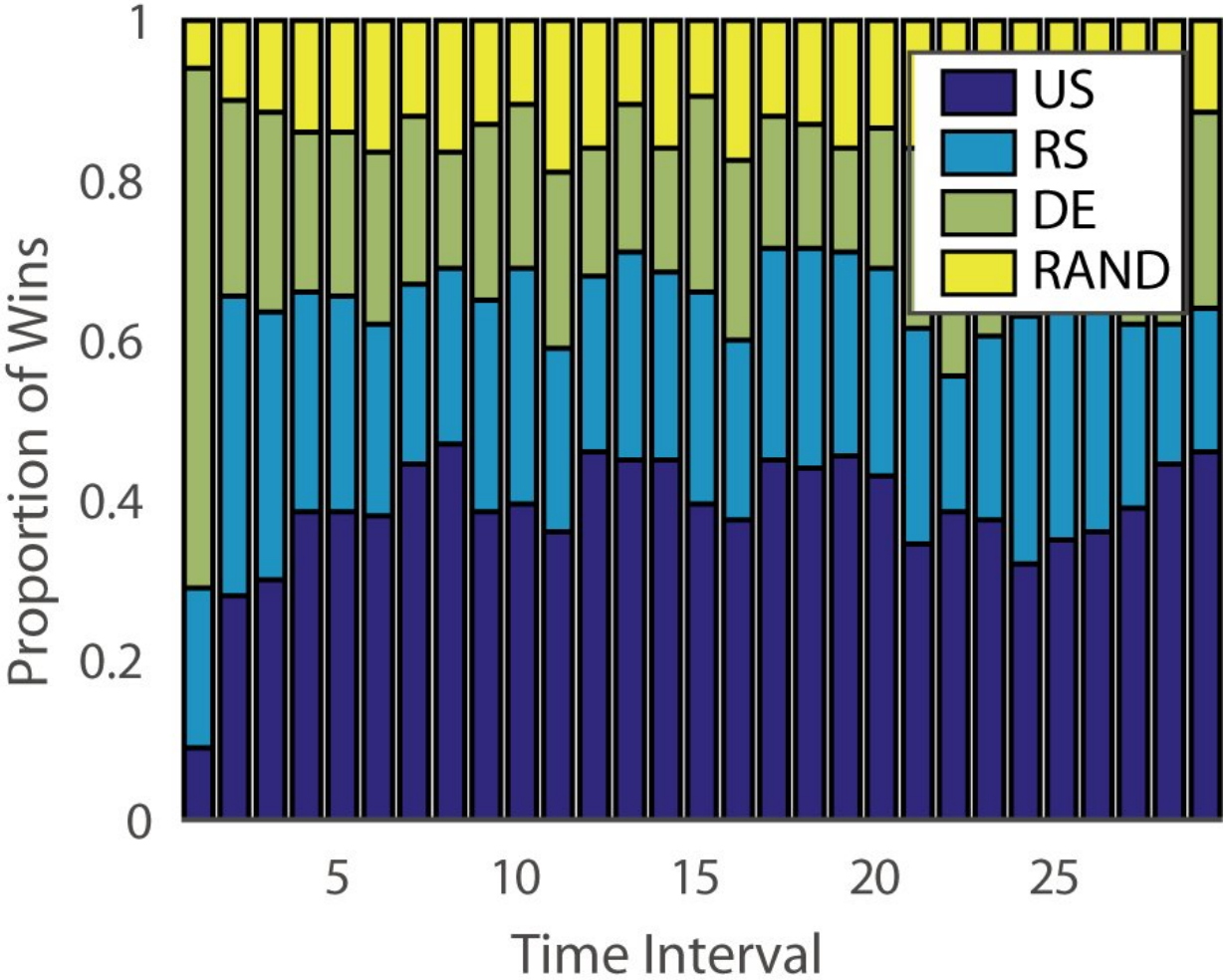# A new dynamic ensemble active learning method based on a non-stationary bandit

