses on the analysis, planning, and management of the budget of a household in a vulnerable situation, with the goal of assessing its financial sustainability and supporting decision-making. In this study, an analysis is conducted of the way in which different chatbots respond to the challenges inherent in this task, assessing the suitability of the responses produced to support household budget analysis and management. The task, entitled "Managing Geo's Family Budget" is based on a fictional yet realistic situation describing the financial context of a household in a vulnerable situation in Portugal, including simulated information on income, expenses, social benefits, and the legislative framework. The task statement includes data from institutional sources, particularly the Social Security system and the Porta 65 program. Solving the task requires exploring, interpreting, and articulating financial information, involving the analysis of different types of income (subsidies and social benefits) and fixed monthly expenses (housing and essential services), as well as considering social tariffs and specific support measures.

The task content refers to the situation of a young couple living in Portugal with a minor child, facing unemployment, limited income, and access to several social support measures (such as reintegration benefits, child allowance, and social tariffs for basic services, among others). The task is divided into three parts. The first part focuses on analysing the family's income, expenses, and monthly balance, that is, calculating and diagnosing their current financial situation. The second part examines the family's ability to plan

future expenses (such as insurance), save, invest, and possibly donate, assessing financial stability and the sustainability of the household budget. The third and final part involves a critical evaluation of the family budget in light of the good practices recommended by the [Simula.pt](#) (n.d.).

3.2. Chatbots and Application Procedure

Solving the task requires chatbots to demonstrate logical reasoning skills, financial interpretation, and clear communication, allowing an evaluation of the responses produced by the chatbots in addressing practical problems with social relevance. To this end, the following models were used: ChatGPT ([OpenAI, 2023](#)), Gemini ([Google, 2024](#)), Microsoft Copilot ([Microsoft, 2023](#)), DeepSeek-Chat ([DeepSeek, 2024](#)), Claude 3 ([Anthropic, 2024](#)), and Amália—Lisbon’s Virtual Assistant ([Câmara Municipal de Lisboa & Unbabel, 2023](#)). Amália was included due to the contextualised nature of the task, which involves a Portuguese family, public support measures, and local financial and social information, without implying technological equivalence between the systems. The six chatbots were accessed through their respective web platforms, using the free versions available at the time of the study, without access to subscriptions or paid versions.

In all cases, the following prompt was used: “Can you solve in detail the following task related to managing the budget of Geo’s family:”, followed by the pasted text of the task presented in Appendix A. The task was carried out using the six chatbots on 1 August 2025 and was performed once for each chatbot, with no follow-up prompts, clarification questions, or additional processes for refining the responses.

3.3. Evaluation Grid and Assessment Procedures

To evaluate the responses produced by the chatbots in solving the task, an evaluation grid was developed and is presented in Table 1. It comprises five dimensions considered relevant to household budget management problems: accuracy, clarity, contextual appropriateness, didactic usefulness, and completeness. Each dimension was operationalised through three criteria, resulting in a total of 15 evaluation criteria.

Table 1. Evaluation grid for chatbot responses.

Dimension	Evaluation Criteria
Accuracy	C1: Calculations performed correctly C2: Conclusions consistent with the results C3: Accurate financial information
Clarity	C4: Clear language C5: Well-organised response C6: Reasoning presented clearly
Contextual appropriateness	C7: Appropriateness of the analysis to the family’s financial situation C8: Appropriateness of the recommendations regarding the family’s financial capacity C9: Consideration of the support measures and constraints presented
Didactic usefulness	C10: Financial concepts explained clearly C11: Calculations presented clearly C12: Financial recommendations justified
Completeness	C13: Response to all elements requested C14: Use of the relevant information provided C15: Appropriate level of detail

The evaluation focused on the complete response of each chatbot to the task, considering its different parts and questions jointly in order to obtain an overall assessment of

the response for each criterion. Each criterion was assessed on a five-point ordinal scale: 1—Very inadequate, 2—Inadequate, 3—Partially adequate, 4—Adequate, and 5—Very adequate. The score was assigned according to the degree to which each criterion was met, with higher scores indicating a more favourable assessment of the response.

To ensure consistent application of the evaluation grid, the three evaluators used the same definitions and rating levels presented in Table 1. The three evaluators were the three authors of the present study. Each evaluator reviewed the responses of all six chatbots in full and independently assigned a score from 1 to 5 for each of the 15 criteria, without knowledge of the assessments provided by the other evaluators.

Following the independent assessment, the ratings were compared across the three evaluators, and any discrepancies were jointly discussed to clarify potential differences in interpretation.

3.4. Data Analysis Procedures

For each chatbot, the mean score for each dimension was calculated based on the ratings assigned by the three evaluators to the three criteria comprising that dimension. The overall mean for each chatbot was calculated as the mean of the five dimensional means, assigning equal weight to the five dimensions. The quantitative analysis therefore enabled comparisons of the chatbots across the 15 criteria and the five evaluation dimensions. The ratings originally assigned by the evaluators were retained for the purposes of the quantitative analysis. The standard deviation was calculated to characterise the dispersion of the aggregated ratings.

In addition to the quantitative analysis of the ratings, a descriptive and comparative analysis of the content of the responses produced by the different chatbots was conducted. This analysis focused on the different parts and questions of the task, considering aspects such as the interpretation of the socioeconomic context, the accuracy of the calculations, the identification and organisation of income and expenditure, the calculation of the monthly balance, the planning of future expenses, saving and investment strategies, the possibility of making donations, recommendations for preparing the household budget, and the application of the recommendations provided by the *simula.pt* platform.

