Predictions through Lean startup? Harnessing AI-based predictions under uncertainty

Santo Raneri and Fabian Lecron
Department of Engineering Innovation Management, Faculty of Engineering, University of Mons (UMONS), Mons, Belgium, and
Julie Hermans and François Fouss
Louvain Research Institute in Management and Organizations (LouRIM), Université catholique de Louvain (UCLouvain), Mons, Belgium

Abstract

Purpose – Artificial intelligence (AI) has started to receive attention in the field of digital entrepreneurship. However, few studies propose AI-based models aimed at assisting entrepreneurs in their day-to-day operations. In addition, extant models from the product design literature, while technically promising, fail to propose methods suitable for opportunity development with high level of uncertainty. This study develops and tests a predictive model that provides entrepreneurs with a digital infrastructure for automated testing. Such an approach aims at harnessing AI-based predictive technologies while keeping the ability to respond to the unexpected.

Design/methodology/approach – Based on effectuation theory, this study identifies an AI-based, predictive phase in the "build-measure-learn" loop of Lean startup. The predictive component, based on recommendation algorithm techniques, is integrated into a framework that considers both prediction (causal) and controlled (effectual) logics of action. The performance of the so-called active learning build-measure-predict-learn algorithm is evaluated on a data set collected from a case study.

Findings – The results show that the algorithm can predict the desirability level of newly implemented product design decisions (PDDs) in the context of a digital product. The main advantages, in addition to the prediction performance, are the ability to detect cases where predictions are likely to be less precise and an easy-to-assess indicator for product design desirability. The model is found to deal with uncertainty in a threefold way: epistemological expansion through accelerated data gathering, ontological reduction of uncertainty by revealing prior "unknown unknowns" and methodological scaffolding, as the framework accommodates both predictive (causal) and controlled (effectual) practices.

Originality/value – Research about using AI in entrepreneurship is still in a nascent stage. This paper can serve as a starting point for new research on predictive techniques and AI-based infrastructures aiming to support digital entrepreneurs in their day-to-day operations. This work can also encourage theoretical developments, building on effectuation and causation, to better understand Lean startup practices, especially when supported by digital infrastructures accelerating the entrepreneurial process.

Keywords Digital entrepreneurship, Lean startup, Artificial intelligence, Effectuation, Causation, Uncertainty

Paper type Research paper

Introduction

The rapid growth of digital technologies has completely transformed the way entrepreneurs build their products, conceive their business models, test their ideas and deal with uncertainty (Ghezzi and Cavallo, 2020; Nambisan, 2017). While still relatively sparse (Kraus *et al.*, 2018), the study of digital entrepreneurship has gained momentum. The first developments were mainly theoretical, with papers laying the framework for future research (Nambisan, 2017; Sahut *et al.*, 2019; Sussan and Acs, 2017). For example, Nambisan (2017) proposed a research agenda calling for an explicit theorization of concepts related to digital technologies in order to enrich theories of entrepreneurship. Sussan and Acs (2017) and Sahut *et al.* (2019) followed this trend by proposing a conceptual framework for the study of the digital entrepreneurship ecosystem and by identifying major research streams respectively. Special issues allowed the emergence of empirical papers (Kaminski and Hopp, 2019; Prüfer and Prüfer, 2020; von Bloh *et al.*, 2020). As exposed in the next paragraph, these papers started to recover the lost ground of research in entrepreneurship on the question of big data and artificial intelligence (AI) (Obschonka and Audretsch, 2020). Indeed, compared to other fields, including management research (Sheng *et al.*, 2017), entrepreneurship researchers have devoted little attention to this emerging trend.

Machine learning techniques – a branch of AI where a model is trained on known data to predict labels of unknown data – increase human cognition in an unprecedented way, allowing the prediction of so far unpredictable events or outcomes. New questions can now be tackled by applying AI-based techniques (Prüfer and Prüfer, 2020). For example, in Kaminski and Hopp (2019), the authors introduced an approach based on neural networks and natural language processing to predict the returns of a crowdfunding campaign. Prüfer and Prüfer (2020) offered an original analysis of the dynamics of demand for entrepreneurial skills in the Netherlands. In von Bloh *et al.* (2020), big data techniques explore the possible link between news coverage of entrepreneurship and regional entrepreneurial activity. Therefore, the contribution of big data and AI techniques allows the field of predictability to be broadened through the development of models based on the history of known data (Amoroso *et al.*, 2017). Situations previously perceived as uncertain – where the probability distributions of the possible cases are unknown (Liu, 2012) – can today be seen as risky – that is, where the possible cases are known as well as the probability of occurrence of each of these cases (Knight, 1921). This is especially important for entrepreneurial decision-making as entrepreneurs often oscillate between situations of risk and uncertainty (Galkina *et al.*, 2021; Sarasvathy, 2001).

As such, decision-making in situations of risk and uncertainty is a key topic for researchers on digital entrepreneurship. Especially, Nambisan (2017) calls for more research about digital infrastructures, that is, the digital resources and tools, including prediction algorithms and digital prototype builder, that help entrepreneurs develop their ventures. It is in line with Obschonka and Audretsch (2020) who called for more research on the role of predictive algorithm for entrepreneurship, taking risk and uncertainty in consideration. It is about exploring the way digital infrastructures change the nature of uncertainty as well as the way entrepreneurs deal with said uncertainty.

However, such predictive technologies strongly fit with a causal "predictive" view of the entrepreneurial process, wherein markets can be measured and discovered (Mansoori and Lackeus, 2020). For entrepreneurs engaging with predictive infrastructures, there is a risk of being "stuck" in a quest for an optimal solution without being able to react to unexpected contingencies. First, entrepreneurs deal with a large variety of domains. In some cases, a core technological resource might be available and provide predicted data about desirability, while other components of the business model (such as the choice of a logistic partner) might require substantial development and adaptability (Sanz-Velasco, 2006). Second, there is a challenge to combine cognitive logics based on a prediction (causal) versus controlled

(effectual) frameworks (Sarasvathy, 2001, 2021). Accordingly, an entrepreneur relying too blindly on a prediction algorithm might be unable to leverage unexpected changes, surprising results and other unpredictable events like a pandemic. Authors such as Galkina and Jack (2021) or Smolka *et al.* (2018) recommend the simultaneous and synergistic use of causal and effectual logics when developing opportunities and creating ventures. However, they concede that synergetic effects depend on specific circumstances and that little is known about how entrepreneurs operationalize blended logics as hybrid practices.

Thus, there is a risk to consider AI techniques as the panacea that mitigates all uncertainty. Prediction algorithms can lead to a paradoxical situation where some elements of uncertainty can be reduced by being converted into a risk, but at the same time, relying on prediction might reduce the capacity of entrepreneurs to deal with unexpected contingencies. The aim of this paper is to develop a framework that accommodates both prediction and effectual practices. The Lean startup approaches (LSAs) are proposed as a frame for hybrid practices. This work elaborates on the role of prediction in LSAs and, thereby, helps practitioners harnessing the power of predictive AI without forfeiting their capacity to deal with unexpected contingencies.

The research question addressed in this article is thus:

RQ. How can entrepreneurs harness prediction technologies while keeping a control-based, effectual approach for product development?

The paper elaborates on the articulation of causal and effectual logics as part of an extended build-measure-predict-learn (BMPL) loop, where the build phase is effectual, the measure and predict phases are by essence causal and the Learning phase is both causal and effectual.

To the best of our knowledge, the present paper is the first to theorize how effectuation and causation can be combined through LSAs, exploiting the potential of both risk-taking through predictive technologies and addressing uncertainty. By developing an AI-based infrastructure as part of the LSA toolkit, this research aims at a twofold contribution. First, it contributes to the conversation about hybrid practices and the key mechanisms that help entrepreneurs blend causal and effectual logics (Galkina *et al.*, 2021; Galkina and Jack, 2021). Second, it contributes to the emerging conversation about digital entrepreneurship by discussing the implication of machine learning on decision-making in situation of risk and uncertainty (Nambisan, 2017). Especially, the build-measure-learn (BML) loop of the Lean startup methodology is revised to add a predictive stage. The new BMPL loop is designed to explicitly articulate causal and effectual logics.

The theoretical background is exposed in the next section, starting with entrepreneurial logics of action, how they relate to LSAs and how they can be augmented by relevant AI-techniques. Building on this background, the next section describes the AI-based BMPL loop: the underlying articulation of logics and its AI-based techniques. Next, the model is applied to a concrete case study, the "EasyTips" entrepreneurial project, thereby providing a primer about how entrepreneurs can build on AI techniques to facilitate opportunity development through LSAs. The results are presented and assessed. Finally, we conclude by discussing results in the light of entrepreneurship research. The main advantages of the methodology – that is, the acceleration of the pace of experiments and facilitation of the evaluation of experiments' results – are discussed along with their implications on uncertainty management. Conclusions draw the limits of the current research and open up paths for future studies.

Theoretical background
### Entrepreneurial logics of action
According to effectuation theory, an effectual logic of action favors a logic of control rather than planning and prediction (Sarasvathy, 2001). Meanwhile, causation logic emphasizes

predicting and planning the future in a strict goal-oriented logic (Smolka *et al.*, 2018). Effectuation and causation differ mainly in five dimensions: views of the future, bases for taking action, predisposition toward risk and resources, attitudes toward outsiders and attitudes to protect themselves from unexpected contingencies (Sarasvathy, 2003; Yang *et al.*, 2019). First, while causation logic sees the future as a continuity of the past that can be planned, effectual entrepreneurs believe that the future can be shaped and is not written. Second, causal entrepreneurs are goal-oriented, and actions are taken to achieve these goals. Effectuation, meanwhile, is characterized by a mean-oriented logic, and goals are determined by these given means. Third, causation focuses on maximizing expected returns, while effectuation focuses on limiting losses to an affordable level. Fourth, the attitude toward outsiders is mainly competitive in causal logic, whereas effectual logic advocates the creation of partnerships. Fifth, causation tries to avoid contingencies as much as possible by using prediction and planning tools, while effectuation views contingencies as inevitable and encourages them to be seen as opportunities.

While effectual and causal logics are often presented as two polarized logics of action, they have been conceptualized as orthogonal dimensions that can be combined. For example, Smolka *et al.* (2018) and Laskovaia *et al.* (2017) show that using a combination of logics have a positive effect on new venture performance. Likewise, Berends *et al.* (2014) find that small firms adopt a combination of causation and effectuation logics during product innovation processes, and Evald and Senderovitz (2013) reveal that SMEs engaging in internal corporate venturing activities combine causation and effectuation logics. This is in line with Andries *et al.* (2013), who argue that, in uncertain environments, new ventures concurrently focusing on causal and effectual logics might be more innovative, and with Vanderstraeten *et al.* (2020), who provided some evidence for this synergistic effect in small businesses. This stream of research suggests that there are synergistic effects to combine logics and even presents such combinations as a survival imperative in dynamic environments (Laine and Galkina, 2017). However, few elaborate on the mechanisms to do so. In spite of their relevance for opportunity development (Galkina *et al.*, 2021), it remains unclear how entrepreneurs can combine causation and effectuation and how hybrid practices can provide entrepreneurs with learning advantages in terms of lower information costs or faster decision-making (Grégoire and Cherchem, 2020).

