An AI-Driven Decision Support System for Product Feature Prioritization in Software Startups

Usman Rafiq[1], Trygve Amundsen[2], Frédéric Pattyn[3], Xiaofeng Wang[1], and Christoph Stöckmann[1]

[1] Free University of Bozen-Bolzano, Bolzano, Italy
usman.rafiq@unibz.it, xiaofeng.wang@unibz.it, christoph.stoeckmann@unibz.it
[2] Independent Researcher
trygve.s.amundsen@gmail.com
[3] Ghent University, Ghent, Belgium
frederic.pattyn@ugent.be

Abstract. Product feature prioritization is widely recognized as a critical challenge for software startups. Fast-paced environments, extreme uncertainty, and limited resources often force these companies to make trade-offs between competing ideas. Our study investigates the potential of an AI-based decision support system for feature prioritization in the context of startups. We used a design science research methodology to design, develop, and evaluate a proof of concept (PoC) artifact for AI-based product feature prioritization. The artifact uses AI to make decisions regarding product features using criteria such as return on investment, confidence, and time-to-value. We demonstrated and evaluated this PoC artifact using semi-structured interviews with seven software entrepreneurs. We analyzed the evaluation data using thematic analysis. The evaluation results indicate that startups perceive the potential to use AI to support the prioritization process. However, adoption depends on improving transparency, explainability, interoperability, and usability in these tools. The study contributes to the literature by providing early insights into how AI-based tools for product feature prioritization are perceived by startups.

Keywords: Artificial Intelligence · Feature Prioritization · Software Product Management · Tech Firms · New Venture · Early-Stage Company

1 Introduction

Software companies strive to deliver high-quality and innovative software products. This responsibility primarily falls to product managers (PMs), who serve as the driving force behind product vision, development, and delivery. PMs are often referred to as 'mini-CEOs', aligning customer needs with business objectives, and working across different functions to ensure product development and success [2].

In the context of software startups, PMs face a unique set of constraints that make their responsibilities even more crucial and challenging. Unlike established companies, which benefit from stable revenue, well-defined customers, a stable product roadmap, historical data, and well-defined processes, startups often operate with limited resources, undefined requirements, and a lack of historical data [2] [4]. Similarly, these companies must quickly validate ideas, implement user feedback, and strive to build a competitive product [4]. Therefore, PMs in such organizations are required to make trade-offs, often based on intuition or incomplete information, balancing short-term survival against long-term vision. Similarly, among others, one of the significant responsibilities of this role is to decide what should be implemented next in the product. This decision-making process, often referred to as requirements or product feature prioritization, is critical and is identified as one of the core responsibilities of PMs in startups [2] [4]. At the same time, artificial intelligence (AI) technologies are becoming more integrated into product development workflows [3], offering new possibilities to automate or enhance decision-making. However, little is known about how AI can specifically support feature prioritization in startups, where intuition and ad-hoc processes often dominate. This study aims to address this gap by understanding how AI-based decision support systems can enhance existing feature prioritization practices. The following research question (RQ) guides this study:

RQ: How can AI-based decision support systems support feature prioritization in startups?

To address this RQ, we used a design science research method [7]. In general, we developed and evaluated a proof of concept (PoC) artifact that showcases an AI-based decision support system for requirements prioritization. It is based on the light framework discussed in [5].

2 Research Method

In this study, we used the design science research (DSR) method. DSR is a problem-solving paradigm that aims to create innovative artifacts to solve real-world problems while simultaneously contributing to theory [7]. A key part of this paradigm is to build and evaluate artifacts. Peffers et al. [6] define a six-step model to implement DSR: problem identification, objective setting for a solution, design and development, demonstration, evaluation, and finally communication.

Based on existing literature (e.g. [1] [5]), we noted that PMs in software startups struggle to prioritize features. Existing practices are ad hoc, manual, and subjective. Therefore, we identified a need and aimed to leverage AI to support this prioritization process.

The objective of our artifact (PoC) were to: (1) to support product managers in prioritizing features using a structured and transparent method, (2) to leverage AI to guide decision-making through a conversational interface (3) to base prioritization on clear criteria like ROI, confidence, and time-to-value, (4) to maintain and organize prioritization results for later review and comparison, (5) to communicate the different features visually on intuitive graphs for quick understanding.

The front end of this PoC was created using Next.JS , with additional technologies including Recharts for custom charts , Axios for API integration, and React-Bootstrap for styling and responsive design. The interface allowed users to view, compare, and interpret prioritization scores through an interactive graph and a data table.

