

Can Designers Take the Driver's Seat? A New User-Centered Process to Design with Data and Machine Learning

Sara Colombo^{a*}, Camilla Costa^b

^aDepartment of Industrial Design, Eindhoven University of Technology

^bSchool of Design, Politecnico di Milano

*Corresponding author e-mail: s.colombo@tue.nl

Abstract:

In this paper, we describe a new design process to use data and machine learning (ML) as design materials in generating new, user-centered adaptive systems. Through a case study, we show the possibilities and limits of designing with ML, how UX and ML aspects need to be handled in parallel when envisioning and developing new solutions, and their mutual influence. We argue that designers can autonomously envision and design user-centered, ML-enabled systems if they acquire basic knowledge of ML principles. However, some steps require close collaboration with ML experts. In this new process, designers are involved in both human- and data-centered activities and should use ad-hoc tools to properly operate in this field. The resulting process described in this paper is characterized by uncertainty and risk of failure, which raise concerns about its applicability in any design context. However, it provides a possible path for design-driven innovation through data and ML.

Keywords: machine learning; user experience; design process; design-driven innovation.

1. Introduction

Machine Learning (ML) can generate better product and service experiences, by enabling the creation of personalized, adaptive, and learning interactive systems (Lee & Shin, 2020). Because UX designers may greatly benefit from ML to generate innovative experiences (Yang, Banovic, et al., 2018), design has recently started to explore ML as a new design material and scholars have advocated for “a research and education agenda for the UX and interaction design communities” (Dove et al., 2017). Yang, Scuito et al. (2018) highlighted the need to develop new user-centered design processes for the integration of data and ML. However, design has been slow in developing new approaches or methods in this area (Yang et al., 2020).

Dove et al. (2017) point out that, within companies, designers are usually involved in the development of ML-based solutions once functional decisions have been made, but they do not contribute to the generation of user-centered innovative systems powered by ML from the start. This may be due to the difficulties designers encounter when designing with ML, including understanding AI capabilities, tackling unpredictable outcomes of AI models, and managing AI errors (Yang et al., 2020). To overcome some of these issues, Yang et al. (2020) mention two possible approaches: i) starting from datasets designers build on users' interactions with an existing solution; ii) leveraging existing AI libraries or models.

While we agree on those being two possible avenues, in this work we intend to explore a third way, where designers envision innovative solutions based on ML even *before* the model, or the dataset, is available, in order to creatively explore and fully leverage the possibilities offered by ML. In this view, the ML model is not the starting point, but it is built subsequently, throughout the design process, to enable the intended experience or solution. We therefore envision a change of perspective, where designers lead a design-driven innovation process and define the requirements of a ML model according to their design goals.

To test the viability of this approach, we performed a Research-through-Design (RtD) process, where we engaged in the design of *Procrastinate no more* - a ML-enabled solution to reduce procrastination, and we reflected on our design activity to generate new theory (Forlizzi, 2010; Volontè et al., 2018). We performed the process as designers, with design tools and methods. Whenever needed, we were willing to involve ML experts. As a result, we outline a new user-centered process for the design of ML-enabled solutions. We discuss how uncertainty surrounding ML concretely affects a user-centered design process, the complexity that it yields, and the knowledge designers need to acquire to operate in this field, in order to reduce the risk of failure. Finally, we identify moments in which the collaboration with ML experts is beneficial, or even necessary.

We encourage the design community to discuss, review, and expand the proposed process and to deepen the exploration of how design can innovate with ML, by leveraging its true potential.

2. Designing with ML: a new process

2.1 Case study: *Procrastinate no more*

Procrastinate no more is a digital solution we designed to help people procrastinate less. This project is used as a case study to illustrate a new type of user-centred process that designers can follow when they develop innovative solutions based on data and ML.

2.2 Design brief

Our design process, based on the Design Council's double-diamond process (2005), started with a research activity including a literature review on procrastination (causes, behaviors, effects on wellbeing), user studies (10 semi-structured interviews and a survey involving 98 subjects), and benchmarking. The resulting design brief aimed at developing a personalized solution that could motivate people to procrastinate less by: i) setting realistic, adaptive, and achievable daily goals (i.e. 'to do' lists), to avoid frustration; ii) reinforcing positive behaviors and generating enjoyable experiences; iii) increasing users' awareness about their procrastination behavior.