One candidate for the operationalization of effectual and causal logics is the LSA. Broadly speaking, LSA applies a Lean management mindset to the problem of finding the right product market fit through multiple iterations (or loops). It advocates for validated learning to grasp what customers want in order to meet their needs. According to Ries (2011), this is done using the scientific method, which consists of designing, running and analyzing experiments in order to prove or reject the startup team's hypotheses. This process is guided by the BML loop (see Figure 1).

First, hypotheses about the product and growth are implemented in a minimum viable product (MVP) during the first build phase. An MVP is "*that version of the product that enables a full turn of the BML loop with a minimum amount of effort and the least amount of development time*" (Ries, 2011, p. 77). Then the MVP is tested by users or potential customers during the measure phase. The information collected and quantified (people's impressions, textual feedback, churn rate, bounce rate, click through rate, conversion rate, etc.) is then analyzed in the learn phase to validate or invalidate hypotheses, and the new knowledge gathered serves as the basis for formulating new hypotheses to test. A new iteration thus starts with the implementation of the new hypotheses. LSAs and their experimentation-oriented mindset are nowadays largely adopted by the practitioners as evidenced by the success of Ries' book and its widespread use in entrepreneurship curricula (Sarasvathy, 2021). It is becoming the standard among entrepreneurs, in aid organizations and in university courses on entrepreneurship (Blank, 2013; Ghezzi, 2019; Mansoori and Lackeus, 2020; Shepherd and Gruber, 2020).
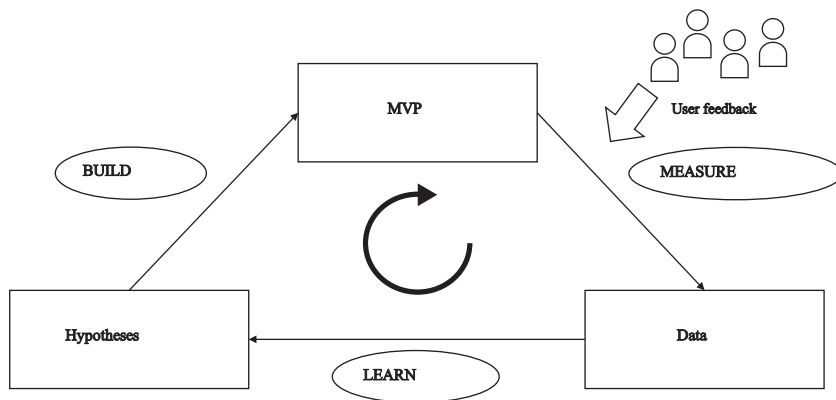
Figure 1.
Overview of the build-
measure-learn loop

As a result, authors such as Sarasvathy (2021) suggest that it would be useful to understand how LSAs enable the operationalization of both causal and effectual logics of action.

According to Frederiksen and Brem (2017, p. 181), *"it is tempting to regard [the Lean startup] as the practical implementation of Sarasvathy's research."* Ghezzi (2019) suggested that combining extremely scarce resources into an MVP is tightly connected to effectuation and bricolage (Baker and Nelson, 2005). He also highlighted that formulating hypotheses about how the world could change and how startups can play a role in it is very similar to the effectuation's view on the future. According to him, even experimenting and testing these hypotheses to gather relevant feedback is consistent with effectuation's principle stating that contingencies should be handled and exploited by entrepreneurs.

Likewise, Yang *et al.* (2019) examined the cognitions behind the Lean startup. They argued that effectuation is mainly associated with search activities, while causation is associated with more execution activities. Search activities are the controlled and proactive processes of attending to, examining and evaluating new knowledge and information (Li *et al.*, 2013). Execution is the process of establishing and accomplishing well-defined plans (Yang *et al.*, 2019). In practice, Yang *et al.* (2019) divided the journey of an entrepreneurial firm into two stages: the early development stage, in which mainly search activities are conducted, and the late development stage, in which mainly execution activities are conducted. This view is consistent with the view of Ghezzi (2019), who defines the turning point for moving from the search stage to the execution stage as the moment when the product-market fit is achieved. This is in line with Blank's view in his book *Customer Development* (Blank, 2005), that a startup should at first focus on the task of finding a sustainable business model (search phase) and then exploit it (execution phase). Therefore, the works of Yang *et al.* (2019) and Ghezzi (2019) argue that entrepreneurs following an LSA start with an effectual cognitive logic and then switch to a causal logic once the product-market fit is reached.

However, LSAs might not be devoid of any causal logic. For instance, Frederiksen and Brem (2017) suggested that the goal of building a successful business might be causal in essence. More fundamentally, it could be argued that the BML loop is rooted in a predictive ontology, whereby the market is knowable and an optimal solution can be found. At the core of LSAs, it is about predicting the market rather than controlling for uncertainty (Sarasvathy, 2021). As suggested by Mansoori and Lackeus (2020), and despite the ambition of Ries to help entrepreneurs create new products and services under conditions of extreme uncertainty, dealing with said uncertainty is a key weakness of LSAs (Mansoori and Lackeus, 2020).

Paradoxically, LSAs are often used to reduce perceived uncertainty rather than to leverage truly unknowable events. Entrepreneurs are stuck in a quest for an optimal solution instead of sealing an affordable one.

This paradox is not inevitable. For instance, Sarasvathy (2021) recently suggested that entrepreneurs following LSAs might start in a causal, prediction-based framework in order to develop their product-fit and only shift to a control-based logic when elaborating their business model. It would mean predicting an optimal product (i.e. causal approach) and then finding an "acceptable" business model (i.e. effectual approach). It follows that LSAs might enable the combination of effectuation and causation sequentially and concurrently. In the next section, we show *how* an AI-based infrastructure can facilitate this process.

Digital infrastructures for entrepreneurs

Digital entrepreneurship entails a large variety of phenomenon. For Nambisan (2017), it focuses on the way new digital technologies, including data analytics and machine learning, "*has transformed the nature of uncertainty inherent in entrepreneurial processes and outcomes as well as the ways of dealing with such uncertainty.*" Nambisan (2017) highlights three distinct but related sources of digitalization: digital artifacts, digital platforms and digital infrastructure. Basically, digital artifacts are the digital components that are part of a new product (or service). Digital platform is about the shared set of services and architecture that host complementary offerings. Digital infrastructures are the digital resources and tools, such as communication tools, social media, cloud computing, automation, AI techniques or digital MVP builder, that help entrepreneurs when developing their venture. The present research is about the latter category. By incorporating AI-based techniques into the LSAs, its contribution is about AI-based (digital) infrastructures and how it can help entrepreneurs to deal with uncertainty when developing their opportunity. To better position this contribution in the broader product design literature, a summary is presented hereafter.

In the last decade, AI-based and big data-based techniques have received an increasing amount of attention in product design. As product design is an important part of the Lean startup methodology, it is interesting to have an overview of extant research. Two main sets of methods are found. The first set deals with searching for the best features to integrate into a specific product, whereas the second set gathers approaches aiming to maximize the affective design of a product. For both sets, certain techniques propose to implement natural language processing methods (text analysis), while others use strictly numerical methods.

When it comes to predicting the best features in product design, the most common approach is to implement a conjoint analysis method (Fan *et al.*, 2017; Luce and Tukey, 1964). The key challenge of conjoint analysis is to find the best product variant in a multi-attribute product space. This is commonly done by asking for customers' ratings about variants of the product in order to build a utility function for each customer. Three other types of conjoint analyses have arisen from studies trying to improve the traditional method: choice-based conjoint analysis (Louviere and Woodworth, 1983), adaptive conjoint analysis (Johnson, 1987) and self-explicated conjoint analysis (Green and Srinivasan, 1978; Rao and others, 2014).

Besides conjoint analysis approaches, other types of methods usually seen in the marketing literature for incorporating customers' preferences into product design and development are quality function deployment and heterogeneous design. Quality function deployment introduces a planning matrix to relate customer preference to design, manufacturing and marketing teams in a firm (Hauser and Clausing, 1988). Heterogeneous design adapts the homogeneous design method by presenting different designs to different customers. Sándor and Wedel (2005) were the first to propose this approach and demonstrated the value of taking prior information about heterogeneity across consumer preferences into account.

More recently, Jiao *et al.* (2019) provided an integration model that addresses the problems of generating feasible configuration plans and selecting the plan that best satisfies the customer requirements. The model uses transaction data for generating configuration plans after analyzing the segmented market characteristics. The customer requirements are then mapped to the configuration plans using a Naïve Bayes classifier. Likewise, Tao *et al.* (2019) presented a method for product design using a digital twin approach. The digital twin – a virtual representation of a physical product – is used to better understand how customers interact with the current version of the product and for virtually testing its improvements under different conditions.

In the context of improving an existing product, some studies intend to predict the best features to include in the next version by analyzing what customers say about the features of the current product on the Internet. For instance, extraction of customer needs was carried out in Ireland and Liu (2018). To start, their framework divided the reviews extracted from Amazon.com into separate sentences, a Naïve Bayes classifier was then applied to determine sentiment polarities and finally feature-sentiment pairs – pairs of words: one sentiment word and one word describing a product feature – were generated and their significance was evaluated. The authors demonstrated through the case study the utility of what they called "data-driven product design." Likewise, Chen *et al.* (2019) implemented several artificial intelligence techniques including feature extraction, sentiment analysis, anomaly and novelty detection, and time-series analysis for analyzing customers reviews and classifying product features.

Affective design is also a method aiming to maximize customer satisfaction in purchasing new products. This technique, based on Kansei engineering (Nagamachi, 1995), is commonly implemented through machine learning methods that deal with survey data and/or big data. A product with good affective design is a product that excites consumers' feelings of wanting to buy it. For instance, Wang *et al.* (2019) exploited online customer reviews from Amazon. com to extract affective opinions. This was done through a heuristic method combining text mining rules and deep learning models. It is thus different from the papers introduced above, since there is no product features extraction phase. The study focused on what people said about the product as a whole, and not on what they said about each feature of the product. The aim here is to understand the customers' comments from an affective perspective. Instead of using existing comments, Ling *et al.* (2014) evaluated different affective responses to new products through an online survey. Finally, Chien *et al.* (2016) provided a data mining framework to capture user experiences of product visual aesthetics. User background information, perception data and user experience reaction were used to train a rule-based model. The role of the resulting model was to predict the user experience reaction from information about the user and her or his perception of some characteristics of the product.

The papers introduced above proposed techniques based on AI, sometimes supported by big data-based methods, to assist the design of new products. While those approaches are promising, they are not always relevant for new opportunity development. Indeed, a significant proportion of these starts with the assumption that a current version of the product is already on the market and is largely distributed – largely enough to be able to gather a large number of textual comments – which is not the case for a new venture. Others assume that similar products are already commercialized, which is not always possible for highly original products. Furthermore, extant research did not fully address several questions that managers or entrepreneurs may ask, such as the following: How can the method be integrated into an entrepreneurial process ? At which steps should we use the algorithm? What should we do with the output of the algorithm? Does the algorithm actually reduce uncertainty? Are the algorithm results accurate? In other words, they fail to propose a methodology so that entrepreneurs can articulate the AI techniques into their management and entrepreneurial processes.

The LSAs try to reach the shared goal of the articles introduced above: guiding the product design to create a product that best satisfies customer expectations. Assisting the LSAs with AI-based techniques might achieve this goal even more effectively, as long as it enables an articulation of causal and effectual logics, taking into consideration that information about new opportunity development is scarce and uncertain. It is about predicting what users who have not given feedback about a certain feature are likely to think about it. In the next section, an extended LSAs "Build-Measure-Predict-Learn" loop is presented, along with its AI-based components.