For demonstration and evaluation, we used guided sessions with seven startup practitioners. Semi-structured interviews were conducted during these sessions. Finally, we used thematic analysis to analyze the evaluation data.

3 Findings

3.1 Overview of the Artifact

Based on the lightweight framework proposed in [5], we designed and developed the proof of concept (PoC) for AI-driven feature prioritization in software startups, which serves as our artifact. The prioritization framework used in this artifact evaluates each input, which could be a specific feature in a project or a standalone project, and then calculates a score and a sub-score from the input values. These values are time-to-market (TTM), time-to-business-value (TTBV), confidence, and return-on-investment (ROI). TTV and TTBV are initially added as days, then mapped to a score, following a mapping table. Similarly, confidence is also calculated using the ICE (impact, confidence, ease) framework for prioritization.

The architecture of the PoC consists of four key components: (1) a front-end interface for visualizing data, (2) a backend server for processing and business logic, (3) a relational database for persistent storage, and (4) a custom ChatGPT-based conversational assistant that enables user interaction.

Figure 1 illustrates the interactions between the different components of the PoC. Additional implementation details, including the user interfaces and user stories that guided the development of this PoC, are available at https://figshare.com/s/cf39a3df61ca022fe0cf.

3.2 Artifact Evaluation

To evaluate our artifact (PoC), we conducted semi-structured interviews with seven practitioners from software startups. In general, our evaluation focused on assessing usability, workflow integration, interpretability of AI-based outputs, and perceived usefulness for product prioritization decisions. Since the roles in startups are really fluid, therefore, our interviewees included not only formal software product managers but also founders, CEOs, and lead software developers, who were directly involved in product decision-making and vision. The

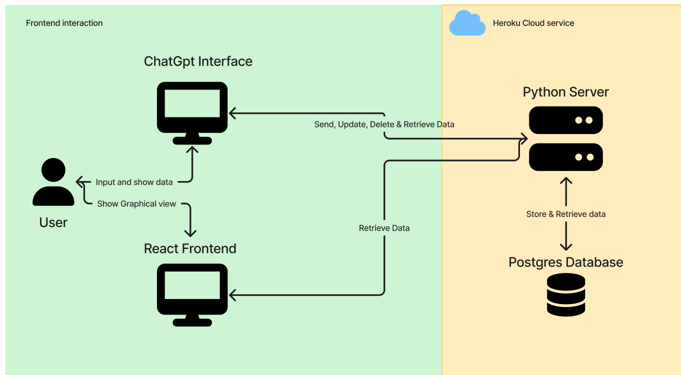

Fig. 1: Architecture of the PoC of AI-Driven Feature Prioritization Tool

characteristics of the interviewees and their startups are summarized in Table 1.

Table 1: Overview of Interviewees and Their Startups

ID	Interviewee Role	Startup Domain	Employees	Startup Stage
A	Product Manager	Digital Health	4	Growth
B	Lead Developer	Commercial Services	6	Growth
C	Founder & CEO	Healthcare	8	Growth
D	CEO & Product Manager	Food Technology	8	Growth
E	Founder & CEO	Software Solutions	7	Stabilization
F	Lead Developer	Software Development	40	Growth
G	Founder & CEO	Software Tools	1	Inception

3.3 Evaluation Results

We organized our evaluation results thematically to reflect key patterns in the data. An overview of the full thematic structure is presented in Figure 2.

Engagement with AI

Perception of AI-based Tools: We noticed mixed perspectives on AI-based prioritization tools and their place in startup workflows. Although most of the interviewees highlighted the clear potential for this tool, some others emphasized its current limitations. In particular, interviewees mentioned a lack of context and judgment. For example, interviewee C expressed his perception in the following words: "*I think it's a great tool ... But having that extra checkpoint to understand if there are any gaps in decision-making around prioritization — I think that's definitely an add-on to the tools available for us.*"