October 30 2018, by Ingrid Fadelli

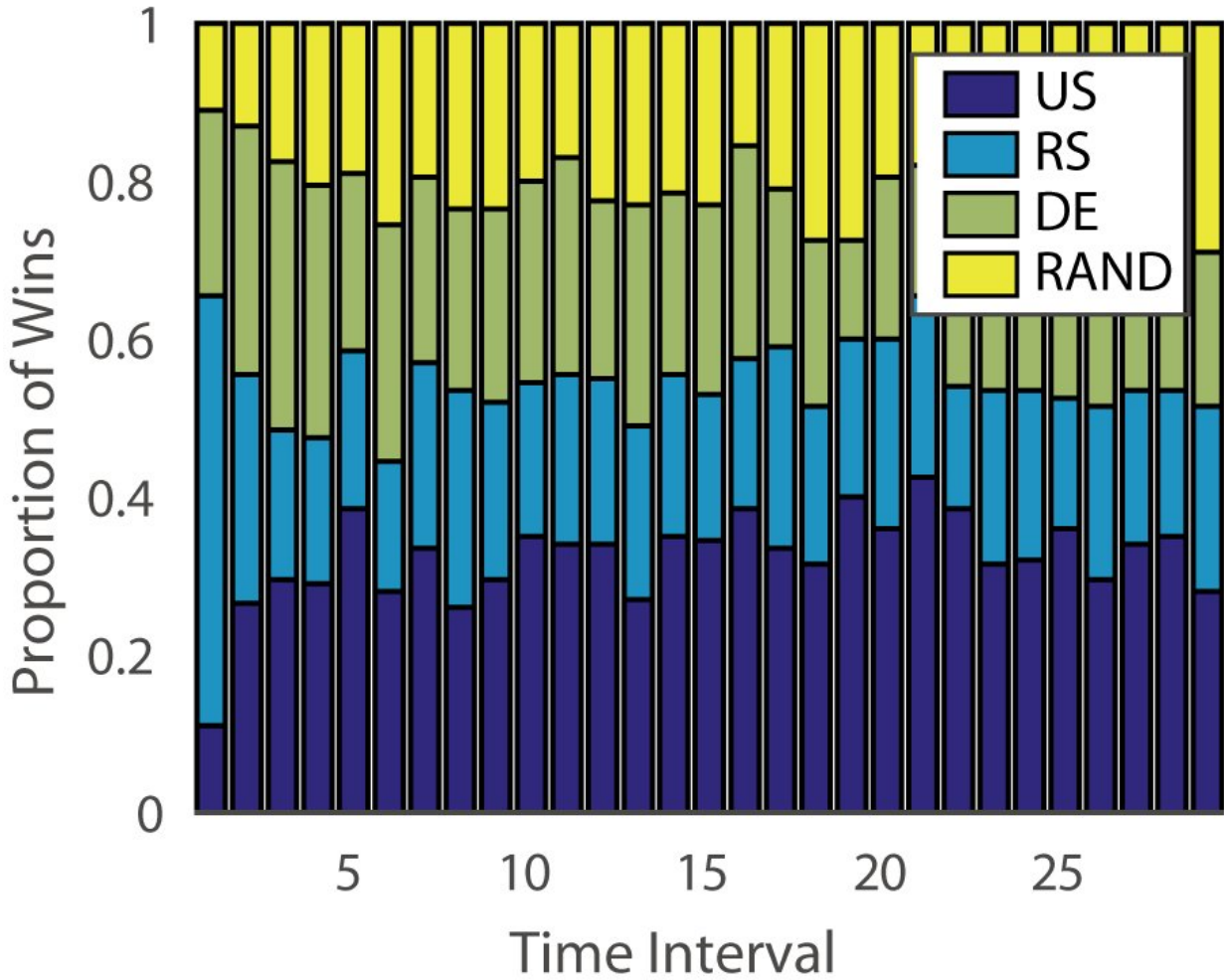Proportion of Wins: "ILPD". Credit: Pang et al.

Researchers at the University of Edinburgh, University College London (UCL) and Nara Institute of Science and Technology have developed a new ensemble active learning approach based on a non-stationary multi-armed bandit and an expert advice algorithm. Their method, [presented in a paper pre-published on arXiv](#), could reduce the time and effort invested in the manual annotation of data.

"Conventional supervised [machine learning](#) is data-hungry, and labelled data can be a bottleneck when data annotation is expensive," Timothy Hospedales, one of the researchers who carried out the study told Tech Xplore. "Active learning supports supervised learning by predicting the most informative data points to annotate so that good models can be trained with a reduced annotation budget."

Active learning is a particular area of machine learning in which a learning [algorithm](#) can actively choose the data it wants to learn from. This typically results in better performance, with significantly smaller training datasets.

Researchers have developed a variety of active learning algorithms that could reduce the costs of annotation, but so far, none of these solutions has proved to be effective for all problems. Other studies have hence used bandit algorithms to identify the best active learning algorithm for a given dataset.