An AI-based build-measure-predict-learn loop for blended logics

According to current research on effectuation, a constellation of behaviors is possible, from sequential use of causal and effectual phases (Yang *et al.*, 2019) to concurrent use of logics (Smolka *et al.*, 2018), with synergetic effects (Galkina and Jack, 2021). Furthermore, LSAs have the potential to articulate both logics of action as a set of hybrid practices. On the one hand, LSAs and effectuation share the basic tenant that a high degree of uncertainty can only be effectively and actively reduced through an experimental process that converts assumptions into facts (Fiet, 2002; Mansoori and Lackeus, 2020; Pfeffer and Sutton, 2006). As such, LSAs have been presented as a set of methods to operationalize effectuation (Frederiksen and Brem, 2017; Ghezzi, 2019). On the other hand, LSAs hold an ontological perspective (a view of the world) that is coherent with a predictive framework, whereby the market can be known and an optimal solution can be found. Recently, Sarasvathy (2021) even suggested that LSA might provide tools that fit in both prediction-based and controlled-based approaches. Hence, LSA is a good candidate to accommodate hybrid practices.

The main goal of this section is to present how entrepreneurs can extend the LSAs' BML loop to harness predictive technologies while keeping a control-based, effectual approach for opportunity development. When implementing the extended BMPL loop, each phase can be qualified as fitting with effectual, causal or both logics, and thus offers a frame to better articulate them. The BMPL loop is illustrated in Figure 2, specifying the underlying logic of each phase. The remainder of this section is dedicated to a more detailed description of the model.

Build: implementing hypotheses

The BMPL loop starts with the formulation of hypotheses about the new product and expected growth, as is the case with traditional LSAs. After identifying the hypotheses that should be tested in the new iteration, they are translated into concrete product design decisions (PDD) to be implemented into the MVP. Imagine a startup offering a mobile application which enables users to share their drawings with friends. If the team uncovers that its users find the app's profile screen too insipid, one hypothesis it can formulate is that users should appreciate putting some funny filters on their profile picture. The PDD can be: "implement a filter adder feature in the profile settings screen with five different filters: a clown, a dog, a cat, a monster and a frog." Then, a preliminary version of the functionality is implemented (the MVP) through a frugal approach (Mansoori and Lackeus, 2020). In line with Ghezzi (2019), the build phase mostly fits with a control-based, effectual logic of action since building an MVP might require assembling available resources in a constrained environment.

Measure: collecting feedback

The role of the measure phase is to assess the extent to which users are satisfied with a new PDD implemented in the product they use. To this end, users rate the new PDD, and the AI
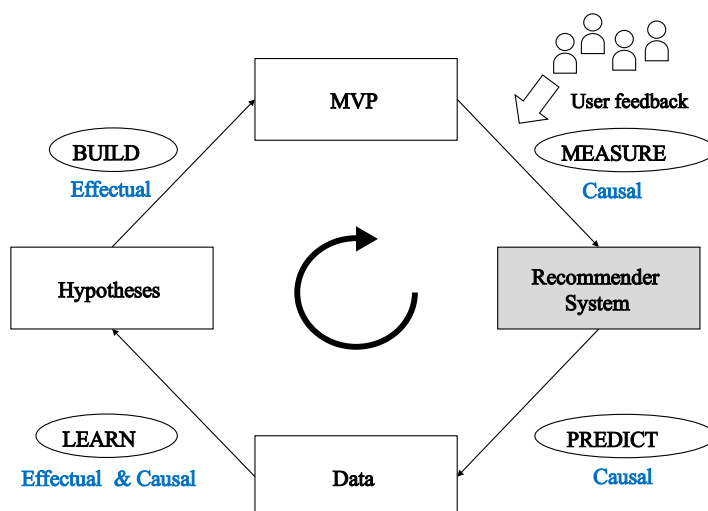
system collects the information. This can be achieved in several ways, such as embedding a 5-star rating system in the digital product to ask users who have tried a feature to rate it; another option would be to request the ratings through an online survey. Although this paper emphasizes the implementation of the BMPL loop in a digital environment, it can be implemented by a company offering physical products, since ratings for each new PDD can be collected via an online survey or interviews after a few interactions with the MVP. The main advantage of a digital environment is the possibility to more quickly implement and test a new PDD in the product.

Ghezzi (2019) convincingly argued that measuring can be part of an effectual view of the world, wherein startup can imagine and create the future. However, as suggested by Mansoori and Lackeus (2020), entrepreneurs adopting an LSA "discover the future" by measuring it. This contrasts with an ontological view of uncertainty in which the future is unknowable in principle (Knight, 1921; see Mansoori and Lackeus, 2020), thus requiring an effectual approach. Instead, LSAs consider that uncertainty can be reduced through adequate information gathering strategies. In line with these argument, the present model suggests that the measure phase better fits with a causal logic of action, since it implies that the market can be known and quantified.

Predict: seeing beyond the collected data
Furthermore, the predict phase might be considered as causal in essence as it relies on the prediction of a knowable market. In the traditional BML loop, the function of the measure phase is to collect feedback that will be assessed during the learn phase. In the BMPL loop, its role is to gather enough feedback from users on new PDDs in order to run a collaborative filtering recommender system (CFRS). CFRS intends to predict what all users would think about the new PDD based on the ratings they gave for other PDDs in the past. A startup team using such a methodology should not wait until a representative part of the users has expressed their opinion on the new PDD, but only until enough feedback is collected to run the CFRS. In order for the CFRS to perform well, the feedback gathered consists of the ratings given by the users on a PDD. Let us denote as $\mathscr{I}$ the set of PDDs for which ratings have been collected, as $\mathcal{U}$ the set of users who have already given feedback and as r_{ui} the rating of user

u for PDD i. We distinguish predicted ratings from the known ones by denoting as $\widehat{r}_{ui}$ the prediction of r_{ui}. The prediction problem is thus to compute [1] $\widehat{r}_{ui} \forall u \in \mathcal{U}$ for whom r_{ui} is not known for a new PDD i.

The objective of this phase is to predict the ratings of users that did not rate the new PDD during the previous phase. This corresponds to the AI-based causal part of the BMPL loop.

The functioning of measure and predict phases is illustrated in Figure 3.

Computing the similarity weights

Computing the similarity weights σ_{uv} between users u and v can be done using (1) through the Pearson correlation (PC) method, which compares ratings by removing effects of mean and variance (Ning *et al.*, 2015). The similarity between user u and user v, denoted by w_{uv}, computed by the PC method is:

$$w_{uv} = \text{PC}(u,v) = \frac{\sum_{i \in \mathscr{I}_{uv}} \left(r_{ui} - \overline{r}_u\right)\left(r_{vi} - \overline{r}_v\right)}{\sqrt{\sum_{i \in \mathscr{I}_{uv}} \left(r_{ui} - \overline{r}_u\right)^2 \sum_{i \in I_{uv}} \left(r_{vi} - \overline{r}_v\right)^2}}, \tag{1}$$

where $\mathscr{I}_{uv}$ is the set of items rated by both u and v, and $\overline{r}_u$ $(\overline{r}_v)$ is the mean rating given by u (by v). Note that $w_{uv} \in [-1, 1]$ when using PC.

However, a problem may arise when $\sum_{i \in \mathscr{I}_{uv}} (r_{ui} - \overline{r}_u)^2 = 0$ or $\sum_{i \in \mathscr{I}_{uv}} (r_{vi} - \overline{r}_v)^2 = 0$. In this case, the PC between u and v cannot be computed. Therefore, when one of these sums is zero, the preference similarity between u and v is computed through the cosine vector (CV) formula. This formula is identical to (1) except that the average rating is not subtracted from the ratings r_{ui}. Nonetheless, as $\text{CV}(u,v) \in [0, 1]$, the following linear transformation is applied:

$$w_{uv} = 2CV(u,v) - 1. \tag{2}$$

Therefore, when a new PDD i is shown to users, the ratings collected in the measure phase are added to the previously gathered ratings to form the set $\mathcal{R}$. Note that when a PDD is the improvement of an existing feature, the system does not remove the PDD corresponding to the previous improvements nor the PDD corresponding to the first implementation of it since these contain important information about users' opinions. After adding new ratings to $\mathcal{R}$, a matrix of similarity weights of size $(n \times n)$ is fulfilled with the w_{uv} computed from $\mathcal{R}$ with (1).

Selecting the neighbors

The next step is to predict what users who have not rated the new PDD i are likely to think about it. As shown in Figure 3, for each missing rating r_{ui}, the prediction computation starts with the selection of the set $\mathcal{N}_i(u)$ of nearest neighbors. The set $\mathcal{N}_i(u)$ contains between k_{min} and k_{max} users who rated i, with the highest values of w_{uv}, where w_{uv} must be greater than a threshold value w_l to be considered.

Once the nearest neighbors of u who rated i are selected, the prediction of her/his rating for i is computed by:

$$\widehat{r}_{ui} = \overline{r}_u + \sigma_u \frac{\sum_{v \in \mathcal{N}_i(u)} w_{uv}\left(r_{vi} - \overline{r}_v\right)\big/\sigma_v}{\sum_{v \in \mathcal{N}_i(u)} |w_{uv}|}, \tag{3}$$

where σ_u and σ_v denote the standard deviation of the ratings given by u and v. Equation (3) uses the Z-score normalization. The objective of Z-score is to convert individual ratings into a more universal scale by taking into account the average and the spread of the ratings that
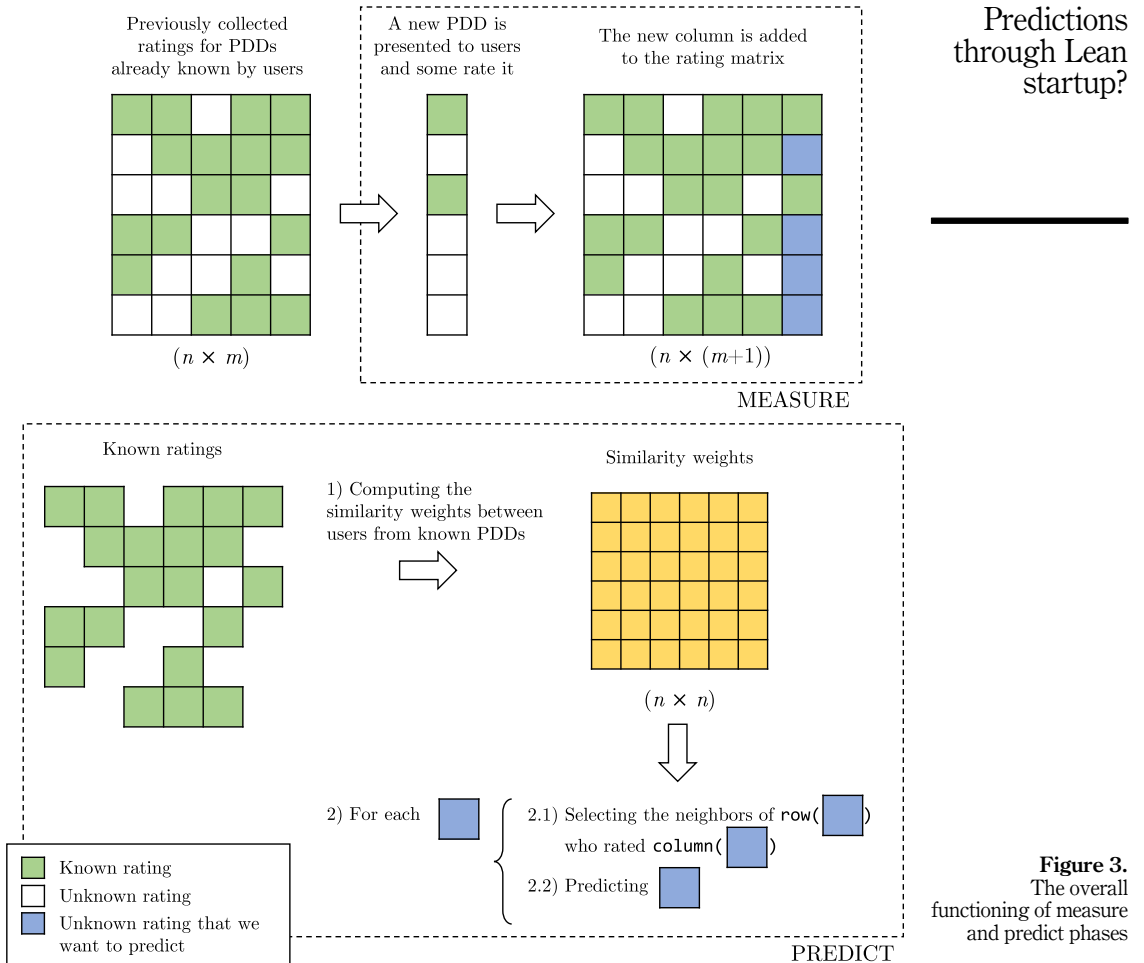
Figure 3.
The overall
functioning of measure
and predict phases