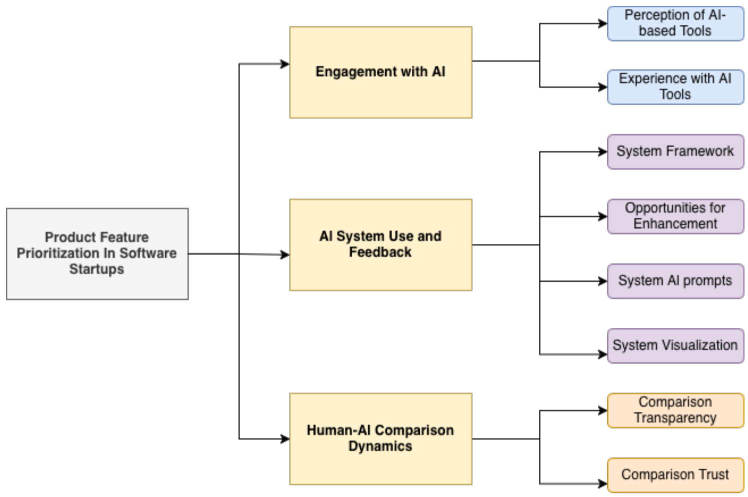

Fig. 2: PoC Artifact Evaluation Results

The general perception among interviewees was that AI could serve as a useful aid, but not by any means a replacement for human decision-making. For those startups that had already implemented some AI systems into their workflow, they described the need for a human element to verify the output. This is in line with the excerpt emphasized by interviewee C. However, a few others added that trust in AI depends greatly on its transparency and how well the AI's process and reasoning were presented.

Experience with AI Tools: None of the interviewees in our data reported using any specific AI-driven tool for feature prioritization. However, interviewees from some startups reported the use of AI tools for other parts of their workflow. These included tools like ChatGPT, Gemini, or Claude to draft project outlines, summarize meeting notices, and conduct market research. For instance, interviewee A described the use of other AI tools: "*I use ChatGPT for that for any kind of research".*

AI System Use and Feedback

System Framework: The PoC was generally well-received; however, interestingly, for many startups, the metrics in this PoC were not known. An interviewee from the startup compared it to their current method of story points, saying that this approach would be more consistent and clear. Another interviewee noted that this method of calculating a combined score of each feature would help reduce the ambiguity in prioritization. It is reported by interviewee G in the following words: "*All the metrics that you described give a very good broad view of the impact of the feature, and [what] you should consider building next.*" Likewise, interviewee

A supported this statement by adding that while different prioritization models may not vary drastically in outcome, this approach provided a clear and effective result: "*Ultimately it boils down to what factors you want to consider... I don't know that there's a whole lot of incremental value between one model versus another, but they all give a systematic way of evaluating what you need to do next. And I think your model works very effectively.*"

Opportunities for Enhancement: Interviewees generally responded positively to the overall experience of the system, but also expressed a few suggestions for improvements and a few dislikes. Many of these suggestions revolved around decreasing friction when inputting the values and clarification on the different metrics.

A recurring theme was the need for greater transparency in how the system actually worked and how it reached its result. As an example, interviewee A asked for: "*clarity on understanding what are those metrics like ... how those metrics are made, like what [and] how it's calculated ... can build some initial trust.*" Other interviewees proposed more technical implementations, such as having it integrated directly into GitLab or GitHub, making the workflow more seamless.

Another pointer we noticed in the data is to make the visualization reflect neglected tasks. Interviewees suggested that this could be done by increasing the size of the circle, the longer it had been in the queue. This could help teams identify forgotten tasks and adjust priorities accordingly.

System AI Prompts: Since our PoC was based on the conversational bot, in the results, we looked at how the interviewees perceived different sets of questions received from it. In general, startups' interviewees considered the prompts clear, helpful, and practical in describing their features. They also highlighted that answering these questions started a deliberate reflection process on the different values. It is reported by the interviewee G: "*It ...forces you to think about every aspect of the feature because, like, if I see the Trello tasks, I think it is OK, it takes a long time... Not that valuable... I will throw it away, but by forcing you to fill out the whole metrics, you reevaluate... Which number should I put there?*"

Likewise, beyond the specific questions on each metric, interviewees also appreciated the possibility of asking follow-up questions to the bot. This was found to be relevant when, for example, the meaning of ROI was unclear or if some examples were needed. This was especially seen as necessary if there were personnel working the tool not fully experienced with product management jargon. Similarly, some interviewees emphasized that certain prompts (RoI, Time-to-value) were difficult to answer and suggested the use of a more category-based scoring system (low-medium-high) to reduce the need for a guesstimate.

Lastly, although the questions were generally good and well-structured, some interviewees raised questions about the efficiency of the tool in terms of asking questions each time.