"The term 'bandit' refers to a multi-armed bandit slot machine, which is a convenient mathematical abstraction for exploration/exploitation problems," Hospedales explained. "A bandit algorithm finds a good balance between effort spent on exploring all slot machines to find out which is paying out most, with effort spent on exploiting the best slot machine found so far."

Proportion of Wins: "german". Credit: Pang et al.

The efficacy of active learning algorithms varies both across problems and over time at different stages of learning. This observation is analogous to playing slot machines, where payout probability changes over time.

"The aim of our study was to develop a new bandit algorithm that improves performance by accounting for this aspect of the active learning problem," Hospedales said.

To tackle this limitation, the researchers proposed a dynamic ensemble active learner (DEAL) based on a non-stationary bandit. This learner builds up an estimate of each active learning algorithm's efficacy online, based on the reward (importance-weighted accuracy) obtained after every annotation of data.

"It does this by using the preference expressed for that point by each active learning algorithm," Kunkun Pang, another researcher who carried out the study, told Tech Xplore. "To deal with the issue of the changing efficacy of active learners over time, we periodically restart the learning algorithm to refresh its active learner preference. With this capability, if the most effective active learning algorithm changes between early and late stages of learning, we can quickly adapt to this change."

In Active Learning, $\begin{cases} \text{Arm = Data Point} \\ \text{Expert = Criterion} \end{cases}$

Advices$=\xi_k^n$, where $n = \{1, ..., N\}$

Arms$= \{1, 2, ..., K\}$
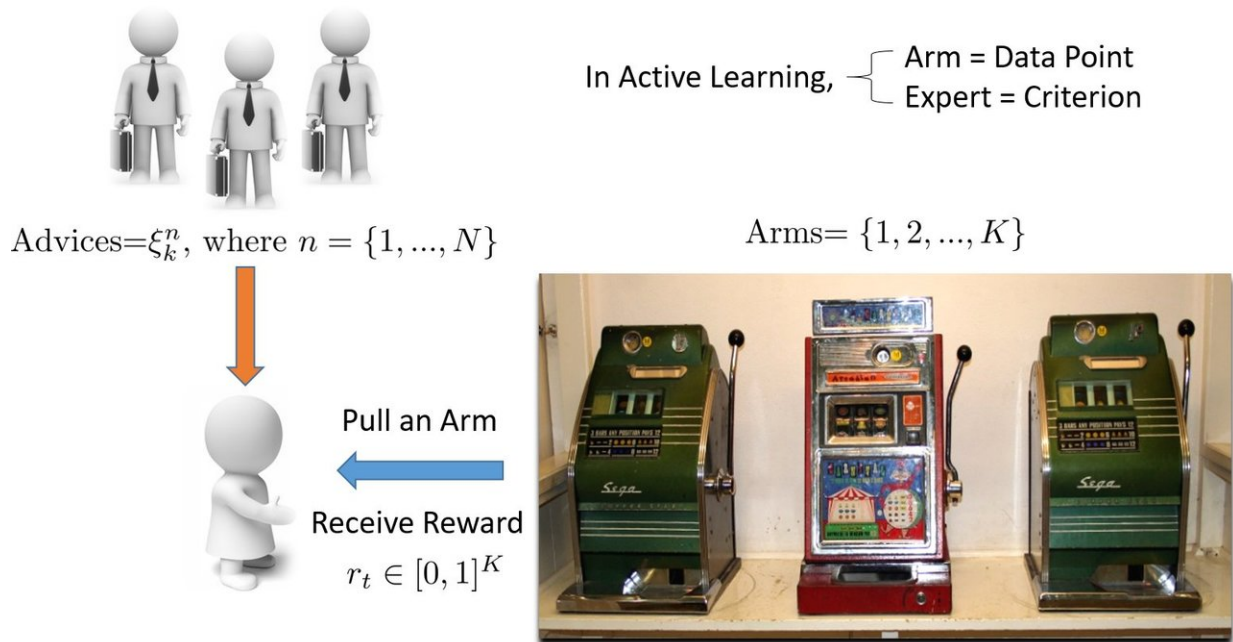
Pull an Arm

Receive Reward
$r_t \in [0,1]^K$

Illustration of multi-armed bandit based active learning approach. Credit: Pang et al.

The researchers tested their approach on 13 popular datasets, achieving highly encouraging results. Their DEAL algorithm has a mathematical performance guarantee, meaning that it there is a high degree of confidence in how well it will work.

"The guarantee relates the performance of our algorithm, which is that of an ideal oracle that always knows the right choice for the active learner," Hospedales explained. "It provides a bound on the performance gap between such a best-case algorithm and ours."

