Conference Paper No: SF-03

Human in the loop design for intelligent interactive systems: A systematic review

N. Arambepola* and L. Munasinghe

Software Engineering Teaching Unit, Faculty of Science, University of Kelaniya, Sri Lanka
nimasha_2019@kln.ac.lk*

Abstract

It is undeniable that modern computers are incredibly fast and accurate. However, computers cannot 'think' (act intelligently) as humans unless it is trained to learn from the past knowledge. Despite their intelligence, humans are comparatively slow in computational tasks. However, the combination of the computational capacity of computers and human intelligence could produce powerful systems beyond the imagination. This concept is called Human-in-the-Loop (HITL) where both human and machine intelligence support the creation of Machine Learning (ML) models. HITL design is an emerging technology which is used in many domains such as autonomous vehicle technology, health systems and interactive system implementations. In this research, we systematically reviewed past research of HITL systems with the objectives of identifying key benefits and limitations of the HITL design. This systematic review was conducted by analyzing 68 research papers published in top-ranked journals and conferences during the past decade. Moreover, the papers were selected using keyword-based searching and references of the most cited HITL research papers. The PRISMA model was used to exclude irrelevant papers, and keyword-based clustering was used to identify the frequent keywords in the selected papers. Although the HITL design often improves the performance of intelligent interactive systems, there are certain drawbacks of this concept when compared to fully manual or fully automated systems such as making decisions with emotional bias and being unable to take actions when demanded. Thus, we comprehensively discuss the approaches proposed by the recent researchers to overcome some of the issues of the existing HITL designs.

Keywords

Human-in-the-loop design, Human intelligence, Intelligent interactive systems, Machine learning, Keyword-based clustering

Introduction

This is the era of Artificial Intelligence (AI) which is the fastest-growing research domain that aims to produce smart solutions to real-world problems. Automated systems such as autonomous cars and intelligent robots are becoming common and essential technologies in day-to-day life. However, some of the recent researchers have investigated the effectiveness of the AI-based systems which includes a part for humans instead of completely automating a system by removing human involvement from the task (Bhardwaj et al., 2014). This concept is called Human-in-the-Loop (HITL) and the main objective of this approach is to provide efficient, intelligent automation for system improvements through human feedback (Holzinger, Valdez and Ziefle, 2016). Here, humans are directly involved in the training, tuning and testing of the ML algorithms. When using a completely automated ML model, there is a certain possibility of having inaccurate results. In HITL design, the general human population can contribute to correct the inaccuracies in machine predictions. HITL systems use supervised ML and active learning together to produce the output. Here, supervised ML is used for the future

predictions through training the algorithms using labelled data. Active learning is used to make the algorithm more accurate and efficient by managing the data (Holzinger, 2016).

Novel technologies have been introduced by investigating the possibility of substituting ML algorithms and optimizing system performance through human experience. For instance, interactive ML techniques have been introduced with novel concepts to improve existing models in different domains. The use of human cognition in the random AI design process is the key benefit to produce outcome design according to human's preferences. It is a common truth that today, people do not have time for choosing their essential daily items. Therefore, AI-based recommender systems have been introduced to make their selections more efficient and easier (Sun et al., 2019). They are commonly used in different domains and fields such as hotels and restaurants, health, fashion, movies, news, and online courses to assist users to discover products with less effort and time. For example, the applicability of the HITL concept for visual analytic tasks was studied by recognizing existing work processes and fitting analytics according to the user feedback (Endert et al., 2014). Nevertheless, most of the existing recommender systems are fully automated systems. There is a huge potential to improve these systems in a more effective way through user-centric design by injecting a human, meaning that a particular domain expert into the system. Thus, we set the main objectives of this literature review as:

1. To identify different domains where HITL design concept can be utilized effectively.
2. To describe how HITL design concept applies in ML based interactive systems.
3. To determine the key benefits and limitations of using HITL for AI-based systems.
4. To explain the approaches that has been proposed to enhance HITL design concept in AI-based systems.

The rest of the paper is organized as follows. In the following section, we describe the flow of our review process with the methodology used. Then, we present the results of our findings with the analysis and the discussion. Finally, we conclude this paper with the future directions of the research.

Methodology

The review design of this work has three main stages. In the first stage, we retrieved research papers and articles from different sources. In the second stage, the contents of each paper were analyzed based on the extracted keywords. Finally, the results were presented as a comprehensive summary.