The first element was key in our decision to use ML to personalize the user's daily 'to do' list.

2.3 Step 1: ML outcome and solution ideation

Based on our knowledge of ML capabilities, we decided to explore the possibility to predict the users' tendency to procrastinate certain activities on a specific day, in order to customize the users' 'to-do' goals accordingly. More specifically, we set out to train a ML model aimed at predicting the user's procrastination index (PI) on each day. The ML model would enable the core feature of our solution, i.e., setting personalized and achievable goals based on the predicted PI for that day.

We brainstormed several ideas on how the predicted PI could be used in an interactive, digital solution. We selected a mobile app where the user records a list of activities to complete every day (wishlist), and the system suggests how many activities they should try to accomplish, based on the predicted PI for that day. The goal should positively challenge the user, by slightly exceeding the amount of predicted completed activities. The mobile app keeps track of the user's progress, motivates them throughout the day, and rewards them at the end of the day, if the personal goal is reached.

Relation between solution ideation and ML outcome

In this initial phase (Step 1, Fig. 1), we had to focus on two mutually influencing activities: i) envisioning what outcome we could reasonably expect from a ML model, and ii) designing the general features of a solution based on the expected ML outcome.

These two activities can be carried out by designers, if they understand the capabilities of ML and its working mechanisms, and if they are aware of its technical limitations (Yang, Banovic, et al., 2018; Yang, Scuito, et al., 2018). Such knowledge is necessary to not treat ML as "magic" (Elish & Boyd, 2018) by setting goals that are technically unachievable.

After defining the expected outcome of the ML model and the general design features, our process started to be characterized by a high level of uncertainty. Indeed, there were no guarantees that the envisioned ML model would be able to accurately predict the PI of a person, on a specific day. However, this goal served as a guide to start exploring ways to build such a model.