users gave (Ning *et al.*, 2015). The rating of neighbor *v* influences more the computation of $\widehat{r}_{ui}$ if the r_{vi} differs from $\overline{r}_v$ and the standard deviation is small. In other words, if a neighbor of *u* rates *i* very differently from its usual, this must intervene in the prediction in an inversely proportional way to σ_v.

As the system is built to predict every missing rating of *i*, predictions have to be computed even if fewer than $k_{\min}$ potential neighbors, that is, with $w_{uv} \geq w_l$, are available. In this latter case, Equation (3) is replaced by:

$$\widehat{r}_{ui} = \frac{1}{2}\left(\overline{r}_u + \overline{r}_i\right) \qquad (4)$$

where $\overline{r}_u$ and $\overline{r}_i$ are the average rating of user *u* and the average rating received by PDD *i*, respectively. This simple model takes into account both user behavior and PDD popularity. The choice of $k_{\min}$ is thus crucial since it affects the proportion of ratings obtained by (9). The optimal values of parameters w_l, $k_{\min}$ and $k_{\max}$ are determined by cross-validation.

Predicting with a normalization scheme
Once the nearest neighbors of u who rated i are selected, the prediction $\widehat{r}_{ui}$ is computed by applying Equation (3). Note that the Z-score normalization scheme can lead to impossible predictions when $\sigma_v = 0$. In this case, σ_v is fixed to 0.1 to specify that v is used to systematically rate in a similar fashion.

After estimating each unknown $\widehat{r}_{ui}$ for i, the collected and predicted ratings about i are gathered in the vector r_i.

Learn: understanding the customers
Finally, the Learn phase starts with tacking stock of the information collected during the experiments. From an effectual perspective, unexpected contingencies can be leveraged and should be seen as opportunities to learn (Funken *et al.*, 2020; Harmeling and Sarasvathy, 2013; Nielsen and Sarasvathy, 2016). However, authors such as Sanz-Velasco (2006) also suggest that causation might be valuable in a learning phase, since it involves choosing from among the given alternatives (like in A/B testing) rather than generating new ones. It might even mean combining both logics when finding solutions that build on existing knowledge while still requiring substantial development (Sanz-Velasco, 2006). While the prior phases suggest a sequential combination of logics, entrepreneur would benefit most of the BMPL loop by engaging in hybrid practices in the Learning phase. The learning opportunities from the AI-based system are described below.

The ratings in r_i give the development team a simple way to understand users' opinions about the new PDD i. If each rating option has a clear meaning as: 1 = disturbing, 2 = useless, 3 = improvable, 4 = useful and 5 = essential (in a 5-star rating scale), the proportions of users giving each possible answer contribute to understanding customer reactions to the feature added. For example, a high proportion of 3s in r_i would indicate that improvements need to be made to the feature associated to the corresponding PDD, but that the idea behind it is well accepted by the majority of users. In contrast, a high proportion of 1s would demonstrate a wrong PDD. In the latter case, the development team would have avoided wasting time to perfectly develop a feature that was highly rejected by the users for several months.

In addition to ratings, a new metric is proposed: the product potential. It measures the proportion of users for whom the PDD i is of interest. For instance, if 85% of the ratings are positive, that is, above a threshold r_l, the product potential is 0.85. This makes it possible to quickly evaluate the enthusiasm of users for a new PDD with a simple indicator. Coupled with traditional indicators, such as churn rate, time on page and bounce rate, the collected and predicted ratings, as well as the product potential can provide valuable information for the next steps of the product design.

These ratings as along with the product potential are simple ways to evaluate the extent to which a new PDD is well received by users. The newly collected ratings are added to the rating matrix R of size $(n \times m)$, where n is the number of users and m the number of PDDs, by adding a new column.

In conclusion, the BMPL loop has the potential to articulate causal and effectual decision-making logics. In the next sections, the model is implemented in a case study, the "EasyTips" project.

Case study
Data set presentation
To estimate the validity of the proposed BMPL loop, we conducted an empirical study with the collaboration of an entrepreneurial project: EasyTips. EasyTips is a mutual aid platform where students can share course content and find relevant information about their questions. The data set was formed through an online survey in which we asked EasyTips users to

evaluate 12 features of the platform. Note that each PDD here corresponds to one feature of the digital platform. The way in which opinions are solicited has a significant impact on the quality of feedback collected. First, it has to be adapted to the purpose of the product. As EasyTips aims to help students with their studies, a good indicator to evaluate a feature is its usefulness. In contrast, for a video game, the questions and the choices might instead focus on the extent to which each feature is entertaining. Second, the ratings gathered should express clear feelings about the features. For instance, directly asking for a rating to evaluate a feature is too vague. Indeed, a 3 can express many different opinions: *"I do not know what to think about this feature," "I do not use this feature much," "I like it but it should be improved," "this feature bugs half the time," "I do not hate it," "It is sometimes useful," "It is useless but it does not bother me"* and so forth. Countless opinions can be summarized by a single rating value. Therefore, the questions were formulated as: *"What do you think about the feature < name of the feature>? "* illustrated by a screenshot of the feature in action. The answer options were:

"harmful to the proper use of the website" – "useless" – "good idea but not like this,"

"useful" – "essential."

The possibility was given to skip a question when the corresponding feature was not known to the user, so the rating matrix is not fully plain. Assuming that the differences in appreciation between each semantic choice can be considered as regular, each answer is transformed into an integral rating. Thereby, each *"harmful to the proper use of the website"* is transformed into a 1, each *"useless"* into a 2, each *"good idea but not like this"* into a 3, each *"useful"* into a 4 and each *"essential"* into a 5.

The EasyTips data set consists of 1,880 ratings from 164 users about 12 features where each feature is associated with a PDD. The proportions of each value among the 1880 ratings are shown in Figure 4. The data set is particular for two reasons: the rating matrix is highly dense (95.53%) and is much smaller than those commonly used in collaborative filtering recommender systems. This data set allows us to carry out credible experiments on a real set of data collected in an entrepreneurial context. The small size of the data set is the main reason that led us to choose a neighborhood-based approach rather than a model-based approach. Indeed, model-based approaches usually perform better but should only be preferred if the model has enough data in the offline phase to train the model. In addition, due to the small number of PDDs, a user-user variant is the only relevant choice since a pertinent set of nearest-neighbors cannot be computed with only 11 possible neighbors.

Simulating the integration of a new product design decision
In order to imitate the measure and predict phases of the BMPL loop with the EasyTips data set, the implemented algorithm loops on each PDD and hides one at a time. For instance, at the first iteration, PDD 1 is seen as the newly implemented PDD, that is,. feature 1 has just been implemented, and the others as the known PDDs, that is, corresponding features are known by the users, for which we already have the ratings.

The first 10% of ratings given to PDD 1 are kept and the others are hidden, that is, removed from $\mathcal{R}$. The similarity weights and the predictions are computed from $\mathcal{R}$. Then 10% more ratings are uncovered, and the process is repeated until 50% of ratings are uncovered. Besides imitating the gradual arrival of ratings from users, this progressive addition of ratings in $\mathcal{R}$ serves for studying the proportion of ratings needed to compute relevant predictions.

Evaluating the results
In order to evaluate the quality of predictions, they were classified into two classes: ratings unfavorable to the new PDD (the negative class), and ratings favorable to it (the positive
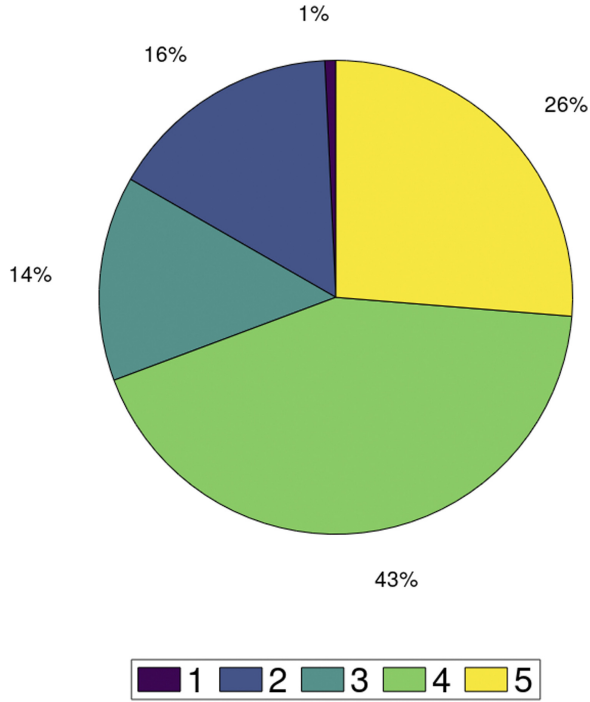
Figure 4.
Proportions of ratings
in the EasyTips
data set

class). A prediction $\widehat{r}_{ui}$ is seen as negative if $\widehat{r}_{ui} < r_l$, and positive otherwise. We set the threshold r_l to 3 − mean absolute error (MAE) – as it corresponds, in a 1-5 scale, to the answer option from which the opinion can be considered as positive, and subtracting the MAE allows to take into account the general precision of the method as the MAE corresponds to the mean error made in absolute value. Such a process avoids considering a predicted rating of 2.9 as a negative rating, whereas it should likely be seen as a 3.

The performance of the algorithm is assessed with both error metrics and accuracy metrics. Therefore, the RMSE (6), MAE (7), precision (8), recall (9) and F-measure (10) are computed with $\mathcal{R}_{test}$ being the set of triplets (u, i, r_{ui}) hidden with i being the PDD considered as new.