The empirical evaluation carried out by Hospedales and his colleagues confirmed that their DEAL algorithm improves active learning performance on a suite of benchmarks. It does this by continuously identifying the most effective active learning algorithm for different tasks and at different stages of training.

"Today, while active learning is appealing, its impact on machine learning practices is limited due to the hassle of matching algorithms to problems and to stages of learning," Hospedales said. "DEAL eliminates this difficulty and provides an approach to tackle many problems and all stages of learning. By making active learning easier to use, we hope it can have a bigger impact on reducing annotation cost in machine learning practice."
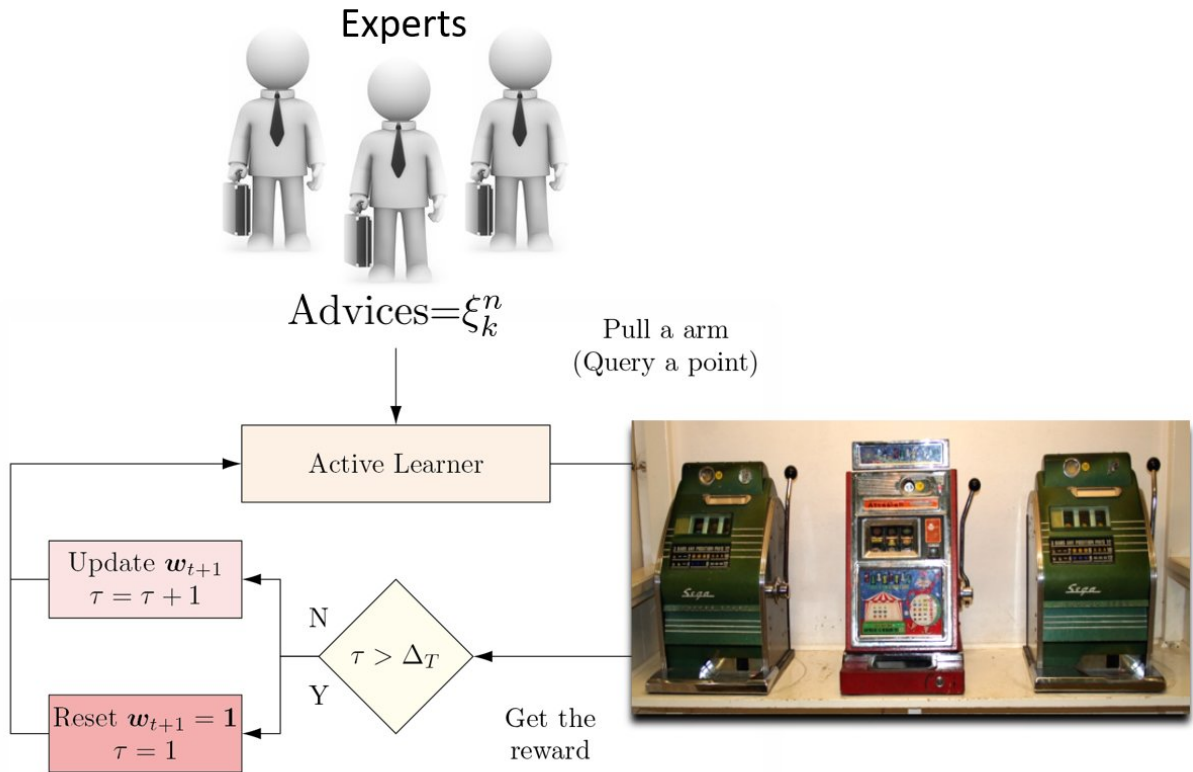
Illustration of DEAL REXP4 algorithm. Credit: Pang et al.

Despite the very promising results, the technique devised by the researchers still has a significant limitation. DEAL does all the learning within a single problem and this results in a 'cold start,' meaning that the algorithm approaches all new problems with a blank slate.

"In ongoing work, we are learning how to annotate on many different problems and eventually transfer this knowledge to a new problem, in order to perform effective annotation immediately with no warm-up requirements," Pang said. "Our preliminary work on this topic has been published and also won the Best Paper prize at ICML 2018 AutoML workshop."

More information: Dynamic ensemble active learning: A non-stationary bandit with expert advice. arXiv: 1810.07778 [cs.LG]. arxiv.org/abs/1810.07778

Meta-Learning Transferable Active Learning Policies by Deep Reinforcement Learning. arxiv:1806.04798 [cs.LG] arxiv.org/abs/1806.04798