The comparative analysis identified similarities and differences between the models, considering, in particular, calculation errors or correct calculations, omissions of relevant information, the appropriateness of recommendations to the family's financial situation, ways of presenting calculations, and differences in the level of detail and contextualisation of the responses. Where relevant, specific characteristics of how the responses were presented were also analysed, namely the use of different mathematical representations and the organisation of information.

The analysis of the responses was based on the content actually produced by each chatbot, thereby complementing the quantitative results of the evaluation grid with a descriptive and comparative assessment of the characteristics of the responses and the differences observed between the models.

4. Results and Discussion

This section describes, interprets, and discusses the responses of the different chatbots to the challenges presented in the task, considering the theoretical references and the adopted framework. The analysis of the results combines the quantitative assessment of the responses, conducted using the 15-criterion evaluation grid applied by the three evaluators, with a descriptive and comparative analysis of the responses produced by the six chatbots.

4.1. Evaluation of Chatbot Responses: Analysis by Criteria and Dimensions

The analysis at the criterion level shows an overall high assessment of the chatbots' responses, although differences can be observed across the aspects evaluated (Table 2). The criteria with the highest mean scores were C2: Conclusions consistent with the results and C5: Well-organised response, both with a mean of 4.89 ($SD = 0.17$). Criteria C4: Clear language and C13: Response to all elements requested also received high ratings ($M = 4.83$, $SD = 0.18$). Criterion C8: Appropriateness of the recommendations regarding the family's financial capacity received the lowest mean score ($M = 4.06$, $SD = 0.25$). This result is particularly relevant, as it shows that, despite the overall high ratings, the appropriateness of financial recommendations to the family's specific financial situation was the weakest aspect of the responses. Criteria C7: Appropriateness of the analysis to the family's financial situation and C15: Appropriate level of detail also received relatively lower ratings ($M = 4.50$, $SD = 0.35$).

Table 2. Evaluation of chatbot responses across criteria and chatbots.

Criteria	ChatGPT	Gemini	Copilot	DeepSeek	Claude	Amália	$M \pm SD$
C1	4.00	5.00	4.00	5.00	5.00	5.00	4.67 ± 0.52
C2	4.67	5.00	4.67	5.00	5.00	5.00	4.89 ± 0.17
C3	4.67	4.67	4.33	4.67	4.67	4.33	4.56 ± 0.17
C4	5.00	4.67	5.00	4.67	4.67	5.00	4.83 ± 0.18
C5	5.00	4.67	5.00	4.67	5.00	5.00	4.89 ± 0.17
C6	5.00	4.67	4.67	4.67	4.67	5.00	4.78 ± 0.17
C7	4.67	4.00	4.33	4.67	4.33	5.00	4.50 ± 0.35
C8	4.00	3.67	4.00	4.33	4.00	4.33	4.06 ± 0.25
C9	4.67	4.33	4.33	4.67	4.67	5.00	4.61 ± 0.25
C10	5.00	4.67	4.67	4.67	4.67	5.00	4.78 ± 0.17
C11	5.00	4.67	4.67	4.67	4.67	5.00	4.78 ± 0.17
C12	4.67	4.67	4.67	4.33	4.33	5.00	4.61 ± 0.25
C13	5.00	4.67	5.00	4.67	4.67	5.00	4.83 ± 0.18
C14	4.67	4.33	4.67	4.67	5.00	5.00	4.72 ± 0.15
C15	4.33	4.33	4.00	4.67	4.67	5.00	4.50 ± 0.35

The dispersion of the criterion ratings across the six chatbots was low, with standard deviations ranging from 0.17 to 0.35, indicating that the mean ratings assigned to the different models showed relatively small differences for most criteria.

The analysis of the five dimensions reveals some differences between the chatbots (Table 3). In the Accuracy dimension, the highest ratings were obtained by Gemini, DeepSeek, and Claude ($M = 4.89$). For Clarity, ChatGPT and Amália received the maximum rating ($M = 5.00$). Amália obtained the highest ratings for Contextual appropriateness ($M = 4.78$), Didactic usefulness ($M = 5.00$), and Completeness ($M = 5.00$). Overall, Amália achieved the highest overall mean score ($M = 4.91$, $SD = 0.12$).

Table 3. Evaluation of chatbot responses across dimensions and chatbots.

Chatbot	Accuracy	Clarity	Contextual Appropriateness	Didactic Usefulness	Completeness	$M \pm SD$
ChatGPT	4.44	5.00	4.44	4.89	4.67	4.69 ± 0.25
Gemini	4.89	4.67	4.00	4.67	4.44	4.53 ± 0.34
Copilot	4.33	4.89	4.22	4.67	4.56	4.53 ± 0.27
DeepSeek	4.89	4.67	4.56	4.56	4.67	4.67 ± 0.14
Claude	4.89	4.78	4.33	4.56	4.78	4.67 ± 0.22
Amália	4.78	5.00	4.78	5.00	5.00	4.91 ± 0.12

The analysis of the five dimensions presented in Figure 1 suggests an overall high assessment of the chatbots' responses, with mean scores ranging from 4.39 to 4.83. Clarity received the highest mean score ($M = 4.83$, $SD = 0.15$), followed by Didactic usefulness ($M = 4.72$, $SD = 0.18$), Accuracy ($M = 4.70$, $SD = 0.25$), and Completeness ($M = 4.69$, $SD = 0.19$). Contextual appropriateness received the lowest mean score ($M = 4.39$, $SD = 0.27$), although it remained above the midpoint of the scale.

