A RESEARCH PROPOSAL: FORECASTING OF USER-INPUT WITH HUMAN-PROXY GENERATED DATA IN A FEDERATED LEARNING SETTING.

A PREPRINT

Microsoft Research
Jakob Serlier
t-jserlier@microsoft.com \ j.l.j.m.serlier@student.tue.nl

February 16, 2022

1 Proposal

This is a research proposal for real-time forecasting of user-input in video games using a Federated Reinforcement Learning (FRL) model. The proposed research consists of two main milestones; firstly, we aim to create a model which is able to predict user-inputs (e.g. mouse-movement or controller input) in gaming in real-time, based on previous inputs and the environment. The second milestone is to train these models on the devices of users (*participants*) using client-server (cross-device) Federated Learning (FL).

While there are many opportunities in combining federated learning and forecasting, we are looking to apply such models for forecasting user-input in gaming specifically. This could, for instance, be helpful in a cloud-gaming context where the network quality plays a key role in end-user experience. Jarschel et al. [2013] showed that in cloud-gaming, both latency and packet-loss are important indicators to the perceived value and experience of end-users. A model which allows to subtly *support* cloud gaming end-users in case their own capacity to deliver continuous input is hampered could improve the gaming experience. This *supporting* is most promising in the context of packet-loss, where the model could *"interpolate"* or take over user-input in short bursts in both an accurate and personalized manner during such disruptions. Generalizing to real-time forecasting in a federated learning, the research setting has many other use-cases, from predicting pedestrian movements in vehicular automation (e.g. Jayaraman et al. [2020]) to forecasting energy fluctuations in the energy grid (e.g. Venkataramanan et al. [2021]) to forecasting in IoT (e.g. works discussed in Nguyen et al. [2021]).

1.1 Federated Learning and human-proxy data generation

Federated learning (FL) offers a privacy-oriented approach to machine learning where the models are trained collaboratively but the data remains decentralized. One of our contributions with this work will be applying our problem setting in an FL environment. We can categorize our approach to FL as cross-device, since we are working with multiple devices, using a client-server architecture with a single orchestrator. Some practical challenges for cross-device such as the need for error-correction and failure of devices (as described in Kairouz et al. [2021]) are considered to be outside of the scope of this research, as we will be focused on the technical aspects of training the forecasting models in the FL setting. We plan on using Microsoft's internal FLUTE framework as a starting point.

Reinforcement learning has shown to be very capable as a tool for learning models in complex environments. For instance, the seminal paper Mnih et al. [2013] showed the possibility of learning viable control policies in the context of Atari games. The use of such models, specifically pre-trained reinforcement learning agents, is a practical approach to generating user-input data by using the agent as a 'human-proxy'. This is an idea seen before in other applications, for instance, Nadiger et al. [2019] presents a work combining a grouping policy, learning policy and federation policy. They present a novel idea of using Deep Q Network (DQN) reinforcement learning for faster personalization in federated settings, personalizing a Non-Player Character (NPC) in the game pong to the approximate skill level of human-proxy (fixed neural network) player.

The use of such a human-proxy is a logical approach for generating user-input data, especially in a federated setting. For simulating the data-heterogeneity commonly found in cross-device Federated Learning, more varied environments as well as human-proxy's can be used. The work of Gulrez and Mansell [2021] shows to be promising in regards to human-proxy data generation. Although it is not a classical RL model (instead using a perceptual control architecture) it is very useful due to the agent requiring no training at all while still showing high performance. Once we move towards the federated setting, which demands more variability in both the agent's performance (for means of personalization of user-input predictions) as well as the environments, different human-proxy's will be used. For instance, the models from Hafner et al. [2020] which are capable of mastery in 55 Atari games could be used to generate more varied data.

1.2 User-input forecasting

For user-input forecasting, we see two promising directions to take; firstly, we can take a similar approach to Oh et al. [2015]. Their work shows the making and evaluating long-term frame predictions for Atari video games, generating realistic frames up to 100 steps ahead of the prediction starting point. A key difference between their predictions ours is that their predictive model is conditioned on the availability of user-input for the predicted frames, whereas this is exactly what we are aiming to forecast. Using a similar encoder-decoder architecture as used in Oh et al. [2015] would allow for forecasting of the environment, where a separate model could predict the next user-inputs.

