Improving Collective Decision-Making Using Confidence and Value Estimates

AXEL ABELS and TOM LENAERTS, Université Libre de Bruxelles and Vrije Universiteit Brussel, Brussels, Belgium
VITO TRIANNI, Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy
ANN NOWÉ, Vrije Universiteit Brussel, Brussels, Belgium

1. INTRODUCTION

In this abstract, we propose two improvements to state-of-the-art collective decision making algorithms in which the learner queries a set of experts for their advice about alternative solutions [Abels et al. 2020]. Whereas the first improvement focuses on the use of quality estimates as an alternative to the classical probability distributions over the alternatives [Auer et al. 2002; McMahan and Streeter 2009; Beygelzimer et al. 2010], the second improvement is to exploit the confidence of experts in order to find the best collective decision [Marshall et al. 2017]. To evaluate the benefits of these additions, we first consider the performance of a classical confidence-weighted majority vote applied to solving contextual multi-armed bandit problems (CMAB). Building on the observation that such a vote is sensitive to the confidence accuracy, we propose an adaptive algorithm, EXP4.P+CON that expands the EXP4.P [Beygelzimer et al. 2010] algorithm with the two previously mentioned improvements. By integrating confidence as priors we obtain a method that is more robust to imperfect confidence estimates.

1.1 Deciding with Expert Advice for contextual multi-armed bandit problems

When *deciding with expert advice* in CMAB [Li et al. 2010; Zhou 2015], a set of experts is queried for their advice about which arm to pull among K possible choices given the contextual information $\vec{x}_{k,t}$ associated with each arm k. Each expert provides advice in the form of a probability distribution over the possible arms. The learner's goal in this problem is no longer to approximate the problem's mapping but rather to determine the best combination of expert advice. State-of-the-art approaches based on EXP4 [Auer et al. 2002], such as EXP4.P [Beygelzimer et al. 2010], iteratively update the weights of experts to compute a weighted average of the received advice. By maintaining context-independent weights, EXP4.P implicitly makes the assumption that expert performance is homogeneous over the context space. This is a limiting assumption, as we argue in the following section.

1.2 Confidence

In general, it is safe to assume experts will perform relatively well in settings for which they have prior experience, which has been shown to improve performance on visual perception and value estimation tasks [Bang et al. 2017; Marshall et al. 2017]. The dominant approach in these tasks is to use confidence-weighted majority votes (WMV). The collective performance can be improved by increasing the importance of experts that express high confidence. If the confidence $c^n_{k,t}$ of an expert n at time t about context $\vec{x}_k$ is expressed in the range $[0, 1]$ (confidences of $1, 0.5$, and 0 correspond respectively to a perfect expert, a random expert, and the worst possible expert) we can follow a WMV by pulling the arm with the highest weighted sum [Marshall et al. 2017]: $\sum_{n \in N} ln(c^n_{k,t}/(1 - c^n_{k,t}))\xi^n_{k,t}$. It should nevertheless be clear that this method heavily relies on accurate confidence estimates.

2. METHODS

In what follows we address the issue of confidence accuracy, introducing an expansion of EXP4.P, i.e. EXP4.P+CON, which works well with both accurate and inaccurate estimates. We also discuss how value advice is introduced to avoid the problem of assigning probabilities to unknown arms.

First, when considering localized expertise in a CMAB problem, experts can be knowledgeable about a subset of the active contexts but agnostic about the remaining contexts. To provide probability advice, experts must make assumptions about unknown arms, which deteriorates the probability distribution assigned over all the arms. In contrast, if advice consists of one value estimate per arm, the uncertainty about some contexts does not affect the given advice for the known arms. Concretely, if $\tilde{f}^n_t$ is expert n's approximation of f at time t, then her value advice for context vector $\vec{x}^n_{k,t}$ at time t is $\xi^n_{k,t} = \tilde{f}^n_t(\vec{x}^n_{k,t})$.

Second, EXP4.P computes the following unbiased gain for each expert which is used to increment expert weights: $\hat{y}^n_t = \xi^n_{k,t} r_t / p_{k,t}$, with $p_{k,t}$ the probability of pulling arm k at time t and $\xi^n_{k,t}$ expert n's probability advice for arm k at time t. When dealing with value advice we hypothesize that an expert with low prediction errors will have low regret and use as gain the negation of (unbiased) squared error between the expert's predicted value and the outcome: $\hat{y}^n_t = -(\xi^n_{k,t} - r_t)^2 / p_{k,t}$. This value iteratively increases the relative weight of the experts with the lowest mean square error. If a confidence estimate $c^n_{k,t}$ is available, we can use it as a prior on the learned weights by modifying EXP4.P to integrate at each time-step the confidence estimate in the aggregation rule (the denominator ensures the weights add up to 1):

$$\sum_{n \in N} \frac{exp(w_{n,t}) c^n_{k,t}/(1 - c^n_{k,t})}{\sum_{n' \in N} exp(w_{n',t}) c^{n'}_{k,t}/(1 - c^{n'}_{k,t})} \xi^n_{k,t}$$