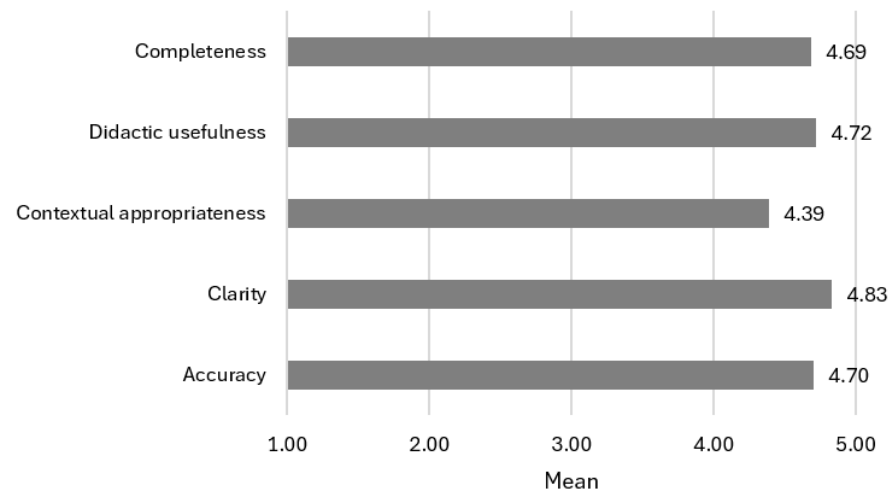


Figure 1. Overall evaluation of chatbot responses by dimension.

The highest rating was obtained for the Clarity dimension, suggesting that the chatbots were generally able to organise and communicate financial information in an understandable way, facilitating the interpretation of the calculations and recommendations. This result is relevant in the context of financial literacy, where understanding information is important for supporting more informed economic decisions (Lusardi, 2008; Vieira et al., 2024).

In contrast, the Contextual appropriateness dimension received the lowest rating, particularly for the criterion concerning the appropriateness of the recommendations to the family's financial capacity, suggesting greater difficulty in adapting the guidance to the household's specific circumstances. This issue is particularly relevant in situations of economic vulnerability, where recommendations should consider the family's specific needs and constraints. The literature highlights the importance of financial literacy in shaping saving behaviours that are aligned with households' circumstances (Lusardi, 2008; Chatterjee et al., 2017), while limitations in language models' ability to integrate multiple pieces of contextual information may hinder the production of responses that are fully adapted to complex situations (Sornlertlamvanich, 2025).

4.2. Interpretation of the Socioeconomic Context of the Task

Regarding the interpretation of the task statement, the different chatbots were able to establish a correct semantic reading of the presented context, recognizing that it referred to a household in a situation of socioeconomic vulnerability. They were able to correctly frame specific terms associated with this context, such as public daycare, social reintegration benefit, first child benefit bracket, risk of extreme poverty, and child guarantee fund, as well as concepts related to budget management, including fixed, variable, and unexpected expenses, savings, investment, balance, house rent, social tariff, annual insurance, insurance premium, financial stability, budget, and retirement savings. This interpretation results from the semantic processing of the context presented in the task by the chatbots. However, it is important to consider that literature recognises limitations associated with the reasoning

and reliability of the responses generated by these systems, particularly in tasks requiring the integration of multiple pieces of contextual information (Sornlertlamvanich, 2025).

Regarding the use of mathematical language, namely numerical and symbolic formatting, some differences were observed among the models. Microsoft Copilot used the full stop as a decimal separator, following the Anglo-Saxon convention, whereas the other models used the decimal comma according to the Portuguese standard. Microsoft Copilot was also the only model to place the euro symbol at the beginning of the number, while all the others placed the euro symbol at the end. In terms of the representation of multiplication, the traditional multiplication symbol ($\times$) predominated, although Amália and Gemini used the asterisk (*), which is more common in programming languages but less intuitive for the general public. In practical terms, these differences may affect the clarity, readability, and suitability of the responses for the target audience, particularly when the level of financial literacy is still limited.

4.3. Determination of the Family's Monthly Income

In solving question 1.1, which consisted of determining the family's monthly income, all chatbots correctly identified the different components of income. Claude was the only model that did not explicitly present the sum of the different amounts used to obtain the final income calculation. However, not all models added the amounts correctly: ChatGPT and Copilot reported a total of €644.96, whereas the correct value, based on the amounts provided in the task, is €643.96. Thus, both models made a €1 error in calculating the total monthly household income. An interesting aspect of the response generated by ChatGPT was the use of the expressions "70% of 237.25 €" and "50% of 237.25 €" to calculate the income of Geo's partner and child, respectively. Although the final result was incorrect, this form of expression is advantageous from a didactic perspective, since the use of the preposition "of" in mathematics indicates multiplication. This language can facilitate the understanding of the type of operation required, particularly for individuals who experience difficulties interpreting percentages in the context of financial problems. It is important to note that learning percentages is associated with persistent difficulties, as it involves proportional relationships that are not always evident in their symbolic notation. This opacity makes it difficult to identify the reference quantity, correctly interpret the "%" symbol, and understand the underlying proportional relationship, often leading to the application of mechanical procedures without conceptual understanding (Lestiana, 2021).

4.4. Calculation of Fixed Expenses for Essential Services

In question 1.2, the chatbots were challenged to calculate the family's monthly expenses related to essential services (housing, electricity, water, and gas), considering the social support measures specified in the task statements. All models correctly identified the four relevant expenses and applied some form of calculation to obtain the total monthly expenses. Microsoft Copilot presented a well-organized table with columns for the base value, reduction percentage, and final value, listing the four expenses in rows; however, it did not explicitly show the calculations performed.