The second approach to user-input forecasting is the use of reinforcement learning or imitation learning. It makes sense to explore these methods as a tool to achieve forecasting in a video-game context due to the state-action-reward nature in a video-gaming context. It is important to mention that, in this context, RL would be used as a tool to imitate the player as accurately as possible- not maximizing the score (assuming the predictive model would take over gameplay). In terms of reinforcement learning, prediction could be achieved by means of inverse RL where the reward is based on the predictive (short- and longrange) accuracies of user-input forecasting. Additionally, a logical approach would be imitation learning. Considering our application, the work of Rhinehart and Kitani [2016] on activity forecasting using Inverse Reinforcement Learning (IRL) is an interesting primer for this train of thought. After discussions with my colleague Lekan Molu (Microsoft Research NYC), his current research on Interaction-Grounded Learning (IGL; first introduced in Xie et al. [2021]) could also prove to be a useful tool for user-input forecasting. In IGL, explicit reward is replaced by implicit feedback by means of a latent reward vector. Regardless of the approach taken in order to forecast user-input, it seems imperative to create a model which is able to *understand* the environment in one way or another. The approaches mentioned in this paragraph are promising, but they are inherently more complex and require more work on my end to research the possible applications and mathematical framework required to tackle user-input forecasting. .

Venturing to the practical side of the proposal, we can divide the work into several subtasks. First and foremost, an environmenet needs to be set up where a video game environment and user-input can be simulated and data can be gathered. Using the popular reinforcement learning framework OpenAI Gym allows for many pre-built RL examples which can serve as human-proxy's, such as those mentioned in the previous paragraphs. Next, the various user-input forecasting approaches should be tested on this environment. Finally we will apply the user-input forecasting in a federated learning environment. For this, both human-proxy agents of different skill-levels, as well as heterogeneous environments will be considered.

References

Michael Jarschel, Daniel Schlosser, Sven Scheuring, and Tobias Hoßfeld. Gaming in the clouds: Qoe and the users' perspective. *Mathematical and Computer Modelling*, 57:2883–2894, 6 2013. ISSN 0895-7177. doi:10.1016/J.MCM.2011.12.014.

Suresh Kumaar Jayaraman, Dawn M Tilbury, X Jessie Yang, Anuj K Pradhan, and Lionel P Robert Jr. Walkaway walkaway cross approach wait approach analysis and prediction of pedestrian crosswalk behavior during automated vehicle interactions. 2020.

Venkatesh Venkataramanan, Sridevi Kaza, and Anuradha M. Annaswamy. Der forecast using privacy preserving federated learning. 7 2021. URL `https://arxiv.org/abs/2107.03248v1`.

Dinh C. Nguyen, Ming Ding, Pubudu N. Pathirana, Aruna Seneviratne, Jun Li, and H. Vincent Poor. Federated learning for internet of things: A comprehensive survey. *IEEE Communications Surveys and Tutorials*, 23:1622–1658, 4 2021. doi:10.1109/comst.2021.3075439. URL `https://arxiv.org/abs/2104.07914v1`.

Peter Kairouz, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, Rafael G.L. D'Oliveira, Hubert Eichner, Salim El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaid Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara

Javidi, Gauri Joshi, Mikhail Khodak, Jakub Konecní, Aleksandra Korolova, Farinaz Koushanfar, Sanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, Rasmus Pagh, Hang Qi, Daniel Ramage, Ramesh Raskar, Mariana Raykova, Dawn Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu, and Sen Zhao. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14, 2021. ISSN 19358245. doi:10.1561/2200000083.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. 12 2013. URL `https://arxiv.org/abs/1312.5602v1`.

Chetan Nadiger, Anil Kumar, and Sherine Abdelhak. Federated reinforcement learning for fast personalization. In *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, pages 123–127, 2019. doi:10.1109/AIKE.2019.00031.

Tauseef Gulrez and Warren Mansell. High performance across two atari paddle games using the same perceptual control architecture without training a preprint. 2021. URL `https://github.com/PCT-Models/PCTagent_Breakout_Atari`.

Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. 10 2020. URL `https://arxiv.org/abs/2010.02193v3`.

Junhyuk Oh, Xiaoxiao Guo, Honglak Lee, Richard Lewis, and Satinder Singh. Action-Conditional Video Prediction using Deep Networks in Atari Games. 2015.

Nicholas Rhinehart and Kris M Kitani. First-person activity forecasting with online inverse reinforcement learning. 2016.

Tengyang Xie, John Langford, Paul Mineiro, and Ida Momennejad. Interaction-grounded learning. 2021.