2.1 Experimental Settings

To allow us to exhaustively test our methods we use a pool of N artificial KernelUCB [Valko et al. 2013] experts which solve an artificial CMAB. We consider a context space of $[0,1]^d$ with $d = 2$. The value landscape is generated following Perlin noise [Lagae et al. 2010]. Values generated in this manner have an average reward of 0.5 and range from 0 to 1. When pulling an arm with context $\vec{x}$ in this space, the reward is sampled from a binomial distribution with probability of success $p(r = 1; \vec{x}) = f(\vec{x})$, where $f : [0,1]^d \to [0,1]$ is the function mapping the context to its value in the value landscape. We simulate prior knowledge by introducing each expert to 100 experiences covering approximately 25% of the context space. To evaluate the effect of imperfect confidence, we introduce noise in an expert's reported confidence. Concretely, if a^n_T is expert n's true confidence, her reported confidence is sampled as follows: $c^n \sim \beta(1 + a^n_T/\eta, 1 + (1 - a^n_T)/\eta)$, with η the noise level.

3. RESULTS AND DISCUSSION

Due to space limitations, the performance for the two extreme cases of many arms with few experts, and few arms with many experts are shown here. Results are given in terms of the distance between the best expert and the CDM's performance, i.e., regret in relation to the best expert. A negative regret indicates that the collective results are better than those produced by the best expert.

The top row of Figure 1 compares the performance of the different algorithms when no confidence is provided. It appears that EXP4.P with value advice performs better than probability advice when $K << N$, a tendency which inverses when this condition is no longer met. When the number of arms is high but the number of experts is low, it is easier for an expert's overestimation of a sub-optimal

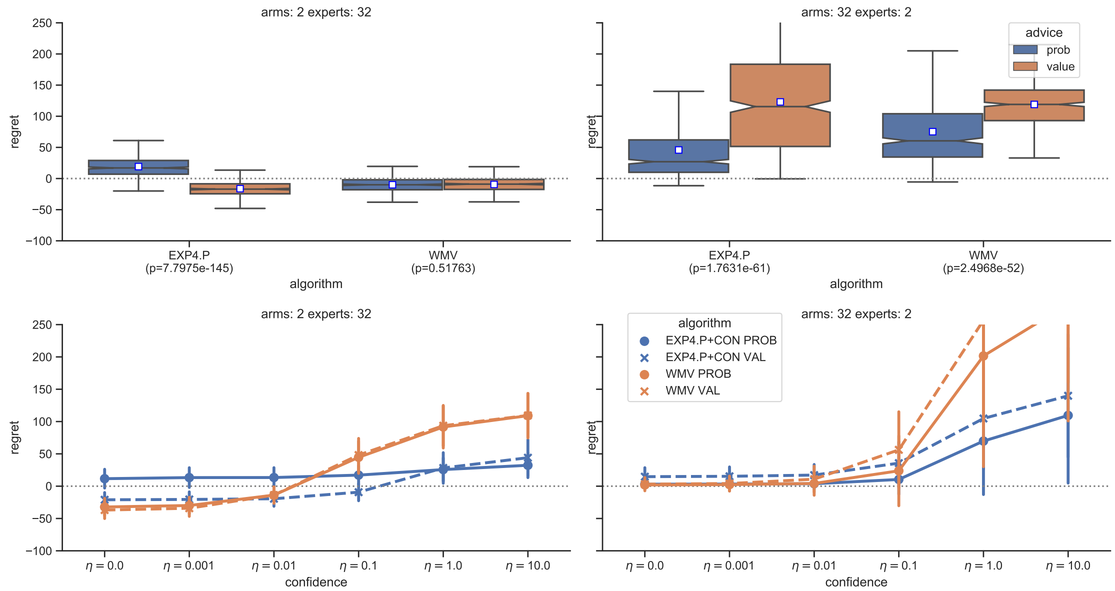

Fig. 1. Performance per advice type and confidence noise. Regret of different aggregation algorithms in function of advice and confidence noise. A value of 0 means the algorithm performs as well as the best expert. This plot presents performance when experts outnumber arms (left) and arms outnumber experts (right). **(top) Performance without confidence grouped by advice type.** The given p-value results from a Wilcoxon test on the results for probability and value advice. **(bottom) Influence of noise on algorithm performance.** For each noise level (η) confidence is sampled from the beta distribution $\beta(1+a/\eta, 1+(1-a)/\eta)$ wherein a is an expert's true confidence. Dashed lines use value advice, full lines use probability advice.

arm to affect the final decision. In contrast, when the number of experts outnumbers the number of arms, the collective variance is reduced (similarly to ensemble methods [Ueda and Nakano 1996]). It is notable that EXP4.P's improvement on a non-adaptive weighted majority vote is not as large as one might expect. In part, this is due to the absence of worse than random experts, which is one of the well known conditions for effective majority votes [Grofman et al. 1983].