The above metrics are mathematical ways to assess the performance of both predictions and classification. However, as the objective of this research is to demonstrate the relevance of the BMPL loop from a business viewpoint, we also evaluate the performance of the algorithm in predicting the product potential PP). For instance, if 85% of ratings collected during the survey for PDD i are higher than r_l, the associated real PP is 0.85. Then, if r_i (i.e. the vector containing known and predicted ratings about i) contains 89% of ratings higher than r_l, the estimated PP is 0.89. The product potential error (PPE) in this example is −0.04. The PPE is computed by:

$$\text{PPE} = \text{PP} - \widehat{\text{PP}} = \frac{\#\text{fn} - \#\text{fp}}{\#\text{ratings}}, \tag{5}$$

where #fn and #fp are the number of false negative and false positive (see Appendix), and #ratings is the number of ratings known plus the number of ratings predicted, which is equal

to the number of ratings collected by the survey for the PDD i. Note that a negative PPE indicates that the algorithm has been too optimistic, while a positive PPE indicates pessimistic predictions.

The active learning version of the algorithm
The key idea behind active learning is that a machine learning model can achieve good accuracy with only a few well selected labeled training instances. The challenge is thus to select the appropriate training instances to label. Active learning is relevant in our case since the labels, that is,. the ratings, are difficult, time-consuming or expensive to obtain (Settles, 2009). The chosen active learning query framework is the following. The active learning build-measure-learn algorithm (AL-BMPLA) computes the similarity weights and $\widehat{r}_{ui}$ for all hidden ratings. The next 10% of ratings to uncover are those for which $\widehat{r}_{ui}$ is closest to r_l. The ratings thus collected are then used to recompute the similarity weights and the predictions. This is done from 10% to 50% of ratings collected. Uncovering ratings predicted close to the limit r_l (also called the "border") is an intelligent way to select the data to label. Indeed, these ratings correspond to the critical instances since a slight modification of the predictions would have produced a different classification. In a digital application, this active learning query framework could allow the selection of the users that should be questioned after collecting some ratings randomly.

Results
The parameters are identical for both the traditional BMPLA and the AL-BMPLA and are summarized in Table 1. The parameter n_c corresponds to the minimum number of common evaluated items by two users u and v to allow the computation of w_{uv}. The values were determined by cross-validation.

Figures 5 and 6 show error metrics, the F-measure and the PPE for each feature for both variants of the algorithm after collecting (i.e. uncovering) 30 and 50% of ratings. Note that the RMSE is always above the MAE in the graphs. It can be noted that for both BMPLA and AL-BMPLA, collecting more ratings tends to reduce the PPE. Looking at the other metrics, the RMSE, MAE and F-measure are not systematically improved with the collection of 20% more ratings when applying the BMPLA. In contrast, it significantly enhances the RMSE, MAE and F-measure with the AL-BMPLA. Together with the fact that metrics are almost always better with AL-BMPLA than with BMPLA, this demonstrates the relevance of the active learning query framework. Note, however, that comparing the metrics with and without active learning is not strictly rigorous from a mathematical point of view. Indeed, the metrics are not determined on the same test set since the hidden ratings are not the same in both cases. However, comparing the metrics with and without active learning from a product design viewpoint enables us to measure the extent to which the approach precisely predicts the opinions of users with the same amount of data collected.

Looking at the results of AL-BMPLA with 30% of known ratings, good results are observed for each metric. The MAE is around 0.4, which means that the ratings are predicted

Parameter	Value
n_c	4
$k_{\max}$	40
$k_{\min}$	1
w_l	0.5

Table 1.
Parameters of the
algorithms

with an average absolute error of 0.4. The RMSE is always higher than the MAE, which indicates that few higher errors exist. The F-measure is always higher than 80% and is well improved when collecting 20% more ratings. Assessing the PPE, we see that for eight PDDs out of twelve, the AL-BMPLA enables us to reach a PPE lower than 0.05 in the absolute value, which means that for those PDDs, after collecting 30% of the ratings, the EasyTips team would have been able to know the proportion of users favorable to the PDD with an error of less than 0.05. Three other PDDs lead to a PPE lower than 0.1 and one leads to a PPE between 0.1 and 0.15 in the absolute value. Note that the four cases are widely improved when collecting 50% of the ratings. It can also be observed that when PPEs are negative, it indicates that the algorithm tends to be too optimistic. It is partially due to the higher proportion of positive ratings in the data set, as seen in Figure 4, and the choice of $r_l = 3 - MAE$. As shown in Figures 5 and 6, the AL-BMPLA performs unequally according to the PDD. When assessing the error metrics, it seems that PDDs 4 and 8 lead to a weaker performance. Graphs on F-measure and PPE indicate that PDDs 4, 8, 9 and 12 present less convincing results. Intuitively, error metrics can be especially affected when the spread of ratings is large, and the two other indicators are more sensitive to the proximity of the average rating with the border r_l. This can be confirmed by inspecting the value of the mean and standard deviation of ratings given to each PDD in Table 2. Therefore, computing the mean and standard deviation of ratings collected for a new PDD allows to determine if predictions are likely to be difficult.

The graphs represented in Figure 7 show the evolution of PPE as a function of the percentage of ratings collected for each PDD. Analyzing the graphs gives a rough approximation of the proportion of ratings to collect when a new PDD appears in the system for a fixed limit PPE. Indeed, by determining both the mean and standard deviation of the new PDD, it is possible to bring it closer to one of the previous PDDs. For example, if the new PDD 13 looks like PDD 2 regarding $\bar{r}_{13}$, we can tell that we need to collect about 20% of 164 ratings to reach a PPE of less than 5%. On the contrary, if the new PDD 13 resembles PDD 8 when comparing their means, we would conclude that the PPE, after collecting approximately 50% of 164 evaluations, is likely to be about 5%.

Figure 5.
Error metrics, F-measure and PPE after collecting 30% of ratings available with BMPLA (blue) and AL-BMPLA (red)
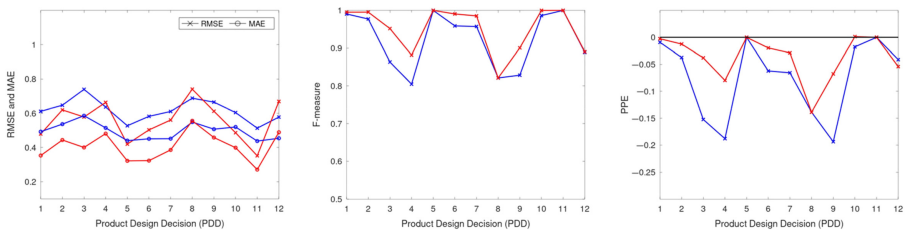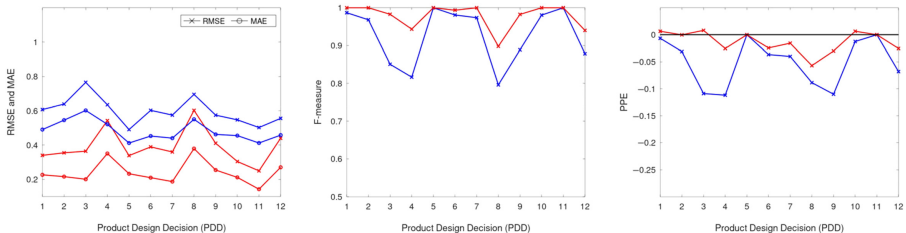

Figure 6.
Error metrics, F-measure and PPE after collecting 50% of ratings available with BMPLA (blue) and AL-BMPLA (red)

Beyond PP, the collection of data about users and the related predictions can provide opportunities for further analysis, such as segmentations. Knowing the ratings collected and predicted as well as the type of features of the EasyTips data set presented in Table 3, the set of users can also be segmented in a data-driven manner. To this end, we gathered all known ratings into a rating matrix R of size $(n \times m)$ and ran the neighborhood-based method to fulfill the unknown elements. When R is plain, a k-means method with $k = 3$ is used to cluster users on the basis of their rating patterns. The parameter k is set at 3, because it is the value for which the clusters are the most interpretable in terms of users' opinions. In Figure 8, the clusters are pinpointed by showing the mean rating that users of each cluster gave to each PDD. Three distinct categories of users can be observed. Cluster 1 represents users who find features of EasyTips useful and are highly satisfied by the website, even though they seem more critical of some features. Cluster 3 gathers users with very clear-cut opinions about EasyTips. They are convinced by both the file sharing and document quality review, but not by social network features. Cluster 2 contains the users who are only convinced by the file sharing features. Therefore, in addition to the advantages presented above, the BMPL loop allows the segmentation of the set of users according to their tastes. Assessing the number of users in each computed cluster provides valuable information too. For instance, in Figure 9,

PDD	1	2	3	4	5	6	7	8	9	10	11	12
$\bar{r}_i$	4.31	4.14	3.37	3.17	4.49	3.78	3.53	3.01	3.34	4.50	4.68	2.91
$\hat{\sigma}_i$	0.68	0.90	0.85	0.97	0.64	0.84	0.92	0.95	0.94	0.71	0.54	0.95

Table 2.
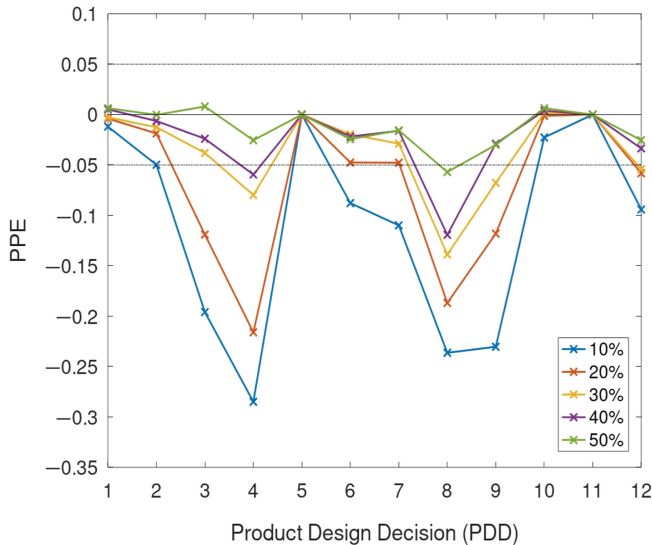Mean and standard deviation of ratings for each product design decision (PDD) of the EasyTips data set

Figure 7.
PPE after collecting 10, 20, 30, 40 and 50% of ratings available with AL-BMPLA

PDD/Feature	File sharing	Document quality review	Social network
	1, 2, 5, 10, 11	6, 7	3, 4, 8, 9, 12

Table 3.
Type of each feature associated to each PDD of the EasyTips data set
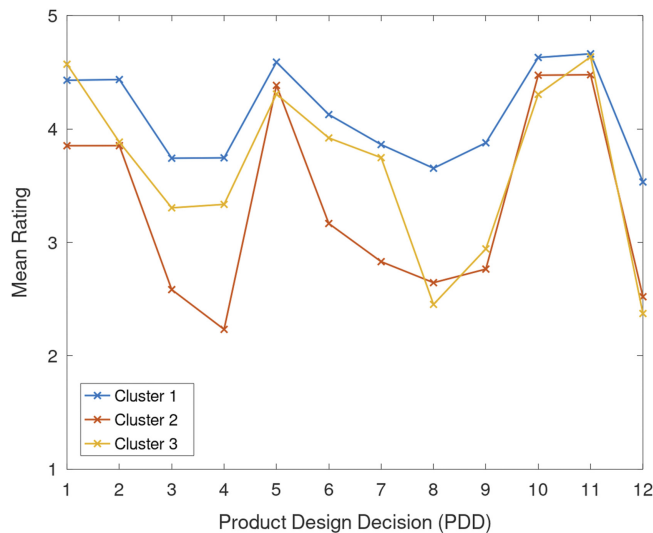
the proportion of users in each cluster is illustrated. According to the pie chart, almost half of users are highly satisfied by EasyTips. However, it becomes clear that features designed to review the quality of documents and for social interactions might need to be rethought. For this purpose, a new iteration might integrate interviews rather than prediction only in order to better understand the reasons for their low ratings (when the market is seen as foreseeable) or to embrace their feedbacks as new contingencies (in an unknowable future).