The most appropriate research papers for this comprehensive literature review were chosen using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) model as shown in figure 1. Here, we collected a total number of 68 research papers and articles related to the fields of intelligent interactive systems, AI and ML applications, and recommender systems that used the HITL design concept. Papers were collected through context keyword searching and choosing the most relevant research papers using the reference section of selected papers that have published in top-ranked journals and conferences during the past decade. The ResearchGate and IEEE Xplore are the main sources used but not limited, to collect research papers. For example, we retrieved papers from Google Scholar. Among the selected papers, 4 duplicate papers and

10 irrelevant research papers were excluded during the identification and screening stages. The remaining 54 papers were eligible for the review but had to exclude 2 papers due to the fact that they are out of the research scope. Then the keyword-based clustering which is an unsupervised ML technique was applied for further analysis. There were five main clusters, after grouping the extracted keywords that are collected from the abstracts of the selected papers and articles. The identified main keywords are machine learning, interactive design, recommender systems, design concepts, Human-in-the-Loop and artificial intelligence.

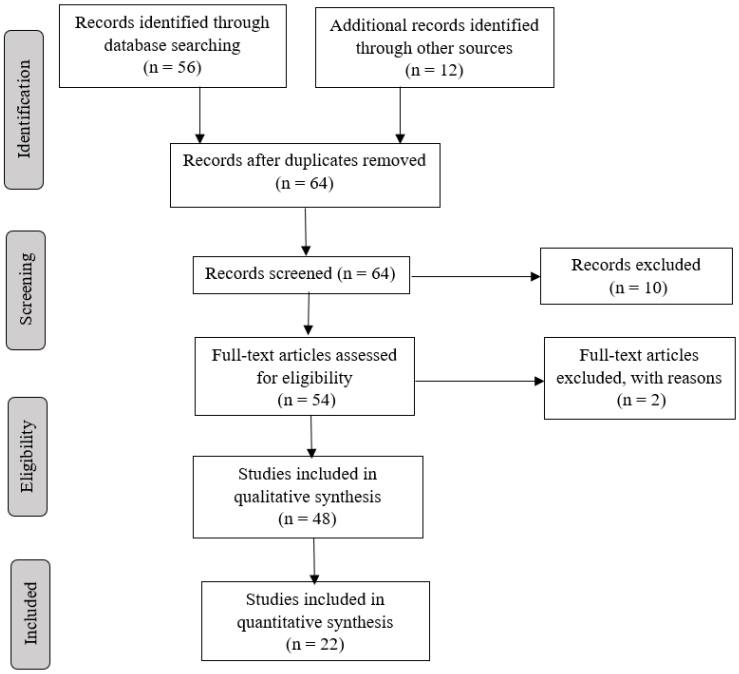

Figure 1. *The approach used to select research papers.*

In the overall review, the HITL design concept which can be used in different software systems was systematically analyzed by examining findings, advantages, limitations of each individual study to provide the knowledge comparatively. The different domains gain significant advantages from using this HITL design concept. Yet, there are certain limitations pertain to this concept, and researchers in the past decade have presented different views to overcome these limitations. Those points were comparatively discussed, and the results were summarized in the next section.

Results and Discussion

According to the literature review, the areas in table 1 are widely used HITL concept. Decision making is very critical in the Healthcare field. Therefore, completely automated systems can be a risk when communicating with patients and their sensitive data. Thus, the Doctor-in-the-loop concept comes into play, and doctors work as expert humans in the domain.

Table 1. *Different areas that use HITL design concept*

Area	Papers from the literature
Health	(Holzinger et al. 2016), (McKinney et al.)
Recommender systems	(Bhardwaj et al. 2014), (Sun et al. 2019), (Holzinger et al. 2016), (Goecks 2020)
Interactive systems	(Emmanouilidis et al. 2019), (Wang et al. 2019a), (Zanzotto 2019), (Wang et al. 2019b), (Goecks 2020), (Endert et al. 2014)

In ML, the HITL design concept comes into play in different scenarios. Following are the identified main cases:

- When there is a lack of data available to train and test the ML model at present
    - In these cases, humans can be used to collect a sufficient amount of data.
    - Humans can make much better judgments in the early stages.
- When there are class imbalances
    - Humans can resolve class imbalance problem if there is any and retrain the ML model.
- When the cost of errors is very high in an ML algorithm.