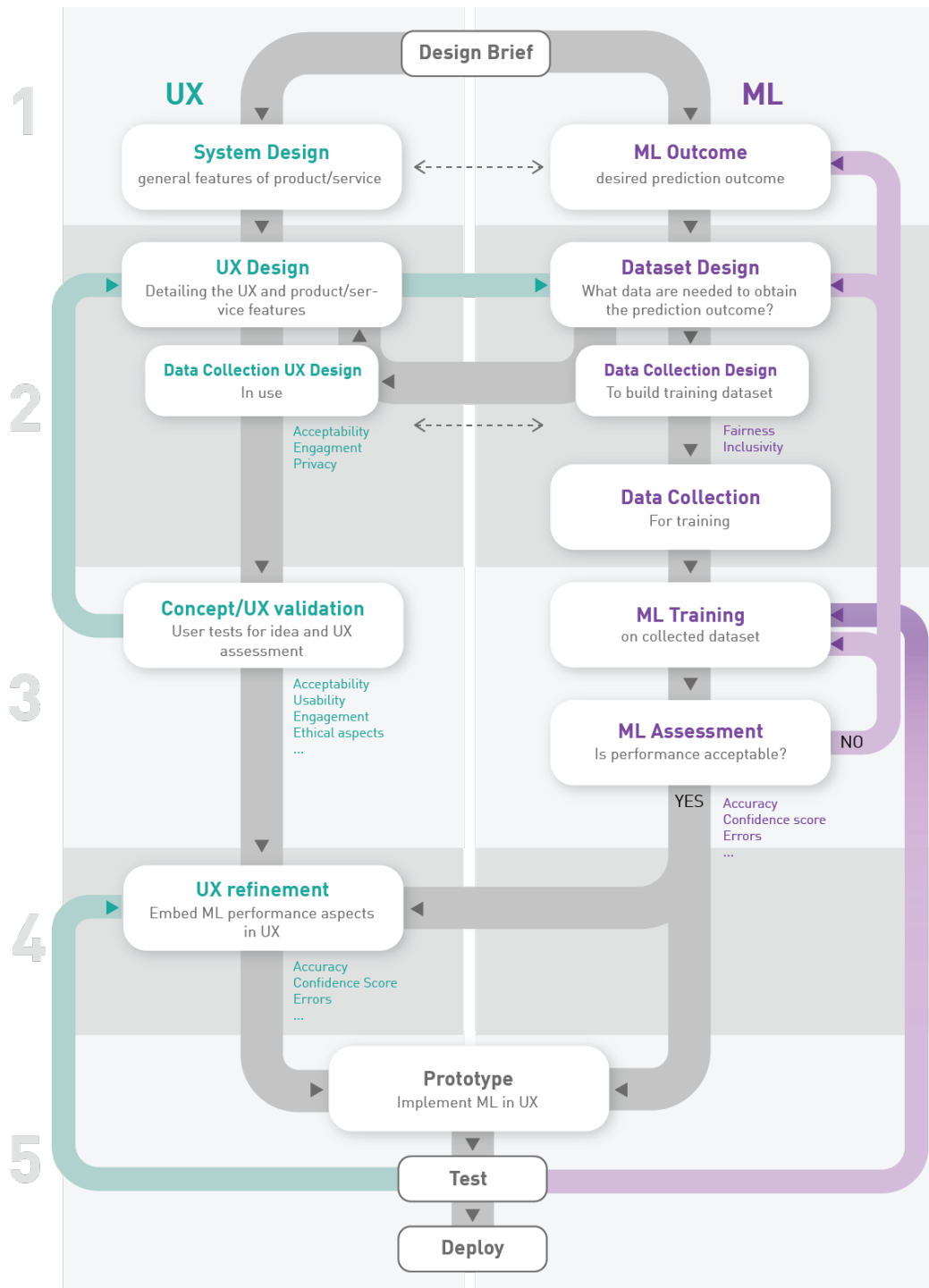


Figure 1. The diagram represents the iterative design+ML process that emerged from our RtD activity. It illustrates how the use of ML requires designers to engage in new activities and tasks, and it shows the relationship between UX and ML aspects. The grey path represents the ideal process. The green and purple paths indicate possible iterations.

2.4 Step 2: UX design and Data design

UX design

After designing the overall solution, we started to detail the UX, by generating a preliminary wireframe (Fig. 2). Two UX elements were particularly difficult to outline: how to explain the user the ML presence and its role in setting personalized goals; how much we wanted the user to trust our

system – which depends on the model's accuracy and performance. The latter was a complex aspect to tackle, as we were not aware yet of the features and performance of the ML model, which we would build in subsequent steps (Step 3, Fig.1). However, we realized how much all these elements could impact the final user experience and could affect its design. It became clear that the UX design could not be concluded in this stage but needed to be finalized after defining the ML technical features.