Discussion

At the core of this paper lies one paradox. As new AI-based infrastructures are available to deal with more uncertainties (Nambisan, 2017), entrepreneurs might get stuck in a quest for optimal solution rather than an affordable one, thus threatening their ability to face remaining unexpected contingencies (Mansoori and Lackeus, 2020). Building on effectuation theory, a possible solution would be to implement LSAs as a source of hybrid practices, instantiating predictive (causal) and controlled (effectual) logics of action. Thus, the proposed model articulates effectual and causal phases while, at the same time, examining its technical implementation in a concrete case study, the "EasyTips" entrepreneurial project. The results given by the BMPL algorithms, and especially by its active learning variant, provide support for the relevance of the BMPL loop for opportunity development. In addition to strong prediction performance, it has the ability to prevent less relevant PDDs. Both qualities provide entrepreneurs with a way to rapidly evaluate the impact of a new PDD and thus generate entrepreneurial learning without waiting for many users to test and evaluate their choices.

Doing so, the model first contributes to the conversation about hybrid practices and the key mechanisms that help entrepreneurs blend causal and effectual logics (Galkina *et al.*, 2021; Galkina and Jack, 2021). It clarifies how LSAs can enable blended logics and shows how the BML loop can be extended to incorporate a predict phase into each iteration. Second, it contributes to digital entrepreneurship research. The current article proposes an algorithm able to predict what users think about a new PDD – that is, the implementation of a new feature or the improvement of an existing feature – with very little collected feedback – that is, ratings. More importantly, it highlights the implication of machine learning on decision-
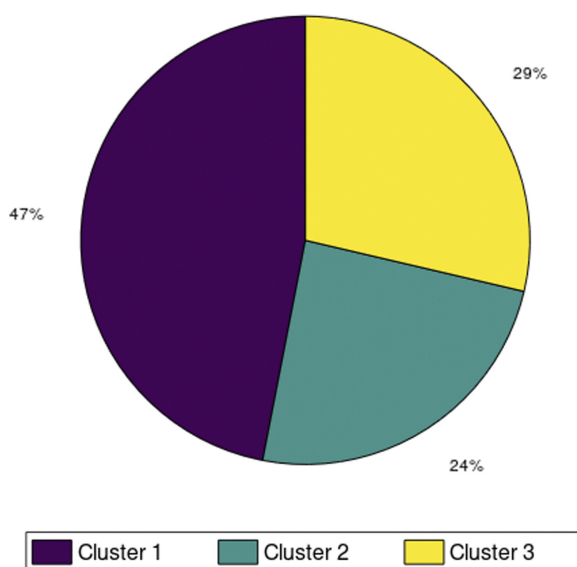
29%

47%

24%

Cluster 1   Cluster 2   Cluster 3

making in situation of risk and uncertainty (Nambisan, 2017). Especially, the model tackles uncertainties in three ways: (1) through epistemological expansion by accelerating the pace of experiments and reducing the time required to perform each iteration, (2) through ontological reduction by transforming previously uncertain elements into "knowable" content by the machine, beyond the limitation of human cognition and (3) through methodological scaffolding, by providing a practical framework that enables hybrid predictive (causal) and controlled (effectual) practices.

First, the BMPL loop allows for reduced uncertainty by compressing the amount of time required to perform an iteration and thus allowing for extended experimentations. Indeed, we provide evidence that entrepreneurs would need to collect ten times fewer ratings to assess the potential of their product (within an acceptable margin of error) while at the same time being able to monitor the strength of predictions' performances. In a real-life application where uncollected ratings are unknown, having such an indication is valuable. In this instance, uncertainty is framed as "epistemological" (Mansoori and Lackeus, 2020) since the future is considered knowable in principle and uncertainty can be mitigated through augmented information gathering strategies (Fiet and Patel, 2008). Entrepreneurs expand their knowledge – and reduce uncertainty – thanks to the AI-enabled acceleration of experimentation and time compression of each iteration.

Second, the AI model enables to bring predictions to situations that were not predictable in the past. For instance, additional feedback (textual feedback, churn rate, click through rate, conversion rate, etc.) and ratings collected during the shortened measure phase can provide with red flags when the system fails at correctly predicting the PP of a new PDD. Likewise, the model can reveal implicit segmentations of users that are unexpected for the entrepreneurs. This is about discovering "unknown unknowns" (Teece and Leih, 2016) that were previously unforeseeable. This enlargement of the field of the predictable is akin to an "ontological" reduction of uncertainty, whereby situations that would have previously been seen as "unknowable in principle" are now knowable through active learning.

Finally, the AI model allows for dealing with uncertainty through methodological "scaffolding" by providing a practical framework that enables hybrid predictive (causal) and

controlled (effectual) practices. The articulation of the prediction phase into a framework that accommodates both effectual and causal logics of action allows to face remaining uncertainty by leveraging unexpected contingencies and embracing a future unknowable in principle.

By exploring "How entrepreneurs can harness prediction technologies while keeping a control-based, effectual approach for product development," this work provides entrepreneurs with a primer on how to implement predictive AI-based infrastructures at the service of their entrepreneurial project. The related indicators, the PP and the PPE, bring significant information to generate entrepreneurial learning. Bringing more assessable metrics certainly facilitates the use of AI-based infrastructure by entrepreneurs. Another practical implication is the reinforcement of collaboration between users and the venture team. The LSAs already call for user and customer involvement in the product and business development (Frederiksen and Brem, 2017). The BMPL loop reinforces it by giving enhanced ways to involve users. Asking users to rate a feature on a 5-star scale rating system is playful and user-friendly, particularly if the rating system is integrated into the digital product. "Easy to provide" feedback is likely to encourage more users to be involved in the co-creation process. Frederiksen and Brem (2017) linked the user involvement advocated by the LSAs with open-innovation, which is positively correlated with innovation performance (Cheng and Huizingh, 2014) and crucial for boosting the speed of internationalization of new digital products (Shaheer and Li, 2020).

Conclusion and future work

This paper proposes a new manner to harness AI-based prediction technologies through the LSAs as a toolkit that accommodates both effectual and causal logics of action. An adaptation of the Lean startup's BML loop integrates an AI-based causal stage, called the predict phase.

This research provides the details of the active learning BMPL algorithm. It allows the prediction of users' opinion, in the form of ratings, of a new PDD from ratings users have given to other PDDs and some randomly collected ratings about the new PDD. Coupled with a simple active learning query framework to guide the rating collection phase, the so-called active learning BMPLA is evaluated against a data set obtained from a real entrepreneurial project. The quality of the algorithm is showcased through its prediction performance and its ability to detect design decisions likely to lead to less convincing results. In addition, the related ratings and the PP are accessible metrics. It enables the acceleration of the opportunity development by shortening the measure phase and facilitating the learn phase, which makes the BMPL loop substantially faster than the BML loop.

Consistently with Galkina and Jack (2021) or Smolka et al. (2018), the proposed approach is a novel way to implement hybrid practices for opportunity development. While the measure and predict phases better fit with a causal logic of action, their integration into the BMPL loop allows for hybrid practices in an iterative way. Uncertainty is managed in three complementary manners: through epistemological expansion by accelerating information gathering, through ontological reduction by turning previously unforeseeable elements into "knowable" content and through methodological scaffolding by providing a framework that enables hybrid practices.

This research also has key limitations. First, the case study focuses on the theoretical development of the model and its technological feasibility from a case study. Further studies can investigate how entrepreneurs deal with epistemological expansion, ontological reduction and methodological scaffolding, and what are the boundary conditions for each of them. Implementing the BMPL loop in a startup and studying its day-to-day functioning would bring valuable information. For example, reactions of the startup team, reactions of users, the impact on the collaboration between users and the startup team and therefore on open-innovation and the impact on team collaboration are some interesting paths to explore.

Especially, the literature on effectuation calls for more studies on the tensions stemming from hybrid practices (Galkina *et al.*, 2021). Since teams are a possible source of effectuation-causation combinations (Ben-Hafaïedh and Ratinho, 2019) and managing teams is another key weakness of LSAs (Mansoori and Lackeus, 2020), it opens intriguing paths for research. Using a digital infrastructure such as the AI-based model developed in this paper might help in creating hybrid practices while at the same time bringing tensions between individuals with different preferences (Alsos *et al.*, 2016). Interestingly, the digital infrastructure might also act as boundary objects to connect team members and develop shared understanding of the opportunity (Shepherd *et al.*, 2021).

Another important limitation is related to the relative acontextual nature of the research. However, it is now established that the need for (and performance of) combined causal and effectual logics differs according to circumstances (Chen *et al.*, 2021; Vanderstraeten *et al.*, 2020). This is especially important since LSAs might be more useful in some sectors (i.e. digital sectors) than in others. Future studies might look at the context-dependent nature of the BMPL, across different sectors and levels of turbulence.

Technical extensions at the service of business opportunity development can also be explored. Extending recommendation techniques to incorporate multi-criteria is considered as one of the important issues for the next generation of recommender systems (Adomavicius *et al.*, 2011). It might be an interesting way to enhance predictions and the interpretability of the active learning BMPL algorithm when dealing with new opportunity with complex value propositions. The time-changing dynamic can also be incorporated into some model-based methods in order to take into account changes in tastes or behavior of users as well as changes in the popularity of PDDs (Koren and Bell, 2015). This is especially important for the study of blended logics, which are considered as essential in dynamic environments (Laine and Galkina, 2017). Another perspective would be to integrate textual reviews of users to enhance the predictions, as in Bao *et al.* (2014) or Lei *et al.* (2016), or even to detect unexpected contingencies in the written comments as sources of new pivots. Finally, prediction technique could be extended to compare concurrent PDDs and to advise on the preferred ones in an A/B testing method. As such, technical expansions would be expanding the digital infrastructures that provide entrepreneurs with learning advantages.

This paper can serve as a starting point for new research into AI-based digital infrastructures and the way they can help entrepreneurs in their day-to-day operations. It can also encourage more research on the use of hybrid predictive (causal) and controlled (effectual) practices through the LSAs. Up until now, LSAs have been framed as an operationalization of effectuation. This work can encourage theoretical developments, building on effectuation and causation, to better understand Lean startup practices, especially when supported by digital infrastructures accelerating the entrepreneurial process.

Note

1. A summary of the mathematical and conceptual elements from recommender systems theory is proposed in Appendix.

References

Adomavicius, G. and Tuzhilin, A. (2005), "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 No. 6, pp. 734-749, doi: 10.1109/TKDE.2005.99.

Adomavicius, G., Manouselis, N. and Kwon, Y. (2011), "Multi-criteria recommender systems", in Ricci, F., Rokach, L., Shapira, B. and Kantor, P.B. (Eds), *Recommender Systems Handbook*, Springer US, Boston, pp. 769-803.

Alsos, G.A., Clausen, T.H., Hytti, U. and Solvoll, S. (2016), "Entrepreneurs' social identity and the preference of causal and effectual behaviours in start-up processes", *Entrepreneurship and Regional Development*, Vol. 28 Nos 3-4, pp. 234-258.