System Visualization: The interviewees perceived the visualization of the final results positively. For example, interviewee G explained how the visual layout supported faster and more objective decisions for their startup: "*The visualization is very helpful because now you can just focus on a very particular subset of the graph. The big dots at the top".*

The ability to compare features side by side was also appreciated in the data. Some interviewees even remarked that this could be adopted for internal communication or for yearly strategic roadmap planning. For example, in the case of interviewee E, the interviewee reflected: " *... I would expect people to want that, because you go back and forth a lot of times on the road map iteration".*

Apart from appreciations, some interviewees felt the need to have greater interaction, to change different metrics for each feature, and to see the change in position and size live. Others mentioned a need for more information on the graph, such as a tool tip or a written description of the calculation, for greater clarity. Although the visual model was valued for its clarity and simplicity, startups wanted to be able to tailor it to their workflow and systems.

Human–AI Comparison Dynamics

Comparison Trust: During the PoC demonstration, the interviewees of the startups also reflected on the trustworthiness of the system. We identified a high level of initial trust in the AI-driven system. Particularly since the AI was used to formalize a structured logic already set in the system, interviewee G explained: "*I would trust that ... just explaining the feature [with words] and then [letting the AI] come up with the metrics that you usually fill in.*"

Startups were more willing to trust it initially, in contrast to if the AI had to figure out and calculate each metric itself.

However, other interviewees voiced a bit of hesitation about the system, especially around contextual limitations. Although an interviewee thought it would make more mistakes than an average person, it would still need human oversight. It is evident from the following excerpt of the interviewee A: "*Having a certain level of experience is important, and why humans need to be part of the equation, because the model may generate something that's not accurate. But you need to be able to eyeball it...*"

Comparison Transparency: We found that while trust in the system was important, the ability to understand and explain the choices and results of the system remained equally important. Our data analysis revealed that being able to justify and explain why one feature is better than another is very important in decision-making processes. The claim was backed by the argument that in real-world product environments, decisions must be communicated, justified, and defended to stakeholders such as engineers, executives, or clients.

For instance, consider the quote of interviewee A: "*Justifying the model suggestions ... becomes kind of the lifeblood of the product managers, how to navigate the politics of moving something forward.*"

4 Discussion and Conclusions

Our findings suggest that AI-driven systems that support feature prioritization can play an important role in supporting decision-making. Startups are already leaning towards the use of AI for several reasons; however, none of them currently focus on feature prioritization. The demonstration of the PoC surfaced that the benefits of such tools are perceived by startups, but they think these are contingent. Overall, startups still do not consider this as a substitute for human decision-making for feature prioritization, and pointed out a few suggestions for improvements in this PoC. Four key improvements are discussed in our findings: (1) While startups were open to integrating AI into their workflow, they suggested connecting our tool directly to third-party applications, such as TeamHub, GitLab, or Slack, to improve integration, (2) Startups were less likely to trust the outcome if they did not know how the AI prioritized features. These results reinforce the importance of explainability in AI systems, (3) Startups appreciated the step-by-step process of talking to a bot in our PoC, as the bot helped them reflect more on the numbers and the specific feature than they otherwise would have. However, some startups with more established processes found that the conversation slowed them down and asked for another way to get input for feature prioritization, (4) Regarding the feature selection criteria, the scoring prioritization method and visual representation were described as useful and effective. This was especially the case for early-stage startups. The visualization of the graph in our PoC stood as a key strength of the system.

Based on these key takeaways, future research should investigate how improvements in integration, explainability, usability, and visualization can enhance the adoption and effectiveness of these AI-assisted feature prioritization tools in startups.

References

1. Melegati, J., Goldman, A., Kon, F., Wang, X.: A model of requirements engineering in software startups. Information and software technology **109**, 92–107 (2019)
2. Melegati, J., Wiese, I., Guerra, E., Chanin, R., Aldaeej, A., Mikkonen, T., Prikladnicki, R., Wang, X.: Product managers in software startups: A grounded theory. Information and Software Technology **174**, 107516 (2024)
3. Ogundipe, D.O., Babatunde, S.O., Abaku, E.A.: Ai and product management: A theoretical overview from idea to market. International Journal of Management & Entrepreneurship Research **6**(3), 950–969 (2024)
4. Pattyn, F.: The critical role of product managers and their responsibilities in software startups: A systematic literature review. American Journal of Engineering and Technology Management **9**(4) (2024)
5. Pattyn, F., Wang, X., Poels, G.: A requirements prioritization method: Vectr
6. Peffers, K., Tuunanen, T., A. Rothenberger, M., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems **24**, 45–77 (2007)
7. Vom Brocke, J., Hevner, A., Maedche, A.: Introduction to design science research. In: Design science research. Cases, pp. 1–13. Springer (2020)