The key benefits that can be obtained by applying the HITL design concept into intelligent interactive systems can be categorized as shown in table 2. It reveals the importance of the HITL concept when designing sophisticated automated systems, especially for the systems which must frequently interact with general users such as multi-agent systems (Dorri et al., 2018). An example for a successful HITL application is Google's search engine which designed using HITL learning. Thus, it provides better service with the access of more users (Johnson, n.d.). Yet, past researchers have proven that the HITL approach is more applicable in the medical domain because medical experts have a high ability in predicting patients' status where medical datasets and ML system along cannot be achieved (Maadi et al., 2021).

Table 2. *Key benefits of HITL design concept*

Benefits	Description/ Use case
Avoiding bias	Most of the ML models are biased because of the biased data. Biased data can be identified as early as possible having a person-in-the-loop.
Increase the amount of rare and limited data	ML models require a large amount of data sets to train and test models accurately. When there is limited amount of data in some areas, humans can be used to collect more data. For example, Facebook keeps HITL for monitoring users and tracking their actions.
Improve efficiency in systems	Recent research has shown that HITL design concept performs better for some cases. For example, it saves time and results are more accurate than AI or human doctors on their own.

Increase safety of humans	Improve the safety through autonomous vehicles to have lesser accidents. In there, professional gamers can be used to simulate actual driving conditions. In addition, humans use to ensure the safety, when manufacturing critical component for field which required high safety level such as aircraft.
Create employment opportunities	Even though the advancement of AI technologies causes the elimination of current jobs, it also generates new jobs, especially in labelling data for ML. Furthermore, it will increase the accuracy of algorithms (Grønsund & Aanestad, 2020).
Subject expert includes in the loop	When a subject expert is generating and preparing data, the system provides more accurate and appropriate results. For example, Doctor-in-the-loop is a commonly use novel paradigm (Holzinger et al. 2016).
Transparent decision making	When replacing human with a robot to provide services for public, it is better to implement the system with HITL design so that people are more familiar with the system rather than a completely automated AI system.

Despites its advantages, it is worth to examine the limitations of this approach. For example, human emotional decisions make bad choices. In addition, human involvement causes to invite human mistakes, and this is considered a form of data poisoning (Carley & Price, 2021). For example, common human errors such as mislabeling critical data which are used to train ML algorithms would incorrectly produce outputs. It decreases the performance of algorithms. A few of the other limitations are what it can be slow down the AI systems because of taking too much time to make decisions and failed to take immediate actions when needed. Even though HITL design improves the efficiency when compared with fully manual systems, in some cases it can be slow down the system when compared with fully automated systems. The explainable ML model is another concern that still does not has a clear definition (Lipton, 2018). In there, the HITL ML concept is used for making decisions mainly in two ways. One way is, making decisions through a specialized professional in the domain (Zanzotto, 2019). For example, a domain specialist doctor can work as an adviser for the Doctor-in-the-loop system in the medical field (Holzinger, Valdez and Ziefle, 2016). On the other hand, machines make decisions based on the knowledge extracted from data. In there, data are produced by people who are not specialized knowledgeable workers. In both cases, these AI-based systems keep the HITL design concept in a direct or indirect way.

Nevertheless, Zanzotto has identified a huge disadvantage of this approach, and its main victim is the knowledgeable person who works as the adviser for those AI systems (Zanzotto, 2019). The concerns are that the stolen knowledge will produce never-ending revenue, not for the real owner of the knowledge but the company it owns. Because it is a one-time knowledge acquiring process from the knowledgeable person. This will become a huge threat for the AI field as it can cause to reduce the skilled workers who are willing to share their exhausting knowledge to implement AI systems. To tackle this issue, a fairer paradigm named Human-in-the-loop Artificial Intelligence (HitAI) has proposed to repay the decisions of AI systems' revenues to real knowledge owners by tracking each interaction with the automated AI system (Zanzotto, 2019). This will be a turning point towards implementing AI-based systems with the HITL design concept. However, the ethical challenges related to HITL design are yet to be addressed. For

instance, security and privacy are important concerns that may violate through the applications of this concept such as social networking and surveillance applications as they directly communicate with people's sensitive data in their daily lives (Nunes et al., 2015). In future, this research can be extended to comparatively study HITL and Human out of the Loop systems by implementing ML models focusing on human cognitive factors. Even though the HITL design concept is being applied for different areas, it is yet to discover the effectiveness of this concept in the e-Learning and educating fields. This will be a successful step towards increasing human cognition through the combination of emerging technologies such as ML, gamification and HITL design approach.