Amoroso, S., Moncada-Paternò-Castello, P. and Vezzani, A. (2017), "R&D profitability: the role of risk and Knightian uncertainty", *Small Business Economics*, Vol. 48 No. 2, pp. 331-343, doi: 10.1007/s11187-016-9776-z.

Andries, P., Debackere, K. and van Looy, B. (2013), "Simultaneous experimentation as a learning strategy: business model development under uncertainty", *Strategic Entrepreneurship Journal*, Vol. 7 No. 4, pp. 288-310.

Baker, T. and Nelson, R.E. (2005), "Creating something from nothing: resource construction through entrepreneurial bricolage", *Administrative Science Quarterly*, Vol. 50 No. 3, pp. 329-366.

Bao, Y., Fang, H. and Zhang, J. (2014), "TopicMF: simultaneously exploiting ratings and reviews for recommendation", *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI'14)*.

Ben-Hafaïedh, C. and Ratinho, T. (2019), "Entrepreneurial behaviour and effectuation: an examination of team formation processes", in *Entrepreneurial Behaviour*, Springer, pp. 91-118.

Berends, H., Jelinek, M., Reymen, I. and Stultiëns, R. (2014), "Product innovation processes in small firms: combining entrepreneurial effectuation and managerial causation", *Journal of Product Innovation Management*, Vol. 31 No. 3, pp. 616-635.

Blank, S. (2005), *The Four Steps to the Epiphany: Successful Strategies for Products That Win*, K&S Ranch.

Blank, S. (2013), "Why the lean start-up changes everything", *Harvard Business Review*, Vol. 91 No. 5, pp. 63-72.

Chen, D., Zhang, D. and Liu, A. (2019), "Intelligent Kano classification of product features based on customer reviews", *CIRP Annals*, Vol. 68 No. 1, pp. 149-152, doi: 10.1016/j.cirp.2019.04.046.

Chen, J., Liu, L. and Chen, Q. (2021), "The effectiveness of effectuation: a meta-analysis on contextual factors", *International Journal of Entrepreneurial Behavior and Research*, Vol. 27 No. 3, pp. 777-798.

Cheng, C.C.J. and Huizingh, E.K.R.E. (2014), "When is open innovation beneficial? The role of strategic orientation", *Journal of Product Innovation Management*, Vol. 31 No. 6, pp. 1235-1253, doi: 10.1111/jpim.12148.

Chien, C.-F., Kerh, R., Lin, K.-Y. and Yu, A.P.-I. (2016), "Data-driven innovation to capture user-experience product design: an empirical study for notebook visual aesthetics design", *Computers and Industrial Engineering*, Vol. 99, pp. 162-173, doi: 10.1016/j.cie.2016.07.006.

Cremonesi, P., Koren, Y. and Turrin, R. (2010) "Performance of recommender algorithms on top-n recommendation tasks", *in Proceedings of the Fourth ACM Conference on Recommender Systems, Association for Computing Machinery*, RecSys '10, pp. 39-46, doi: 10.1145/1864708.1864721.

Evald, M.R. and Senderovitz, M. (2013), "Exploring internal corporate venturing in SMEs: effectuation at work in a new context", *Journal of Enterprising Culture*, Vol. 21 No. 3, pp. 275-299.

Fan, T., Golder, P.N. and Lehmann, D.R. (2017), "Innovation and new products research: a state-of-the-art review, models for managerial decision making, and future research directions", in Wierenga, B. and van der Lans, R. (Eds), *Handbook of Marketing Decision Models*, Springer International Publishing, pp. 79-116.

Fiet, J.O. (2002), *The Systematic Search for Entrepreneurial Discoveries*, Praeger.

Fiet, J.O. and Patel, P.C. (2008), "Entrepreneurial discovery as constrained, sytematic search", *Small Business Economics*, Vol. 30 No. 3, pp. 215-229, doi: 10.1007/s11187-006-9010-5.

Frederiksen, D.L. and Brem, A. (2017), "How do entrepreneurs think they create value? A scientific reflection of Eric Ries' Lean Startup approach", *International Entrepreneurship and Management Journal*, Vol. 13 No. 1, pp. 169-189, doi: 10.1007/s11365-016-0411-x.

Funken, R., Gielnik, M.M. and Foo, M.-D. (2020), "How can problems be turned into something good? The role of entrepreneurial learning and error mastery orientation", *Entrepreneurship Theory and Practice*, Vol. 44 No. 2, pp. 315-338, doi: 10.1177/1042258718801600.

Galkina, T. and Jack, S. (2021), "The synergy of causation and effectuation in the process of entrepreneurial networking: implication for opportunity development", *International Small Business Journal*, 02662426211045290, doi: 10.1177/02662426211045290.

Galkina, T., Atkova, I. and Yang, M. (2021), "From tensions to synergy: causation and effectuation in the process of venture creation", *Strategic Entrepreneurship Journal*.

Ghezzi, A. (2019), "Digital startups and the adoption and implementation of lean startup approaches: effectuation, bricolage and opportunity creation in practice", *Technological Forecasting and Social Change*, Vol. 146, pp. 945-960.

Ghezzi, A. and Cavallo, A. (2020), "Agile business model innovation in digital entrepreneurship: lean startup approaches", *Journal of Business Research*, Vol. 110 No. C, pp. 519-537, doi: 10.1016/j. jbusres.2018.06.

Green, P.E. and Srinivasan, V. (1978), "Conjoint analysis in consumer research: issues and outlook", *Journal of Consumer Research*, Vol. 5 No. 2, pp. 103-123, doi: 10.1086/208721.

Grégoire, D.A. and Cherchem, N. (2020), "A structured literature review and suggestions for future effectuation research", *Small Business Economics*, Vol. 54 No. 3, pp. 621-639.

Gunawardana, A. and Shani, G. (2015), "Evaluating recommender systems", in Ricci, F., Rokach, L. and Shapira, B. (Eds), *Recommender Systems Handbook*, Springer US, Boston, pp. 265-308.

Harmeling, S.S. and Sarasvathy, S.D. (2013), "When contingency is a resource: educating entrepreneurs in the Balkans, the Bronx, and beyond", *Entrepreneurship Theory and Practice*, Vol. 37 No. 4, pp. 713-744, doi: 10.1111/j.1540-6520.2011.00489.x.

Hauser, J.R. and Clausing, D. (1988), "The house of quality", *Harvard Business Review*, Vol. 66 No. 3, p. 63.

Ireland, R. and Liu, A. (2018), "Application of data analytics for product design: sentiment analysis of online product reviews", *CIRP Journal of Manufacturing Science and Technology*, Vol. 23, pp. 128-144, doi: 10.1016/j.cirpj.2018.06.003.

Jiao, Y., Yang, Y. and Zhang, H. (2019), "An integration model for generating and selecting product configuration plans", *Journal of Intelligent Manufacturing*, Vol. 30 No. 3, pp. 1291-1302, doi: 10. 1007/s10845-017-1324-4.

Johnson, R.M. (1987), "Adaptive conjoint analysis", *Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*.

Kaminski, J.C. and Hopp, C. (2019), "Predicting outcomes in crowdfunding campaigns with textual, visual, and linguistic signals", *Small Business Economics*, Vol. 55, pp. 627-649, doi: 10.1007/ s11187-019-00218-w.

Knight, F.H. (1921), *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston, New York.

Koren, Y. and Bell, R. (2015), "Advances in collaborative filtering", in Ricci, F., Rokach, L. and Shapira, B. (Eds), *Recommender Systems Handbook*, Springer US, Boston, pp. 77-118.

Kraus, S., Palmer, C., Kailer, N., Kallinger, F.L. and Spitzer, J. (2018), "Digital entrepreneurship: a research agenda on new business models for the twenty-first century", *International Journal of Entrepreneurial Behavior and Research*, Vol. 25 No. 2, pp. 353-375, doi: 10.1108/IJEBR-06-2018-0425.

Laine, I. and Galkina, T. (2017), "The interplay of effectuation and causation in decision making: Russian SMEs under institutional uncertainty", *International Entrepreneurship and Management Journal*, Vol. 13 No. 3, pp. 905-941.

Laskovaia, A., Shirokova, G. and Morris, M.H. (2017), "National culture, effectuation, and new venture performance: global evidence from student entrepreneurs", *Small Business Economics*, Vol. 49 No. 3, pp. 687-709.

Lei, X., Qian, X. and Zhao, G. (2016), "Rating prediction based on social sentiment from textual reviews", *IEEE Transactions on Multimedia*, Vol. 18 No. 9, pp. 1910-1921, doi: 10.1109/TMM.2016.2575738.

Li, Q., Maggitti, P.G., Smith, K.G., Tesluk, P.E. and Katila, R. (2013), "Top management attention to innovation: the role of search selection and intensity in new product introductions", *Academy of Management Journal*, Vol. 56 No. 3, pp. 893-916, doi: 10.5465/amj.2010.0844.

Ling, S.H., San, P.P., Chan, K.Y., Leung, F.H.F. and Liu, Y. (2014), "An intelligent swarm based-wavelet neural network for affective mobile phone design", *Neurocomputing*, Vol. 142, pp. 30-38, doi: 10.1016/j.neucom.2014.01.054.

Liu, B. (2012), "Why is there a need for uncertainty theory", *Journal of Uncertain Systems*, Vol. 6 No. 1, pp. 3-10.

Louviere, J.J. and Woodworth, G. (1983), "Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data", *Journal of Marketing Research*, Vol. 20 No. 4, pp. 350-367, doi: 10.1177/002224378302000403.

Luce, R.D. and Tukey, J.W. (1964), "Simultaneous conjoint measurement: a new type of fundamental measurement", *Journal of Mathematical Psychology*, Vol. 1 No. 1, pp. 1-27, doi: 10.1016/0022-2496(64)90015-X.

Mansoori, Y. and Lackeus, M. (2020), "Comparing effectuation to discovery-driven planning, prescriptive entrepreneurship, business planning, lean startup, and design thinking", *Small Business Economics*, Vol. 54, pp. 791-818, doi: 10.1007/s11187-019-00153-w.

Nagamachi, M. (1995), "Kansei Engineering: a new ergonomic consumer-oriented technology for product development", *International Journal of Industrial Ergonomics*, Vol. 15 No. 1, pp. 3-11, doi: 10.1016/0169-8141(94)00052-5.

Nambisan, S. (2017), "Digital entrepreneurship: toward a digital technology perspective of entrepreneurship", *Entrepreneurship Theory and Practice*, Vol. 41 No. 6, pp. 1029-1055, doi: 10.1111/etap.12254.

Nielsen, K. and Sarasvathy, S.D. (2016), "A market for lemons in serial entrepreneurship? Exploring type I and type II errors in the restart decision", *Academy of Management Discoveries*, Vol. 2 No. 3, pp. 247-271, doi: 10.5465/amd.2014.0108.

Ning, X., Desrosiers, C. and Karypis, G. (2015), "A comprehensive survey of neighborhood-based recommendation methods", in Ricci, F., Rokach, L. and Shapira, B. (Eds), *Recommender Systems Handbook*, Springer US, Boston, pp. 37-76.

Obschonka, M. and Audretsch, D.B. (2020), "Artificial intelligence and big data in entrepreneurship: a new era has begun", *Small Business Economics*, Vol. 55, pp. 529-539, doi: 10.1007/s11187-019-00202-4.

Pfeffer, J. and Sutton, R.I. (2006), "Evidence-based management", *Harvard Business Review*, Vol. 84 No. 1, pp. 62-74.