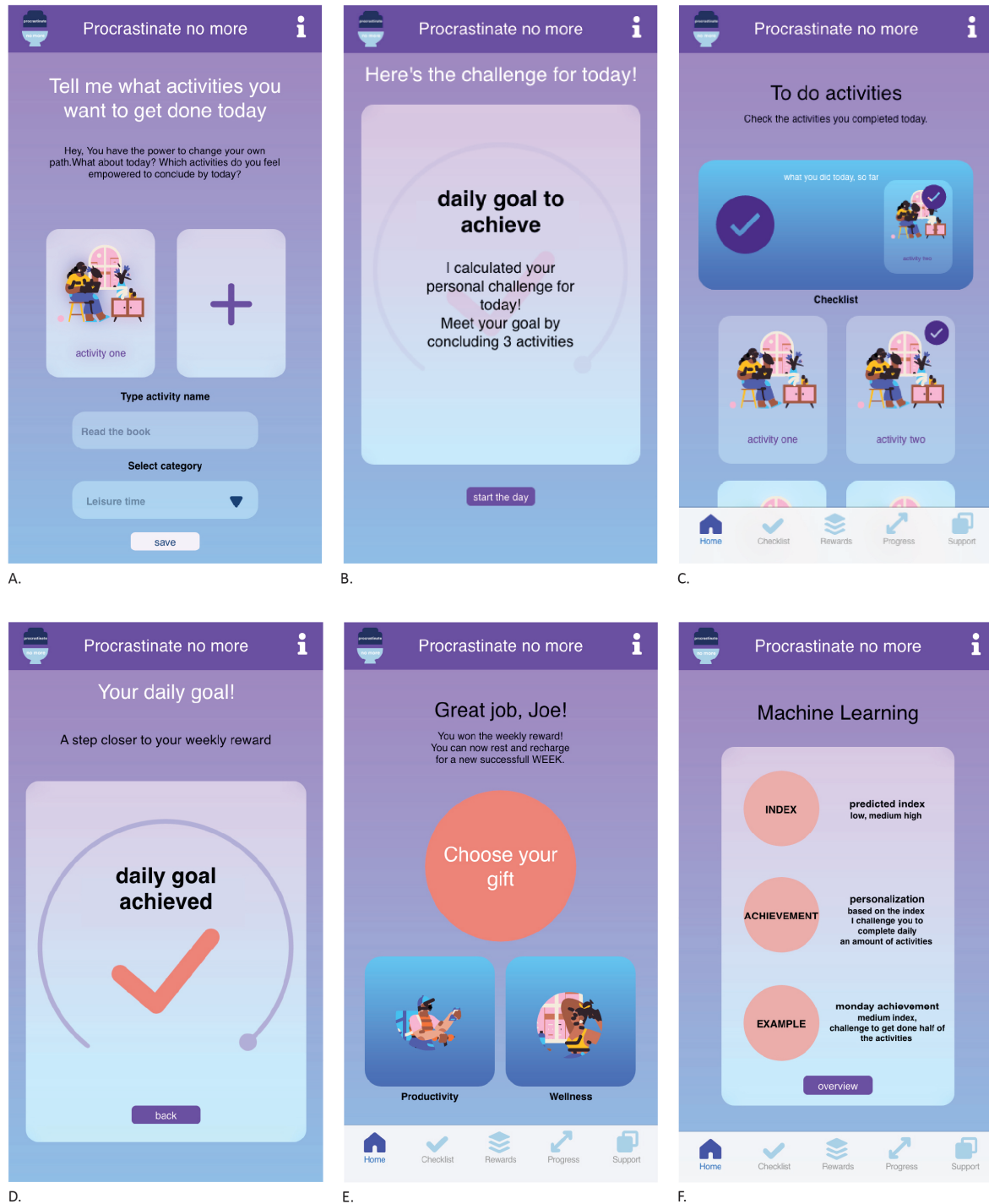


Figure 2. Procrastinate no more, example UX screens. The user lists up to six activities they want to complete by the end of the day, by adding their name and category (A); The ML system predicts the daily PI and sets the goal accordingly (B); The user marks the activities completed throughout the day (C); When the day is over, the system rewards the user if they reached the daily goal (D). At the end of the week, the system rewards the user with a gift to support their fight against procrastination (E). The user can get more information on how ML works (F).

Dataset design

In parallel with the UX design, we started to build a ML model to predict the user's daily PI. We decided to use supervised learning, structured data, and classification algorithms, for which we needed to create a training dataset made of labelled instances. A training dataset includes instances that are used by the ML algorithm to learn that a set of personal or contextual elements (*attributes*) lead to a certain outcome (*class*). The patterns learned by ML will subsequently be used to make predictions. To keep our model simple, instead of predicting the overall daily PI, we decided to predict if each activity added to the wishlist would be procrastinated or not. From there, we could infer the daily procrastination index, if needed. Therefore, the prediction for each individual activity (procrastinated or not) was our *class* (Fig.3). To proceed, we needed to define what attributes could influence our class, based on our knowledge of the phenomenon under study - procrastination.

The primary and secondary research performed at the beginning of our project scoped the main procrastination patterns and uncovered the elements that could influence the user's (conscious or unconscious) decision to procrastinate. We learned that one's procrastination behavior can be influenced by contextual factors - such as the day of the week (weekday/weekend), the weather, the type of activity that should be performed and its duration, but also by personal factors, such as age and personality (e.g., if one procrastinates due to stress, tiredness, or laziness). (Steel, 2001)

As a result, the following three clusters of elements were identified as factors that could potentially influence the procrastination of a certain activity. They became the *attributes* of our dataset, and included:

- User's profile data (age, gender, type of procrastinator);
- Daily contextual data (weather, day of the week, hours of sleep, level of physical activity of the previous day);
- Activity features (type of activity, duration).

The dataset is reported in Figure 3.