The house rent, subject to a 50% reduction under the Porta 65 program, was correctly adjusted by all models, although with variations in presentation. Claude was more direct, indicating the initial rent value, mentioning the support, and presenting the final value. ChatGPT, DeepSeek-Chat, and Gemini used the mathematical expression " 0.5×200 ", whereas Amália used the expression " $(1 - 0.5) \times 200$ ".

Regarding electricity, all models applied the multiplication factor 0.662, corresponding to the amount to be paid after applying the support measures. Claude and DeepSeek used this multiplication directly. Amália and Gemini described the process in two steps:

" $50.00 \text{ €} * (1 - 0.338) = 50.00 \text{ €} * 0.662 = 33.10 \text{ €}$ ". ChatGPT also performed the calculation in two steps: first calculating the reduction, " $50 \times 0.338 = 16.90 \text{ €}$ ", and then subtracting it from the initial value, " $50 - 16.90 = 33.10 \text{ €}$ ". Both approaches are appropriate and useful from a didactic perspective. The method used by Amália and Gemini highlights the origin of the multiplication factor by directly applying the formula " $\text{value} \times (1 - \text{percentage})$ ", promoting an algebraic understanding of the reduction. By contrast, the approach used by ChatGPT isolates the reduction before applying it, which may be more intuitive for those who are beginning to work with percentages. Both approaches make the reasoning more transparent than simply applying the factor 0.662 directly. These characteristics have been highlighted in the literature, with some researchers arguing that such approaches can contribute to improving the understanding of ideas while fostering users' confidence and motivation (Opesemowo & Adewuyi, 2024).

Regarding water, the 50% reduction was correctly interpreted by all models, although only some explicitly indicated the values and the multiplication operation. Claude and ChatGPT presented the final value while indicating the reduction. Amália and Gemini detailed the calculation steps " $20.00 \text{ €} * (1 - 0.50) = 10.00 \text{ €}$ ", whereas DeepSeek-Chat performed the calculation " $20 \times 0.50 = 10 \text{ €}$ ".

For gas, all models applied the correct reduction (multiplying the value by 0.688). The response patterns remained consistent with those observed for electricity.

Finally, regarding the total expense value, only Claude presented the final amount without explicitly showing the sum of the partial values.

4.5. Estimation of Variable Expenses and Organisation of Categories

In question 1.3, the models were asked to estimate other monthly household expenses beyond the fixed ones. All simulated expenses generated by the models are presented in Table 4, with values ranging from €335 (Microsoft Copilot) to €470 (Gemini). These differences result from the number and type of categories considered, as well as from the values assigned to each of them.

Regarding the organization of the responses, all models presented the values in a structured and clear way. Claude and Amália chose to categorize expenses into groups such as "fixed and additional," "variable," or "unexpected," which facilitated reading and comprehension. ChatGPT, Microsoft Copilot, DeepSeek-Chat, and Gemini used tables, varying in how they classified and described the categories.

In terms of identifying and describing expenses, all models provided at least a brief explanation for each item included. Amália, Microsoft Copilot, DeepSeek-Chat, and Gemini were the most complete, identifying and describing all listed expenses.

Despite the diversity of categories and values presented in Table 4, none of the models explicitly considered the possibility of support from institutions such as Caritas, Food Banks Against Hunger, or other social organisations that frequently assist families in situations of socioeconomic vulnerability. This omission reveals a common limitation among the analysed models: the lack of integration of real social and institutional aspects, particularly those that directly affect consumption habits and the basic needs of low-income families. For example, the chatbots did not explore the possibility that the amounts allocated to food could be supplemented through food assistance provided by charitable organisations, namely through the provision of essential goods such as rice, pasta, beans, or canned food, which could constitute a relevant alternative in a context of economic vulnerability. In such cases, the monetary expenditure on food may be significantly lower than the estimates produced by the models.

Table 4. Simulation of monthly variable expenses (in euros) suggested by the models.

	Claude	Amália	ChatGPT	Microsoft Copilot	DeepSeek	Gemini
Telecommunications (mobile phone/internet)	25	40		25	30	30
Basic health insurance	15					
Public transportation	40	50	60	40	50	50
Food	180	250	250	180	200	300
Clothing	30			25		
Clothing and footwear (occasional)		20	25		30	30
Health (pharmacy, non-reimbursed consultations)		30	30	20	30	20
Household hygiene		25				
Hygiene and health	25					
Leisure and personal expenses						20
Leisure/entertainment	20					
Leisure				15	20	
Leisure/small emergencies		10				
Leisure/communications			20			
School materials/daycare (extra expenses beyond public daycare)		15				
Child expenses (diapers, etc.)	25				40	20
Car insurance					15	
Unexpected expenses					50	
Emergency fund		10				
Minimum monthly savings			50	30		
Minor repair/medication			20			
Home maintenance (occasional)		5				
Reserve for unforeseen expenses	20					
Total	380	455	455	335	465	470

Including these types of support would have a direct impact on estimating expenses related to food and hygiene, for example, and highlights the importance of calibrating models with knowledge of the local context. This gap illustrates a broader limitation of chatbots: although they are capable of generating correct responses and structuring lists of expenses coherently, they tend to produce generalized results without considering more specific real-world information unless it is explicitly provided by the user. The reliability and relevance of responses can be improved when chatbots are complemented by structured external knowledge and semantic validation mechanisms (Sornlertlamvanich, 2025). Thus, the results reinforce the importance of calibrating the use of these models with local and institutional information, particularly when tasks involve estimating consumption and basic needs in situations of vulnerability.