A first crucial observation is that when perfect confidence estimates are provided (Figure 1, bottom, $\eta = 0$), the WMV outperforms EXP4.P+CON for both advice types when the number of experts is large. Because EXP4.P+CON is adaptive and we are dealing with a stochastic problem there is a possibility for it to diverge from the ideal weights provided by perfect confidence. As the noise level increases, the performance of the non-adaptive WMV degrades more rapidly than the performance of EXP4.P+CON. This supports our hypothesis that EXP4.P+CON is more robust to noisy confidence. It should however be noted that if confidence is expected to be very noisy, one should simply ignore the provided confidence to fall back on the performance without confidence.

In this paper we explore the inclusion of expert confidence for deciding with expert advice. The confidence-weighted majority vote offers a straightforward way of integrating such confidence estimates. However, our results confirm that such a WMV is sensitive to noisy confidence estimates. To alleviate this, we proposed EXP4.P+CON, which builds on the state-of-the-art EXP4.P algorithm for deciding with expert advice by using confidence estimates as priors on the expert weights. In doing so we obtain an algorithm which benefits from accurate confidence estimates and is more robust to imperfect estimates than the WMV. Confidence with high noise remains a problem however, suggesting that a method which purposefully identifies when confidence is noisy might provide further improvements.

REFERENCES

Axel Abels, Tom Lenaerts, Vito Trianni, and Ann Nowé. 2020. How Expert Confidence Can Improve Collective Decision-Making in Contextual Multi-armed Bandit Problems. In *Submitted to the 12th International Conference on Computational Collective Intelligence (ICCCI 2020)*.

P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. 2002. The Nonstochastic Multiarmed Bandit Problem. *SIAM J. Comput.* 32, 1 (2002), 48–77. DOI:http://dx.doi.org/10.1137/S0097539701398375

Dan Bang, Laurence Aitchison, Rani Moran, Santiago Castanon, Banafsheh Rafiee, Ali Mahmoodi, Jennifer Lau, Peter Latham, Bahador Bahrami, and Christopher Summerfield. 2017. Confidence matching in group decision-making. *Nature Human Behaviour* 1 (05 2017). DOI:http://dx.doi.org/10.1038/s41562-017-0117

Alina Beygelzimer, John Langford, Lihong Li, Lev Reyzin, and Robert E. Schapire. 2010. An Optimal High Probability Algorithm for the Contextual Bandit Problem. *CoRR* abs/1002.4058 (2010). http://arxiv.org/abs/1002.4058

Bernard Grofman, Guillermo Owen, and Scott L. Feld. 1983. Thirteen theorems in search of the truth. *Theory and Decision* 15, 3 (01 Sep 1983), 261–278. DOI:http://dx.doi.org/10.1007/BF00125672

Ares Lagae, Sylvain Lefebvre, Rob Cook, Tony DeRose, George Drettakis, David S Ebert, John P Lewis, Ken Perlin, and Matthias Zwicker. 2010. A survey of procedural noise functions. In *Computer Graphics Forum*, Vol. 29. Wiley Online Library, 2579–2600.

Lihong Li, Wei Chu, John Langford, and Robert E Schapire. 2010. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*. ACM, 661–670.

James A.R. Marshall, Gavin Brown, and Andrew N. Radford. 2017. Individual Confidence-Weighting and Group Decision-Making. *Trends in Ecology & Evolution* 32, 9 (2017), 636 – 645. DOI:http://dx.doi.org/https://doi.org/10.1016/j.tree.2017.06.004

H. Brendan McMahan and Matthew Streeter. 2009. Tighter Bounds for Multi-Armed Bandits with Expert Advice. In *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*.

Naonori Ueda and Ryohei Nakano. 1996. Generalization error of ensemble estimators. In *Proceedings of International Conference on Neural Networks (ICNN'96)*, Vol. 1. IEEE, 90–95.

Michal Valko, Nathaniel Korda, Rémi Munos, Ilias N. Flaounas, and Nello Cristianini. 2013. Finite-Time Analysis of Kernelised Contextual Bandits. *CoRR* abs/1309.6869 (2013). http://arxiv.org/abs/1309.6869

Li Zhou. 2015. A Survey on Contextual Multi-armed Bandits. *CoRR* abs/1508.03326 (2015). http://arxiv.org/abs/1508.03326