User personal details attributes					Wishlist attributes		Daily contextual data attributes				Class
User	Gender	Lazy	Stress	Not organized	Activity to procrastinate	Duration	Sleep hours	Vitality	Weather	Week/weekend	Procrastinated
1	male	3	2	2	dailyroutine	<15	<5	sedentary	sunny	weekday	yes
1	male	3	2	2	studentduty	>90	<5	sedentary	sunny	weekday	no
1	male	3	2	2	studentduty	>90	<5	sedentary	sunny	weekday	yes
1	male	3	2	2	leisuretime	60-90	<5	sedentary	sunny	weekday	yes
1	male	3	2	2	dailyroutine	30-60	<5	sedentary	sunny	weekday	yes
1	male	3	2	2	work	60-90	<5	sedentary	sunny	weekday	yes
2	male	3	2	2	dailyroutine	<15	<5	slightlyactive	cloudy	weekday	yes
2	male	3	2	2	homeduty	15-30	<5	slightlyactive	cloudy	weekday	yes
2	male	3	2	2	studentduty	>90	<5	slightlyactive	cloudy	weekday	yes
2	male	3	2	2	studentduty	>90	<5	slightlyactive	cloudy	weekday	no
2	male	3	2	2	leisuretime	60-90	<5	slightlyactive	cloudy	weekday	yes
2	male	3	2	2	work	60-90	<5	slightlyactive	cloudy	weekday	yes
3	male	3	2	2	dailyroutine	<15	5,6	sedentary	cloudy	weekday	yes
3	male	3	2	2	work	30-60	5,6	sedentary	cloudy	weekday	yes
3	male	3	2	2	studentduty	>90	5,6	sedentary	cloudy	weekday	yes
3	male	3	2	2	mandatoryduty	<15	5,6	sedentary	cloudy	weekday	yes
3	male	3	2	2	studentduty	15-30	5,6	sedentary	cloudy	weekday	no
3	male	3	2	2	leisuretime	30-60	5,6	sedentary	cloudy	weekday	yes
4	male	3	2	2	dailyroutine	<15	5,6	sedentary	cloudy	weekday	no
4	male	3	2	2	dailyroutine	15-30	5,6	sedentary	cloudy	weekday	yes
4	male	3	2	2	studentduty	60-90	5,6	sedentary	cloudy	weekday	yes
4	male	3	2	2	leisuretime	>90	5,6	sedentary	cloudy	weekday	yes

Figure 4. A sample of the Procrastinate no more dataset. The dataset is composed of 11 attributes and one class. Each instance corresponds to one activity inserted in the wishlist.

The dataset design turned out to be a balancing act. Selecting the right attributes is essential to building a reliable ML model – if we select attributes that are loosely connected to the predicted outcome, the ML model may not be able to predict it accurately. On the other side, it is important to state that the ML model can also be built to test some new hypotheses that are not present in

literature, as ML can uncover unexpected and unpredictable relations between attributes and outcome.

In modelling the phenomenon, designers should consider if and how the data related to the selected attributes can be collected to build the training dataset (e.g. through sensors, users' inputs, or other sources – such as API) and how they can be collected once the solution is deployed. The two collection strategies (to populate the training dataset and to feed the model during use) indeed, might differ, leading to the design of two separate user experiences.

Data collection for ML training

In order to train the ML model, we needed to populate our dataset. First, for each attribute, we listed possible ways to collect related data, e.g. through APIs, wearable tracking devices, or direct user's inputs (e.g. through a survey or chatbot). Then, we selected the most suitable method, which in our case was manual data collection through a user diary. We involved 8 subjects and asked them to fill in a digital diary with both contextual and activity data every day for 14 days. Profile data were collected only once. Every day, starting from Day 2, we asked the users to confirm if the activities entered on the previous day had been completed, in order to obtain our class labels (procrastinated or not).

The resulting training dataset consisted of 306 instances (rows), each one corresponding to one activity added to the wishlist (see Fig.3). Although we knew the dataset could be too small to achieve accurate predictions, we decided to proceed with it and to add more instances later, if needed.

Because of the scale of our study, we did not consider elements such as fairness and inclusivity in the data collection, which however are a crucial part of any dataset design and data collection process. Designers can, and should, bring a user-centered perspective to this activity, not only by planning engaging and effective ways to collect training data, but also by designing inclusively and by envisioning strategies to reach out to different types of users, to reduce the chances of building biased datasets.