Another relevant aspect concerns transportation. All models referred to public transportation, seemingly overlooking an essential piece of information included in the task, specifically in question 2.1, which states that in the following month, the family will need to pay the annual car insurance. Only DeepSeek-Chat mentioned this insurance, suggesting a value of €15, although without further elaboration. This point highlights a limitation related to the short-term memory or internal coherence of the models: since they had access to the full text of the task, it would be expected that they could maintain consistency of information across its different sections. This point highlights a difficulty in maintaining the consistency of information across the different parts. As noted by Liu et al. (2023), the ability to maintain coherent information throughout long, multi-step tasks remains a challenge for language models, which often struggle to retain and use information located in different parts of the context, particularly in the middle sections of the text.

4.6. Calculation of the Family's Monthly Balance

In question 1.4, the models calculated the family's monthly balance by subtracting total expenditure from the income considered in each response. The values obtained were generally consistent with the expenditure estimates provided by each model in the preceding questions. However, it should be noted that ChatGPT and Copilot made an error in calculating the family's total income, using €644.96 instead of the correct amount of €643.96, which affects the accuracy of their final balances. The differences in the balances show that small differences in the income considered can also affect the final assessment of the family's financial capacity.

The final balances ranged from €10.22 (Gemini) to €146.22 (Microsoft Copilot), reflecting the differences in the estimated amounts for monthly expenses. This dispersion highlights how the assumptions adopted during the expense estimation phase (question 1.3) significantly influence the outcome.

4.7. Car Insurance Planning and Savings Strategies

In question 2.1, which required estimating an approximate value for annual car insurance and providing savings advice, all models demonstrated the ability to interpret the request appropriately, recognising the importance of planning for this periodic expense and proposing saving strategies.

The values assigned to the insurance varied across models. Claude estimated a value between €300 and €400 per year, proposing a monthly saving of €35. Amália considered a value of €300 (third-party insurance), suggesting a monthly saving of €25, with emphasis on financial discipline and risk management. One of the most relevant aspects of its response was the recommendation that any surplus should be directed toward an emergency fund.

ChatGPT estimated the insurance at €250 and suggested saving around €25 per month for insurance and other emergencies, with recommendations for simple, low-risk investment options, including savings certificates, savings accounts, and even low-risk Retirement Savings Plans (PPRs). Microsoft Copilot and DeepSeek-Chat opted for a lower estimate (€180 per year), suggesting a monthly saving of €15, reflecting a more economical scenario adjusted to the available budget. DeepSeek adopted a pragmatic approach by including this value in the previous table of expenses (question 1.3) and recommended reducing spending in categories such as leisure or clothing. Gemini presented a range of values between €250 and €400 and, despite recognizing the family's limited monthly balance (€10.22), advised saving between €25 and €35, which would require reassessing budget priorities.

In general, the models recognised the need for regular saving to meet this periodic expense, although differences were observed in the appropriateness of the amounts suggested in relation to the family's financial capacity.

All models indicated that savings for car insurance should be made regularly and prioritized low-risk, high-liquidity solutions, such as savings accounts and savings certificates, as well as reductions in expenses focused on non-essential goods. Amália was particularly complete in distinguishing between savings for the insurance and the creation of a more robust emergency fund. This question made it possible to verify that, despite small differences in estimated values and levels of detail, all models demonstrated the ability to correctly identify the need for planned savings and recommended strategies tailored to the financial reality of a family with limited income. The results show that chatbots provided information consistent with the principles of financial literacy by recommending regular savings and prioritizing low-risk, high-liquidity financial instruments, such as savings accounts and savings certificates, in which predictability and liquidity play a central role in household budget management (Vieira et al., 2024). The recommendation to reduce non-essential expenses also reinforces the ability of chatbots to support informed

financial decision-making processes, promoting more conscious consumption practices. Despite variations in the suggested values and the level of detail in the responses, all models demonstrated skills in analysing the situation and generating recommendations appropriate to the presented context, highlighting their potential to support learning and the understanding of financial concepts (Opesemowo & Adewuyi, 2024).

4.8. Assessment of the Possibility of Donations

Question 2.2 invited the models to reflect on whether the family, considering its economic situation, could allocate part of its income to donations. All models reached the same essential conclusion: it is not advisable to make financial donations at this time, given the family's fragile budgetary situation and their risk of extreme poverty.

Claude, DeepSeek-Chat, and Microsoft Copilot adopted a concise and objective response, clearly stating that, given the scarcity of resources, the priority should be financial stability and the creation of an emergency fund. Claude further added that building financial reserves should come before any gesture of solidarity.

Amália and Gemini stood out for the richness of their argumentation. Amália contextualized the family's situation in detail, noting that the monthly balance of €25.22 is almost entirely absorbed by the need to save for car insurance, and that priorities such as food, health, transportation, and education must come before any intention to donate. Gemini followed a similar line of reasoning, reinforcing that ensuring basic needs and creating a financial reserve should be the family's main focus.

ChatGPT also shared this perspective, adding that at a later stage, when greater financial stability is achieved, small donations (around €1 to €2) could be considered. It also mentioned examples of organisations such as the Food Banks Against Hunger and Caritas.

Finally, Microsoft Copilot and Gemini referred to alternative forms of solidarity that do not involve financial expenditure, such as volunteering or donating time.

Overall, the models demonstrated a realistic reading of the family's economic situation. This interpretation results from chatbots' capabilities in interpreting user-provided information, recognising patterns, and inferring semantic relationships, allowing them to generate contextualized responses in natural language (Sornlertlamvanich, 2025).