Conclusion

Even though a considerable number of industries are practicing HITL ML, it is still a relatively new area of AI. Therefore, this research focused on studying how HITL design has been utilized in different AI systems, and their key benefits and limitations through a systematic literature review. Moreover, the approaches proposed by past researchers to overcome some of those disadvantages and limitations were discussed. This systematic analysis was conducted by collecting altogether 68 research papers from different sources. The results of this study can be utilized to effectively expand this design concept for future AI systems and to explore enhancement to overcome the limitations identified in the existing HITL intelligent interactive systems.

References

Bhardwaj, A., Jagadeesh, V., Di, W., Piramuthu, R., & Churchill, E. (2014). Enhancing Visual Fashion Recommendations with Users in the Loop. *ArXiv:1405.4013 [Cs]*. https://arxiv.org/abs/1405.4013

Enabling the human in the loop: Linked data and knowledge in industrial cyber-physical systems. (2019). *Annual Reviews in Control*, *47*, 249–265. https://doi.org/10.1016/j.arcontrol.2019.03.004

Endert, A., Hossain, M. S., Ramakrishnan, N., North, C., Fiaux, P., & Andrews, C. (2014). The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems*, *43*(3), 411–435. https://doi.org/10.1007/s10844-014-0304-9

Goecks, V. G. (2020). Human-in-the-Loop Methods for Data-Driven and Reinforcement Learning Systems. *ArXiv:2008.13221 [Cs, Stat]*. https://arxiv.org/abs/2008.13221

Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, *3*(2), 119–131. https://doi.org/10.1007/s40708-016-0042-6

Holzinger, A., Valdez, A. C., & Ziefle, M. (2016). Towards Interactive Recommender Systems with the Doctor-in-the-Loop. In *Mensch & Computer Workshopband*. https://doi.org/10.18420/muc2016-ws11-0001

Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM*, *61*(10), 36–43. https://doi.org/10.1145/3233231

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. C., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert,

F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., & Ledsam, J. R. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89–94. https://doi.org/10.1038/s41586-019-1799-6

Sun, W., Khenissi, S., Nasraoui, O., & Shafto, P. (2019). Debiasing the Human-Recommender System Feedback Loop in Collaborative Filtering. *Companion Proceedings of the 2019 World Wide Web Conference*. https://doi.org/10.1145/3308560.3317303

Todd, D. (2020, January 10). *Douglas Todd: Robots replacing Canadian visa officers, Ottawa report says*. Vancouversun. https://vancouversun.com/opinion/columnists/douglas-todd-robots-replacing-canadian-visa-officers-ottawa-report-says#:~:text=from%20our%20team.-

Wang, P., Peng, D., Li, L., Chen, L., Wu, C., Wang, X., Childs, P., & Guo, Y. (2019). Human-in-the-Loop Design with Machine Learning. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 2577–2586. https://doi.org/10.1017/dsi.2019.264

Zanzotto, F. M. (2019). Viewpoint: Human-in-the-loop Artificial Intelligence. *Journal of Artificial Intelligence Research*, *64*, 243–252. https://doi.org/10.1613/jair.1.11345

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems*, *29*(2), 101614. https://doi.org/10.1016/j.jsis.2020.101614

Carley, S. S., & Price, S. R. (2021). Analyzing a human-in-the-loop's decisions for the detection of data poisoning. Proc. SPIE 11746, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III, 1174616. https://doi.org/10.1117/12.2586260

Nunes, D. S., Zhang, P., & Sá Silva, J. (2015). A Survey on Human-in-the-Loop Applications Towards an Internet of All. *IEEE Communications Surveys Tutorials*, *17*(2), 944–965. https://doi.org/10.1109/COMST.2015.2398816

Johnson, J. (n.d.). *What Is Human in The Loop (HITL) Machine Learning?* BMC Blogs. Retrieved August 20, 2021, from https://www.bmc.com/blogs/hitl-human-in-the-loop/

Maadi, M., Akbarzadeh Khorshidi, H., & Aickelin, U. (2021). A Review on Human–AI Interaction in Machine Learning and Insights for Medical Applications. *International Journal of Environmental Research and Public Health*, *18*(4), 2121. https://doi.org/10.3390/ijerph18042121

Dorri, A., Kanhere, S. S., & Jurdak, R. (2018). Multi-Agent Systems: A Survey. *IEEE Access*, *6*, 28573–28593. https://doi.org/10.1109/access.2018.2831228