Data collection during use & UX design

The training dataset provided more detail on the data the model needs to receive during use, i.e. once trained and deployed, to predict the outcome. In this phase, it was important to decide what data needed to be collected through the mobile app vs other sources, and how to collect them, as that would affect the design of the UX. For instance, we discussed if the wishlist activities were to be collected as manual inputs, or voice inputs, and in what steps of the interaction (e.g. homepage, or as prompts), in order to smoothly integrate the data collection during use in the overall UX. Designing the experience of the data collection during use was an essential part of the overall UX design. At the same time, defining how to collect data through other sources (e.g. web APIs, connected devices or sensors) gave a clear view of the ecosystem that had to be built for the ML model to work.

Relation between data design and UX design

The UX design is influenced by the data the ML model needs to receive as input (attributes) in order to predict the outcome (class) during use. Although such data is based on the original training dataset, it may differ from it. Indeed, the training phase might reveal that not all features (attributes) have an influence on the outcome, as explained in the following sections. This would require the UX to be fine-tuned after the ML training phase, in order to collect only relevant data *during use*.

2.5 Step 3: Concept validation and ML training

ML Training

After populating our dataset, we started the ML training phase. This phase can be approached by designers that have little experience of ML with tools for non-experts. We used Weka¹ (Smith Tony C. and Frank, 2016), because it allows to test different ML algorithms on the same dataset, while providing information on the performance of each trained model. Although it has some technical limitations, it was useful for us to feed our dataset to different training algorithms, including One R, J48, Linear regression, and Naïve Bayes.

ML Outcome assessment

None of the trained ML models reached a satisfying level of accuracy. Naïve Bayes performed the better, but its accuracy reached only 60%, possibly due to the limited number of instances in the dataset.

We therefore decided to use a second tool, ILLMO² (Martens, 2014; 2021), in collaboration with a data scientist and ML expert. We did not achieve better accuracy in the prediction, but we extrapolated some insights related to i) the attributes that seemed to play a role in the prediction (i.e. user's gender, activity duration); ii) the fact that individual users have unique behavioral patterns, therefore it might be worth training the model on each single user (which would likely extend the training phase); iii) the rate of false positives (i.e. predicting that an activity will be procrastinated, when it will not) being much higher than false negatives.

Such elements were found to potentially impact the UX on multiple levels:

- The data to collect *during use* might differ from the original training dataset (as not all attributes might actually affect the outcome);
- A different data collection strategy would be necessary to build individual ML models
- Different types of errors (false positives vs false negatives) can greatly affect the UX and need to be carefully considered in designing the solution.

This step was affected by several iterations, as the dataset needed to be pre-processed and adapted to the different ML tools we used. Moreover, unsatisfying results in the ML training required us to try other algorithms, and eventually made it clear that the dataset should be extended to reach higher accuracy. If a richer dataset did not result in the desired performance, new attributes would need to be selected, which would have consequences on Step 1 and 2 as well.

Concept and UX validation

Concurrently to the ML training and assessment, we built and tested a preliminary UX mockup with users, in order to validate:

- Overall design concept - How do users react to adaptable goals?
- UX aspects - Is the solution understandable, usable, engaging, etc.?
- Ethical aspects connected to ML implementation in our mobile app, including:
 - Explainability - What do users want to know about how daily goals are personalized?

¹ <https://www.cs.waikato.ac.nz/ml/weka/>

² <https://researchoutreach.org/articles/illmo-a-new-platform-for-interactive-statistics/>

- Transparency - What do users want to know about how data is being collected and used, and how it contributes to determining the goals?
- Trust - Would users trust the system in setting personal goals? How much should they trust it?
- Feedback - Would users want to give feedback on the ML predictions?
- Overall acceptability of the solution and pain points, especially connected to data/ML:
- Data collection methods - If manual, does it require too much effort? If based on sensors, does the user feel monitored?
- Use of data - Are users concerned about ML fairness and inclusivity?
- Behavior change - Are users comfortable with ML affecting their behavior?

We performed 10 user tests to validate our concept. Results are not reported in detail, however the preliminary UX assessment turned out to be very beneficial, because it provided insights on the users' level of acceptance of the solution, what they would like to understand of the adaptive features and the ML model, and how comfortable they felt with sharing data through the system and with the idea of their behavior being influenced by ML.

Relation between ML training and UX validation

Testing the UX might uncover issues that could require iterating back to previous steps of the process. For instance, if users had difficulties in understanding some features of the solution, the UX would need to be redesigned (Step 2). If they were uncomfortable with what or how data is collected, designers should modify the dataset structure. This would also require adjusting the ML model, as other types of data would be needed to train the model, ultimately affecting both the data design and the ML training phases (Step 3). Although iterating is common in any design process, when designing with ML, such iterations have more far-reaching implications and increase the complexity of the process.