4.9. Recommendations for Preparing a Household Budget

Question 2.3 focused on the importance of a household budget and how it could be recommended to the family. All the AI models analysed were unanimous in emphasising the fundamental relevance of a household budget, particularly in the context of limited income and economic vulnerability. However, the depth, clarity, and practical applicability of the suggestions varied significantly across the different chatbots.

The Amália model stood out for the detail and depth of its response. It highlighted that a budget is a tool for control, awareness, and planning, enabling the family to identify deficits and make informed decisions. It proposed a practical methodology adapted to the family's reality, including the careful categorization of expenses (fixed, variable, and unexpected), suggestions for simple tools (such as spreadsheets or free applications), and emphasised the importance of monthly reviews and the involvement of all family members. Claude and Gemini also provided structured responses, recognising the budget as an essential instrument to avoid debt and plan for the future. Both proposed concrete steps: recording income and expenses, categorising spending, setting savings goals, and conducting monthly reviews. Claude additionally suggested creating an emergency fund and using mobile applications as a practical recommendation. ChatGPT complemented these ideas by recommending saving 20% of income, although this value may be unrealistic in the context of Geo's family.

Microsoft Copilot presented a clear and objective response structured in two parts: importance and recommendations. It suggested the 50/30/20 rule as a budgeting methodology (essential expenses, discretionary spending, and savings), although this rule may not be feasible for low-income families.

DeepSeek-Chat, in turn, focused on debt prevention as the central objective of budgeting. It proposed practical measures such as the use of financial management applications (e.g., DABOX), reducing non-essential spending, and prioritizing savings for unforeseen expenses.

Overall, all models demonstrated an understanding of the strategic role of the household budget, particularly in contexts of economic risk. Responses ranged from more technical and detailed explanations (Amália and Gemini) to more concise and functional answers (DeepSeek-Chat and Microsoft Copilot). These results show that the chatbots were able to interpret the information provided and present contextualised financial recommendations aimed at prioritising basic needs, building savings, and preventing over-indebtedness, aspects addressed in the literature on financial literacy (Lusardi, 2008; Chatterjee et al., 2017; Vieira et al., 2024; Sornlertlamvanich, 2025).

4.10. Evaluation of the Household Budget Based on the Recommendations of the Simula Platform

The third part of the task focused on applying the recommendations of the Simula platform (Simula.pt, n.d.) to the financial situation of the family under study. This platform proposes an ideal division of monthly income into three categories: up to 50% for fixed expenses, up to 30% for variable expenses, and at least 20% for savings and investment. However, with a monthly income of €643.96, all models concluded that the family cannot fully meet these targets.

All models recognised that fixed expenses fell within the recommended limits, ranging between 25% and 38% of monthly income. This outcome is largely made possible by public support measures, such as the Porta 65 program and social tariffs for essential services.

However, variable expenses emerge as the most problematic element. The amount allocated to these expenses significantly exceeds the 30% recommended by Simula.pt (n.d.), ranging between €300 and €445, even though they mainly correspond to essential expenditures.

Savings are also well below the recommended level: the values range between €10 and €100, corresponding to only 1.5% to 15% of income, which makes it difficult to build an emergency fund.

Given this reality, the models recommend measures such as reducing non-essential expenses, starting gradual savings, and seeking additional sources of income. Claude was the only model that proposed a revised budget closer to the targets suggested by Simula.pt (n.d.), with 37.8% allocated to fixed expenses, 38.8% to variable expenses, and 23.3% to savings. Even so, the reduction in variable expenses remained insufficient to fully comply with the guidelines. The other models did not present alternative budget simulations, limiting themselves to a descriptive analysis of the expenses estimated in the previous question.

5. Conclusions

This study made it possible to compare the responses by different chatbots in solving a budget management task applied to a family in a situation of economic vulnerability. The results show that, despite differences in calculation accuracy, clarity of language, and level of detail in the responses, the models analysed were able to interpret the essential aspects of the socioeconomic context, structure the information, perform basic calculations,

and propose savings and budget organisation strategies aligned with the principles of financial literacy.

The ratings assigned by the three evaluators using the evaluation grid indicated an overall high assessment of the chatbots' responses, with Clarity receiving the highest mean score, while Contextual appropriateness received the lowest mean score, mainly due to the lower appropriateness of the recommendations to the family's specific financial capacity. However, the results should be interpreted as an assessment of the quality of the responses produced by the chatbots in the task analysed and do not constitute evidence that their use leads to actual improvements in users' financial literacy. In this sense, chatbots may serve as relevant aids in supporting financial literacy tasks, namely by helping families to navigate their economic decisions, organise income and expenditure, better understand the calculations involved, and explore saving strategies. This potential for support should not, however, be confused with a demonstrated effect on learning or the development of financial literacy, the assessment of which requires specific studies involving users.

The study also revealed several weaknesses, including occasional calculation errors, omissions of institutional support measures, difficulties in integrating contextual information, and limitations in maintaining coherence throughout extended tasks. These issues reinforce the need for mediated and critical use of such tools. Thus, the results do not allow us to conclude that chatbots, on their own, constitute sufficient tools for household budget management or the development of financial literacy, but they suggest potential for them to function as complementary tools, provided that they are used with critical verification and, where necessary, human mediation.

From a practical perspective, the results of this study suggest that chatbots can be integrated into financial education strategies, social inclusion programs, and initiatives aimed at supporting families, provided that they are accompanied by appropriate guidance and mechanisms that encourage the critical verification of the information provided. In educational contexts, these tools may be explored as complementary pedagogical resources, facilitating the understanding of financial and mathematical concepts and encouraging reflection on everyday economic decisions. Their use should, however, take into account the limitations identified, particularly the possibility of errors, omissions, or recommendations that are not well suited to the user's specific context.

