Reinforcement learning for personalization: a systematic literature review (Abstract)

Floris den Hengst[1], Eoin Martino Grua[⋆,1], Ali el Hassouni[⋆,1], and Mark Hoogendoorn[⋆,1]

Vrije Universiteit Amsterdam, `<initial>.<affix>.<surname>@vu.nl`

Abstract. This compressed contribution presents a survey into reinforcement learning (RL) for personalization.

Keywords: Reinforcement Learning, Contextual Bandits · Personalization.

1 Introduction

When products and services are adapted to individual tastes, they become more appealing, desirable, informative, etc. to the intended user than one-size-fits all alternatives. Digital systems enable such *personalization* on a grand scale. The key enabler is data. While the software is identical for all users, the system's behavior can be tailored based on experiences with individual users. Reinforcement learning (RL) has been attracting attention for personalization. An overview of RL for personalization, however, is lacking.

This contribution summarizes our systematic review and categorization of 166 papers describing RL solutions to personalization problems, problem contexts and evaluation strategies across domains [1]. It thus aids researchers and practitioners in identifying relevant related work, promotes the understanding of the usage of RL and identifies challenges across domains. The data used and a tool[1] for exploring it have been made available.

2 Systematic Literature Review

We performed a systematic literature review following the PRISMA standards. We queried five databases on keywords similar to 'reinforcement learning', 'contextual bandit' and 'personalization' and found 983 publications. Titles and abstracts and subsequently full texts were assessed for eligibility, resulting in 166 included papers. For included papers, data on the problem context, solution architecture and evaluation strategy were extracted.

⋆ Authors contributed equally

1 Data exploration tool at https://florisdh.nl/rl4personalization/

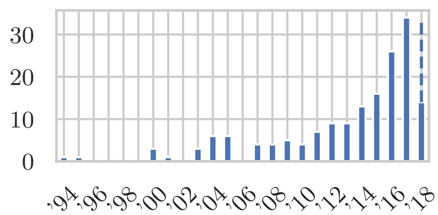

Fig. 1: Number of publications per year, dashed indicates projection for full year.

Domain	#	Domain	#
Health	44	Transport	9
Entertainment	30	Energy	6
Commerce	28	Other	5
Education	25	Smart Home	4
Domain Indep.	11	Communication	4

Table 1: Number of publications per domain.

3 Results

Figure 1 shows a marked increase in publications over time. Table 1 shows that RL is used for personalization in various domains. We continue by looking at problem contexts, solution patterns and evaluation strategies.

Problem In most publications (130/166), users do not provide feedback to the system explicitly, but feedback is derived from various measurements that indicate suitability of system behavior. Data on user responses are available in a minority of cases (66/166) and safety concerns are mentioned in a reasonable number of works (30/166) whereas privacy is not (9/166).

Solutions The RL framework can be used for personalization in various ways: learning a single policy across all users is the most popular approach (91/166). User traits can be included in the state representation in this approach (51/91). The next most popular approach is to represent each user as a separate environment and learn a policy per user (59/166). Hybrid approaches, such as a policy for every group of users, are less popular (11/166). Only a small fraction compares different approaches (5/166). Combining these approaches is an interesting direction for future work, e.g. to increase the level of personalization as more data is obtained

When analyzing the most popular algorithms, we found generic and well-established ones such as Q-learning, contextual bandits and Sarsa to be most popular. Besides these, we find little algorithm re-use. In recent years, approaches that include function approximation with deep neural networks such as DQN or DDQN are becoming more popular.

Evaluation We reviewed the usage of live or real-life data for evaluation and found that the number of studies with such an evaluation is increasing. This indicates that RL has become sufficiently robust to apply in contexts that involve humans. However, the relative number of works that include a realistic evaluation is not increasing. Furthermore, we find that little works compare multiple algorithms. This indicates that the field is growing, but not yet maturing.

References

1. den Hengst, F., Grua, E.M., el Hassouni, A., Hoogendoorn, M.: Reinforcement learning for personalization: A systematic literature review. Data Science (Preprint), 1–41