Depending on the available resources and the time constraints of the project, the concept validation and ML training can be performed in parallel (as in our case), or in line, first validating the idea, then proceeding with the ML training.

2.6 Step 4 and 5: UX refinement and ML implementation

Once the UX is validated and the ML model is trained, the UX design can be finalized based on the technical features and performance aspects of the selected ML model (accuracy, types of errors, confusion matrix, etc. - Figure 1, Step 4). The goal is to design a system that can fail 'gracefully' (Google PAIR, 2019).

Following this step, the ML model should be implemented in the UX prototype, and the whole system would be available for final tests with users. In our process, we did not engage in this activity, as we aimed at creating a proof of concept and validating it in Step 3.

3. Discussion

3.1 ML knowledge for UX designers

This paper explores a new design process for using ML as a design material and represents a variant of the traditional user-centered design process. Based on the knowledge gained in our RtD project, we claim that, to carry out similar processes, designers need to:

- Acquire knowledge on ML capabilities, learning approaches, and main algorithms;
- Learn to design with data (designing datasets based on the phenomenon under study, planning data collection strategies both for ML training and during use);
- Understand ML performance aspects (e.g., accuracy score, confusion matrix);
- Recognize ethical issues and learn how to minimize them.

3.2 Relation between UX designers and computer scientists

In our process, we, as designers, ideated and developed a ML-enabled solution independently in Step 1 and 2, with the support of a computer scientist (CS) only in Step 3. We envision CSs' contribution continuing in the following phases, i.e., testing, refining and deploying the solution. The proposed process requires a constant inclusion of ML knowledge, tools, and tasks along the five steps, although ML experts can be involved only when a more technical expertise is crucial.

In Step 1, designers define the ML outcome, and sketch a solution responding to a design brief and the target user. This activity can be carried on by designers, as long as they have basic knowledge of ML capabilities.

In Step 2, designers need to select the ML approach and to design an inclusive, non-biased, and effective data collection strategy for ML training. Designers play a decisive role also in designing how data are collected *during use*, to create engaging and acceptable user experiences. In this step, CSs can help in identifying the best ML approach.

In Step 3, designers and CSs work in parallel on different goals: CSs train and test the ML model, while designers validate the concept and UX, also to detect issues connected to the integration of ML aspects in advance (e.g., privacy, data collection, explainability, acceptance). Although designers can tackle ML training tasks with non-expert tools, our case study shows that a technical contribution is highly beneficial in order to build an up and running ML model, and to correctly assess the outcomes.

In Step 4, designers finalize the UX based on technical aspects, by taking into consideration ML errors and by designing ways to fail gracefully. Finally, in Step 5, they collaborate with CSs to integrate ML in the final prototype, for final user tests.

3.3 Complexity and risks

From our RtD process, we learned that CSs and designers can benefit from each other and work together to contain the risk of failure. However, uncertainties and ambiguities characterize the process, increasing the chances of unsuccess or iterations in the following steps:

- ML outcome definition: misjudging ML capabilities and identifying unrealistic outcomes would result in failure;
- Dataset design: unfitting, limited, or biased training datasets may yield inaccurate predictions;
- Concept/UX validation: design concepts may be rejected by users, requiring to re-design UX details, data collection strategies *during use*, or training datasets;
- ML training: the ML model could be inaccurate, making it necessary to test other algorithms, to modify data - and consequently the UX - or even to abandon the project, if the expected performance is not achieved;
- ML implementation and testing: errors and low accuracy can make the UX experience unacceptable, requiring designing new ways to fail gracefully, or - more effortfully, to improve the ML model.

The cost of these iterations may be much higher than in traditional design processes because any change in ML aspects can deeply impact UX elements, and vice-versa, making both UX and ML decisions tentative and temporary.

3.4 Limitations

Our process was abstracted from the reported RtD study and from knowledge stemming from the authors' prior experiences in designing with ML, but it would benefit from further validation through more examples and cases. The dataset used to train our ML model was limited, and it led to low accuracy, therefore preventing us from implementing the model into the final prototype, and fully testing our solution.