Furthermore, the task analysed in this study shows potential for application in educational contexts. Considering the mathematical and financial concepts involved, such as percentage calculations, expense analysis, and household budget planning, it may be particularly suitable for secondary or higher education students and may also be adapted for other educational levels. In these contexts, the use of chatbots can support problem-solving activities, stimulate critical discussion about financial decisions, and contribute to the simultaneous development of mathematical reasoning and financial literacy. This pedagogical potential represents an avenue for application that should be subject to specific investigation and cannot be directly inferred as a learning effect from the findings of the present study.

From a theoretical perspective, the study contributes to the literature by deepening the understanding of the potential and limitations of systems based on artificial intelligence in the domain of financial literacy. It shows that the usefulness of these tools depends not only on their technical capabilities but also on how information is presented and the context in which they are used.

This study presents some limitations, namely the use of a single budget management scenario and a limited number of chatbot models, which restricts the generalization of the results. Moreover, each chatbot completed the task only once, with no multiple executions of the same scenario or comparisons of responses obtained at different points in time. Given

the non-deterministic nature of these systems and their ongoing development over time, the results therefore reflect the responses produced under the specific conditions and on the date on which the queries were conducted. Another limitation concerns access to external information in Part 3 of the task: although the task included the URL of the simula.pt platform, Internet access and actual consultation of this webpage were not controlled experimental conditions. Therefore, it cannot be ensured that all chatbots directly consulted this source in real time.

Future research may explore multiple family scenarios with different income levels, household compositions, and degrees of economic vulnerability, as well as analyse chatbot performance over longer interactions. It would also be relevant to conduct multiple executions of the same tasks, test different prompts and conditions for access to external information, and use more robust inter-rater assessment procedures, thereby enabling the stability and reproducibility of the responses obtained to be assessed. Additionally, it will be important to investigate the impact of using these tools on the actual development of users' financial literacy, as well as the role of human mediation in maximizing their benefits.

Author Contributions: Conceptualization, E.S., C.P. and F.O.T.; methodology, E.S.; software, E.S.; validation, E.S., C.P. and F.O.T.; formal analysis, E.S.; investigation, E.S., C.P. and F.O.T.; resources, E.S., C.P. and F.O.T.; data curation, E.S.; writing—original draft preparation, E.S., C.P. and F.O.T.; writing—review and editing, E.S., C.P. and F.O.T.; visualization, E.S., C.P. and F.O.T.; supervision, E.S., C.P. and F.O.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Appendix A

Task—Managing Geo's Family Budget

Geo is 24 years old, and his partner is 23. They live in Portugal and have recently completed their vocational training courses and are currently looking for employment. They have a 3-year-old child who attends a public daycare center.

Couple's Income:

- Geo receives the social reintegration benefit of €237.25.
- Geo's partner receives 70% of Geo's reintegration benefit.
- Because they have a child, they receive 50% of Geo's benefit.
- Geo's child has been assigned the 1st bracket of the child allowance, amounting to €72 per month.
- As the family is at risk of extreme poverty, they receive an additional €50 for the child, corresponding to the Child Guarantee Fund.

Couple's Expenses in February 2025:

House rent: €200 per month, but they benefit from a 50% reduction in rent because they applied to the Porta 65 program.

Electricity: €50, to which a 33.8% social tariff reduction will be applied.

Water: €20, to which a 50% social tariff reduction will be applied.

Gas: €30, to which a 31.2% social tariff reduction will be applied.

Part 1: Calculation and Financial Diagnosis of the Family

1.1. Based on the information provided, determine Geo's family income.

1.2. Calculate the fixed expenses listed above.

1.3. Identify other possible expenses (fixed, variable, unexpected, savings, or investments) that you consider relevant to simulate for this family. For each expense, indicate an approximate value, considering the current cost of living in Portugal.

1.4. Considering the previous items, determine the family's balance for the month under analysis.

Part 2: Financial Planning and Sustainability of the Household Budget

2.1. Considering that next month the family will need to pay the annual car insurance, estimate an approximate value for the insurance premium and indicate whether you would advise the family to save some money.

If so, how much should they save, and what investments would you suggest?

2.2. Do you believe that this family has sufficient financial stability to allocate part of its income to donations?

If so, what would be an appropriate amount, and which institutions could receive these donations?

2.3. Do you consider it important to prepare a household budget? How would you recommend that the family prepare it?

Part 3: Evaluation Based on Good Practices

The platform Simula is an independent service that provides free online simulators designed to help Portuguese citizens understand various fiscal and financial aspects. On this platform, it is possible to find recommendations and advantages associated with a well-prepared household budget.

In this part of the task, you are asked to consult <https://simula.pt/artigos/orcamento-familiar.html>, accessed on 1 August 2025. Prepare a budget for this family and determine whether it is possible for the family to follow the recommendations provided by the platform.

Consider the following guidelines:

- Fixed monthly expenses should ideally represent up to 50% of household income and include housing costs (rent or mortgage), essential services (water, electricity, gas, and telecommunications), insurance (home, car, health, and life), education (tuition, school materials, and training), and loan payments.
- Variable expenses should be limited to 30% of the monthly budget. These include spending on food, transportation, health, clothing, and leisure.
- Savings and investments should represent at least 20% of income. These may include retirement savings plans (PPR), savings certificates, or savings with specific purposes such as “saving for holidays” or “saving to renovate the house”, among others. Within these savings, there should be an emergency fund for unexpected situations, ideally corresponding to six times the amount of monthly fixed expenses.