Prüfer, J. and Prüfer, P. (2020), "Data science for entrepreneurship research: studying demand dynamics for entrepreneurial skills in The Netherlands", *Small Business Economics*, Vol. 55, pp. 651-672, doi: 10.1007/s11187-019-00208-y.

Rao, V.R., others (2014), *Applied Conjoint Analysis*, Springer, Berlin, Heidelberg.

Ricci, F., Rokach, L. and Shapira, B., others (2015), *Recommender Systems Handbook*, Springer, Boston.

Ries, E. (2011), *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*, Crown, New York.

Sándor, Z. and Wedel, M. (2005), "Heterogeneous conjoint choice designs", *Journal of Marketing Research*, Vol. 42 No. 2, pp. 210-218, doi: 10.1509/jmkr.42.2.210.62285.

Sahut, J.-M., Iandoli, L. and Teulon, F. (2019), "The age of digital entrepreneurship", *Small Business Economics*, Vol. 56, pp. 1159-1169, doi: 10.1007/s11187-019-00260-8.

Sanz-Velasco, S.A. (2006), "Opportunity development as a learning process for entrepreneurs", *International Journal of Entrepreneurial Behavior and Research*, Vol. 12 No. 5, pp. 251-271.

Sarasvathy, S.D. (2001), "Causation and effectuation: toward a theoretical shift from economic inevitability to entrepreneurial contingency", *Academy of Management Review*, Vol. 26 No. 2, pp. 243-263.

Sarasvathy, S.D. (2003), "Entrepreneurship as a science of the artificial", *Journal of Economic Psychology*, Vol. 24 No. 2, pp. 203-220, doi: 10.1016/S0167-4870(02)00203-9.

Sarasvathy, S.D. (2021), "The middle class of business: endurance as a dependent variable in entrepreneurship", *Entrepreneurship Theory and Practice*, doi: 10.1177/10422587211015983.

Settles, B. (2009), "Active learning literature survey", Technical Report, University of Wisconsin-Madison.

Shaheer, N.A. and Li, S. (2020), "The CAGE around cyberspace? How digital innovations internationalize in a virtual world", *Journal of Business Venturing*, Vol. 35 No. 1, p. 105892, doi: 10.1016/j.jbusvent.2018.08.002.

Sheng, J., Amankwah-Amoah, J. and Wang, X. (2017), "A multidisciplinary perspective of big data in management research", *International Journal of Production Economics*, Vol. 191, pp. 97-112, doi: 10.1016/j.ijpe.2017.06.006.

Shepherd, D.A. and Gruber, M. (2020), "The lean startup framework: closing the academic–practitioner divide", *Entrepreneurship Theory and Practice*, Vol. 45 No. 5, pp. 967-998, doi: 10.1177/1042258719899415.

Shepherd, D.A., Souitaris, V. and Gruber, M. (2021), "Creating new ventures: a review and research agenda", *Journal of Management*, Vol. 47 No. 1, pp. 11-42.

Smolka, K.M., Verheul, I., Burmeister–Lamp, K. and Heugens, P.P. (2018), "Get it together! Synergistic effects of causal and effectual decision–making logics on venture performance", *Entrepreneurship Theory and Practice*, Vol. 42 No. 4, pp. 571-604.

Sussan, F. and Acs, Z.J. (2017), "The digital entrepreneurial ecosystem", *Small Business Economics*, Vol. 49 No. 1, pp. 55-73, doi: 10.1007/s11187-017-9867-5.

Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., Guo, Z., Lu, S.C.-Y. and Nee, A.Y.C. (2019), "Digital twin-driven product design framework", *International Journal of Production Research*, Vol. 57 No. 12, pp. 3935-3953, doi: 10.1080/00207543.2018.1443229.

Teece, D. and Leih, S. (2016), "Uncertainty, innovation, and dynamic capabilities: an introduction", *California Management Review*, Vol. 58 No. 4, pp. 5-12.

Vanderstraeten, J., Hermans, J., van Witteloostuijn, A. and Dejardin, M. (2020), "SME innovativeness in a dynamic environment: is there any value in combining causation and effectuation?", *Technology Analysis and Strategic Management*, Vol. 32 No. 11, pp. 1277-1293.

von Bloh, J., Broekel, T., Özgun, B. and Sternberg, R. (2020), "New(s) data for entrepreneurship research? An innovative approach to use Big Data on media coverage", *Small Business Economics*, Vol. 55, pp. 673-694, doi: 10.1007/s11187-019-00209-x.

Wang, W.M., Wang, J.W., Li, Z., Tian, Z.G. and Tsui, E. (2019), "Multiple affective attribute classification of online customer product reviews: a heuristic deep learning method for supporting Kansei engineering", *Engineering Applications of Artificial Intelligence*, Vol. 85, pp. 33-45, doi: 10.1016/j.engappai.2019.05.015.

Yang, X., Sun, S.L. and Zhao, X. (2019), "Search and execution: examining the entrepreneurial cognitions behind the lean startup model", *Small Business Economics*, Vol. 52 No. 3, pp. 667-679, doi: 10.1007/s11187-017-9978-z.

Appendix
Mathematical and conceptual elements from recommender systems theory

Recommender systems address the concern of guiding users in a personalized way to content or products that might interest to them. They play an important role in applications such as Netflix, Amazon, YouTube, Spotify and Facebook, thanks to more than twenty years of research in this field.

Therefore, the challenge of matching users or consumers with appropriate content or products has become crucial in order to enhance users' satisfaction and loyalty (Adomavicius and Tuzhilin, 2005). The general principle is to predict user u's interest for an item i based on information about what u usually appreciates. Recommender systems are usually classified into the following categories: content-based, collaborative filtering and hybrid approaches (combining the first two categories). While content-based techniques recommend items on their similarity to items that the user has expressed an interest in the past, collaborative filtering approaches recommend items to a particular user based on what other users with similar tastes have previously liked. For their better performance in serendipity (Ning *et al.*, 2015), the amount of data that can be processed (Ricci *et al.*, 2015) and its simplicity, we chose to implement a collaborative filtering recommender system (CFRS). For more information about recommender systems the reader is invited to refer to the book by Ricci *et al.* (2015).

Mathematical notations

We define the problem as the prediction of the rating of user u for a new item i. We distinguish predicted ratings from the known ones by denoting as $\widehat{r}_{ui}$ the prediction of r_{ui}. Referring to the set of users and the set of items as $\mathcal{U}$ and $\mathcal{I}$, we can denote as $\mathcal{U}_i$ all users who rated i and as $\mathcal{I}_u$ all items rated by u. $\mathcal{R}$ denotes the set of available triplets (u, i, r_{ui}). The number of users and items are denoted by n and m, respectively.

The computation of $\widehat{r}_{ui}$ is done in the present paper by a neighborhood-based approach. Neighborhood-based recommender systems form a subcategory of CFRS in which predictions are directly made from the collection of known ratings without computing a model. Neighborhood-based approaches are known to perform better on small data sets and be justifiable (Ning *et al.*, 2015). Predicting a rating with such a method consists of computing similarity weights, selecting neighbors and calculating $\widehat{r}_{ui}$ with a normalization scheme.

Computing the similarity weights

There are two types of neighborhood-based approaches: user-user and item-item methods. In a user-user neighborhood-based method, $\widehat{r}_{ui}$ is computed using the ratings given to the item by a subset of users, called neighbors, that have rating patterns similar to u. In an item-item neighborhood-based method, $\widehat{r}_{ui}$ is computed using the ratings of u for items similar to i, called neighbors, that is, items rated similarly by users of the system. In this research, we only consider the user-user alternative since, for the reasons set out in Section 4, it is the most relevant approach for the empirical study conducted. The rest of this section is therefore limited to user-user approach. It is not restrictive since the item-item variant can easily be deducted from the following explanation.

Neighbors are determined by computing the similarities between u and other users. Similarity weights play two important roles in a neighborhood-based recommender system: they enable us to find the nearest neighbors and provide a means to give more or less importance to these neighbors in the prediction calculation. Several approaches can be used to compute the similarities: cosine vector, Pearson correlation (PC), adjusted cosine, mean squared difference, Spearman rank correlation and so forth. We preferentially use the PC method since it compares ratings by removing effects of mean and variance (Ning *et al.*, 2015). The similarity between user u and user v, denoted by w_{uv}, computed with the PC method, is given by (1).

Selecting the neighbors

From these similarities, the prediction computation consists first in selecting the nearest neighbors of user u who rated the new item i. The set of nearest neighbors of u who rated i is denoted by $\mathcal{N}_i(\mathcal{U})$, and is fulfilled by applying a series of conditions detailed in Section 3.

Predicting with a normalization scheme

Once the nearest neighbors of u who rated i are selected, the prediction of its rating for i is computed by (2). Equation (2) uses the Z-score normalization. The objective of Z-score is to convert individual ratings into a more universal scale by taking into account the average and the spread of the ratings that users gave (Ning *et al.*, 2015). The rating of neighbor v further influences the computation of $\widehat{r}_{ui}$ if the r_{vi}

differs from $\bar{r}_v$ and the standard deviation is small. In other words, if a neighbor of u rates i very differently from its usual, this must intervene in the prediction in a way that is inversely proportional to σ_v.

Evaluating the results

Assessing the performance of a neighborhood-based recommender systems is usually performed by either computing errors metrics or accuracy metrics (Cremonesi *et al.*, 2010). Both start by dividing the data set into two parts: the training set – that is, triplets (u, i, r_{ui}) considered as known and thus used to compute similarity weights – and the test set – that is, the set of triplets (u, i, r_{ui}) for which r_{ui} is hidden and computed by (2). Error metrics evaluate how precise the predicted ratings are, while accuracy metrics evaluate how precise the classification is. Classification can consist of considering a prediction $\hat{r}_{ui}$ as positive if $\hat{r}_{ui} \geq r_l$ – with r_l being a threshold – and negative otherwise.

Common error metrics are the root mean squared error (RMSE) and the mean absolute error (MAE) (Gunawardana and Shani, 2015):

$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{R}_{\text{test}}|} \sum_{r_{ui} \in \mathcal{R}_{\text{test}}} \left(\hat{r}_{ui} - r_{ui}\right)^2},$$ (6)

$$\text{MAE} = \frac{1}{|\mathcal{R}_{\text{test}}|} \sum_{r_{ui} \in \mathcal{R}_{\text{test}}} |\hat{r}_{ui} - r_{ui}|,$$ (7)

where $\mathcal{R}_{\text{test}}$ is the set of triplets (u, i, r_{ui}) and r_{ui} is hidden.

Assessing the classification quality is done with accuracy metrics (Gunawardana and Shani, 2015):

$$\text{Precision} = \frac{\#tp}{\#tp + \#fp},$$ (8)

$$\text{Recall} = \frac{\#tp}{\#tp + \#fn},$$ (9)

where $\#tp$, $\#fn$, $\#fp$ and $\#tn$ are the number of true positives, false negatives, false positives and true negatives, respectively. The meanings of tp, fn, fp and tn are summarized in Table A1.

	$\hat{r}_{ui} \geq r_l$	$\hat{r}_{ui} \geq r_l$
$r_{ui} \geq r_l$	tp	fn
$r_{ui} < r_l$	fp	tn

Table A1.
True positive, false negative, false positive and true negative meaning

The harmonic mean of precision and recall, commonly named *F*-measure, is often examined to summarize the effects of both metrics:

$$F - \text{measure} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$ (10)

Corresponding author
Fabian Lecron can be contacted at: fabian.lecron@umons.ac.be