4. Conclusions

Giving designers new tools and methods to design with data and ML would foster design-driven and user-centered innovation in this field, with the creation of adaptive systems that are not just optimized through ML, but fully built on its potential from the start.

From a business viewpoint, it will be relevant to investigate if such a high-risk process is acceptable for companies, and to what extent they would be willing to adopt it, knowing the uncertainties that characterize it. On the other hand, we also argue that such a process would give designers the ability to join from the start, or even lead, innovation processes within companies.

Finally, we argue that designer-entrepreneurs (Colombo et al., 2017) might highly benefit from such an approach in generating innovative adaptive, personalized, and learning solutions, which could enable the creation of design startups hinged on data and ML.

References

- Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science*, 13(3), 219-224.
- Colombo, S., Cautela, C., & Rampino, L. (2017). New design thinking tools for the next generation of designer-entrepreneurs. *The Design Journal*, 20(1), S566–S580.
- Design Council. (2005). The 'double diamond' design process model.
- Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX design innovation: Challenges for working with machine learning as a design material. *Conference on Human Factors in Computing Systems-Proceedings*, 2017-May, 278–288. <https://doi.org/10.1145/3025453.3025739>
- Elish, M.C., & Boyd, D. (2018). Situating methods in the magic of Big Data and AI. *Communication Monographs*, 85(1), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- Google PAIR. (2019). People + AI Guidebook. [Pair.Withgoogle.Com/Guidebook](https://pair.withgoogle.com/guidebook).
- Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, 63(2), 157–170. <https://doi.org/10.1016/j.bushor.2019.10.005>
- Martens, J. B. (2014). Interactive statistics with Illmo. *ACM Transactions on Interactive Intelligent Systems*, 4(1). <https://doi.org/10.1145/2509108>
- Martens, J. B. (2021). Comparing experimental conditions using modern statistics. *Behavior Research Methods*, 53(3), 1240-1261.

- Smith Tony C. and Frank, E. (2016). Introducing Machine Learning Concepts with WEKA. In S. Mathé Ewy and Davis (Ed.), *Statistical Genomics: Methods and Protocols* (pp. 353–378). Springer New York. https://doi.org/10.1007/978-1-4939-3578-9_17
- Steel, P., Brothen, T., & Wambach, C. (2001). Procrastination and personality, performance, and mood. *Personality and individual differences*, 30(1), 95-106.
- Yang, Q., Banovic, N., & Zimmerman, J. (2018). Mapping machine learning advances from HCI research to reveal starting places for design innovation. *Conference on Human Factors in Computing Systems-Proceedings*, 2018-April. <https://doi.org/10.1145/3173574.3173704>
- Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating how experienced UX designers effectively work with machine learning. *DIS 2018-Proceedings of the 2018 Designing Interactive Systems Conference*, 585–596. <https://doi.org/10.1145/3196709.3196730>
- Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020, April 21). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Conference on Human Factors in Computing Systems-Proceedings*. <https://doi.org/10.1145/3313831.3376301>
- Volonté, P., Rampino, L., & Colombo, S. (2018). The Specificity of Design Research: How Practice-Based Design Knowledge Can Enter the Great Archive of Science. In *Advancements in the Philosophy of Design* (pp. 319-345). Springer, Cham.
- Zimmerman, J., Oh, C., Yildirim, N., Kass, A., Tung, T., & Forlizzi, J. (2021). UX Designers Pushing AI in the Enterprise: A Case for Adaptive UIs. *Interactions*, 28(1), 72–77.
- Zimmerman, J., Stolterman, E., & Forlizzi, J. (2010, August). An analysis and critique of Research through Design: towards a formalization of a research approach. In *proceedings of the 8th ACM conference on designing interactive systems* (pp. 310-319).

Author Bios:

Sara Colombo PhD, Assistant Professor at TU/e, Industrial Design. She leads the Inclusive Design and Thoughtful Technology squad. Her research explores the relation between user experience, ethics, and AI. She worked as a researcher and lecturer at Politecnico di Milano, MIT, and Northeastern University.

Camilla Costa Industrial and UX Designer. She received a bachelor's degree in Industrial Design and a master's degree in Design and Engineering from Politecnico di Milano. She is interested in how design can foster innovation and generate positive impacts on people's lives.

Acknowledgements: We would like to thank prof. Jean-Bernard Martens for his help with training ML models in the ILLMO environment with our dataset, and for his insights on the models performance and results.