References

- Anthropic. (2024). *Claude 3* [AI model family]. Available online: <https://www.anthropic.com/index/claude-3-family> (accessed on 1 August 2025).
- Brounen, D., Koedijk, K., & Pownall, R. (2016). Household financial planning and savings behavior. *Journal of International Money and Finance*, 69, 95–107. [CrossRef]
- Câmara Municipal de Lisboa & Unbabel. (2023). *Amália—Assistente virtual de Lisboa* [AI-powered chatbot]. Available online: <https://www.lisboa.pt/atualidade/noticias/detalhe/amalia-assistente-virtual-da-camara-de-lisboa> (accessed on 1 August 2025).
- Chatterjee, S., Fan, L., Jacobs, B., & Haas, R. (2017). Risk tolerance and goals-based savings behavior of households: The role of financial literacy. *Journal of Personal Finance*, 16(1), 66–77.
- Chumakova, T., Ivanova, M., & Kasyanov, I. (2019). Financial literacy and the need for financial education: Evidence and implications. *Swiss Journal of Economics and Statistics*, 155(1), 1–12. [CrossRef]

- DeepSeek. (2024). *DeepSeek-chat* [Open-source large language model]. Available online: <https://www.deepseek.com/> (accessed on 1 August 2025).
- Gomes, C. V., Borges Sales, K. M., Leite, I., & Nogueira Nunes, I. (2024). Financial education for decision making in family budget: A literature review from a behavioral perspective. *Environmental & Social Management Journal/Revista de Gestão Social e Ambiental*, 18(10), e09041. [CrossRef]
- Google. (2024). *Gemini* [Conversational AI]. Available online: <https://gemini.google.com/> (accessed on 1 August 2025).
- Lestiana, H. T. (2021). What are the difficulties in learning percentages? An overview of prospective mathematics teachers' strategies in solving percentage problems. *Indonesian Journal of Science and Mathematics Education*, 4(3), 331–343. [CrossRef]
- Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12, 157–173. [CrossRef]
- Lusardi, A. (2008). *Household saving behavior: The role of financial literacy, information, and financial education programs*. NBER working paper no. 13824. National Bureau of Economic Research. [CrossRef]
- Martins, R., Antunes, M., Silva, P. P., & Fortunato, A. (2021). Tarifa social de energia: Gênese, incidência e lições. *Notas Económicas*, (53), 85–102. [CrossRef] [PubMed]
- Microsoft. (2023). *Microsoft copilot* [AI assistant]. Available online: <https://copilot.microsoft.com/> (accessed on 1 August 2025).
- Moura, P., Fonseca, P., Cunha, I., & Morais, N. (2024). Diagnosing energy poverty in Portugal through the lens of a social survey. *Energies*, 17(16), 4087. [CrossRef]
- Nogueira, M. C., Almeida, L., & Tavares, F. O. (2025). Financial literacy, financial knowledge, and financial behaviors in OECD countries. *Journal of Risk and Financial Management*, 18(3), 167. [CrossRef]
- OECD. (2014). *PISA 2012 results: Students and money: Financial literacy skills for the 21st century* (Vol. VI). OECD Publishing. [CrossRef]
- OECD. (2017). *PISA 2015 results (Vol. IV): Students' financial literacy*. OECD Publishing. [CrossRef]
- OECD. (2024). *PISA 2022 results (Vol. IV): How financially smart are students?* OECD Publishing. [CrossRef]
- OpenAI. (2023). *ChatGPT* [Large language model]. Available online: <https://chat.openai.com/> (accessed on 1 August 2025).
- Opesemowo, O. A. G., & Adewuyi, H. O. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(7), em2478. [CrossRef] [PubMed]
- Santos, E., Ratten, V., Diogo, A., & Tavares, F. (2022). Positive and negative affect during the COVID-19 pandemic quarantine in Portugal. *Journal of Science and Technology Policy Management*, 13(2), 195–212. [CrossRef]
- Silva, A. M., & Powell, A. B. (2015). Educação financeira na escola: A perspectiva da organização para cooperação e desenvolvimento econômico. *Boletim Gepem*, (66), 3–19. [CrossRef]
- Simula.pt. (n.d.). *Orçamento familiar—Como criar um orçamento familiar equilibrado?* Available online: <https://simula.pt/artigos/orcamento-familiar.html> (accessed on 1 August 2025).
- Sornlertlamvanich, V. (2025). Enhancing large language models: Alleviating knowledge deficiency with external knowledge and semantically aware reasoning (SAR). *Knowledge and Information Systems*, 67(10), 8767–8785. [CrossRef]
- Tavares, F., Santos, E., Tavares, V., & Ratten, V. (2020). The perception and knowledge of financial risk of the Portuguese. *Sustainability*, 12(19), 8255. [CrossRef]
- Vieira, K. M., Matheis, T. K., Lehnhart, E. D. R., & Tavares, F. O. (2024). Digital financial knowledge scale (DFKS): Insights from a developing economy. *International Journal of Financial Studies*, 12(4), 120. [CrossRef]
- Wong, C. K., & Salleh, M. N. M. (2023). Personal finance and budgeting mobile application, “CashSave”. *Applied Information Technology And Computer Science*, 4(1), 1372–1387.
- Zhu, J., Xin, J., & Zhang, N. (2026). Artificial intelligence embedding and enterprise competitiveness in the embodied intelligence industry: The mediating role of competitive structure reconfiguration. *Administrative Sciences*, 16(7), 352. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.