# Machine Learning 2019: Human / AI interaction loop training as a new approach for interactive learning with reinforcement-learning agents: Neda Navidi - Montreal, Canada

**Neda Navidi**
*Montreal, Canada*

Human / AI interaction loop training as a new approach for interactive learning with reinforcement-learning: Reinforcement-Learning (RL) in various decision-making tasks of Machine-Learning (ML) provides effective results with an agent learning from a stand-alone reward function. However, it presents unique challenges with large amounts of environment states and action spaces, as well as in the determination of rewards. This complexity, coming from high dimensionality and continuousness of the environments considered herein, calls for a large number of learning trials to learn about the environment through RL. Imitation-Learning (IL) offers a promising solution for those challenges, using a teacher's feedback. In IL, the learning process can take advantage of human-sourced assistance and/or control over the agent and environment. In this study, we considered a human teacher and an agent learner. I was once inquired by a colleague within the Logic Office here at Stanford in the event that robot performers will ever exist, to which I answered that they may — sometime in the not so distant future — but as it were on the off chance that we to begin with figure out what it implies to have robot logicians. The trade was in fact a bit tongue-in-cheek, but it revealed a blind-spot within the way we conversation almost the longer term of AI: in our propensity to inquire whether or when a given errand will be taken over by robotization, it is simple to disregard the more profound issue of what such a takeover would mean. This may an justifiable oversight when we're considering around fabricating, clerical work or indeed driving a car. We're less concerned with how these errands are finished and more concerned with the result — by and large measured in taken a toll, speed and security. But when we envision "automating" a interest like music making, we're constrained to adjust the item of work with something more profound — the meaning we determine from the method of doing itAs

an Relate Teacher at Stanford University's CCRMA (Center for Computer Investigate in Music and Acoustics — articulated "karma"), I'd like to think I have a one of a kind point of view on a address like this. I plan programming dialects for music, disobedient and melodic toys like Ocarina for the iPhone, coordinate the Stanford Portable workstation Ensemble, and investigate VR/AR plan for music. I am a portion of the Stanford Human-centered AI activity, and my understudies and I inquire about the plan of frameworks that in a general sense implant innovation and human interaction, within the interest of modern apparatuses for melodic and other shapes of human inventiveness. Upheld by a 2016 Guggenheim Partnership, I composed Guileful Plan: Innovation in Look of the Grand, a comedian book approximately the forming of innovation — and how innovation shapes us.As Michael Polanyi once famous around the implied nature of human information, we know more than ready to tell. This is often portion of what gives AI, and profound learning in specific, its unimaginable appeal; its capacity to spot designs in complex wonders that oppose rule-based depictions implies it can basically get it them "for us". Tragically, it too makes it tempting to think of AI as a "Big Ruddy Button" — a innovation that dependably conveys the proper answers whereas stowing away the method that leads to them. Take the idea of imaginative fashion, for occasion. Most of us would concur it's "a thing", but would discover it precarious to characterize, and indeed more troublesome — in the event that not incomprehensible — to expressly program. It's about comical to see these pictures side-by-side. On the cleared out could be a depiction of a merry minute, worlds apart from the apprehension captured on the correct: an existential emergency, a response to the craziness of advanced life, or be that as it may you translate Edvard Munch's The Shout of Nature. But once combined, all sense of meaning goes out the

window. Is our upbeat couple still on get-away? And are they truly not frightened by the stream of magma behind them, or the red hot atmosphere over them?! The style-infused excursion photo is wonderful to see at, but its meaning is underwhelming, casual, kitschy. The Shout, in differentiate, could be a work of Craftsmanship, welcoming us to reflect, to feel. One picture essentially says "oh, I've been there,'' whereas the other shouts, quietly, "I've been there." It's the distinction between fashion — something AI is creating an noteworthy get a handle on of  and meaning something indeed we people still battle withAt last, and maybe most critically, there's the reality that we don't fair esteem the item of our work; we regularly esteem the method. For illustration, whereas we appreciate eating ready-made dishes, we moreover appreciate the act of cooking for its inborn encounter — taking crude fixings and forming them into nourishment. Or take music; we may have get to to more of it than ever some time recently within the shape of recordings numerous of which speak to the exceptional apex of the craftsmanship — but we haven't halted singing, playing and composing for ourselves. From the most punctual days of radio and recording to advanced music, and presently gushing, there's remains—through different innovative innovations an inborn bliss to the movement of making music. It's clear there's something worth protecting in numerous of the things we do in life, which is why mechanization can't be diminished to a straightforward parallel between "manual" and "automatic." Instep, it's about searching for the correct adjust between perspectives that we would discover valuable to mechanize, versus errands in which it might remain meaningful for us to take an interest. As simple because it can be to grasp the extremes — to surge into robotizing everything or to demand on mechanizing nothing perfect arrangements frequently exist someplace in between, as a duality between computerization and human interaction, between independent innovation and the instruments we use basically, the human-in-the-loop approach reframes a mechanization issue as a Human-Computer Interaction (HCI) plan issue. In turn, we've broadened the address of "how do we construct a more astute system?" to "how do we consolidate valuable, important human interaction into the system?" This kind of plan is at the center of inquiring about in areas like Intelligently Machine Learning, in which cleverly frameworks are planned to expand or upgrade the human, serving as a device to be used through human interaction..The teacher takes part in the agent's training towards dealing with the environment, tackling a specific objective, and achieving a predefined goal. Within that paradigm, however, existing IL approaches have the drawback of expecting extensive demonstration information in long-horizon problems. With this work, we propose a novel approach combining IL with different types of RL methods, namely State–action–reward–state–action (SARSA) and Proximal Policy Optimization (PPO), to take advantage of both IL and RL methods. We address how to effectively leverage the teacher's feedback – be it direct binary or indirect detailed – for the agent learner to learn sequential decision-making policies. The results of this study on various OpenAI-Gym environments show that this algorithmic method can be incorporated with different RL-IL combinations at different respective levels, leading to significant reductions in both teacher effort and exploration costs.