Identifying Top-k Players in Cooperative Games via Shapley Bandits

Patrick Kolpaczki[1], Viktor Bengs[1] and Eyke Hüllermeier[2]

[1]Paderborn University, Germany
[2]University of Munich (LMU), Germany

1. The Shapley Value

The notion of a cooperative game, in which players can form coalitions to accomplish a certain task, is a versatile concept with countless practical applications. In the context of (supervised) machine learning, individual features can be seen as players and feature subsets as coalitions — the task here is to train a model with high predictive performance [1, 2]. The connection to explainable AI is established by the question of which proportion of the collective benefit in predictive performance is to be assigned to each individual feature.

Formally, a *cooperative game* is characterized by a pair (N, ν) containing a set of *players* $N = \{p_1, \ldots, p_n\}$ and a *value function* $\nu : \mathcal{P}(N) \to \mathbb{R}$, where $\nu(\emptyset) = 0$ by definition. The players can form *coalitions* $S \subseteq N$ and obtain a combined benefit given by $\nu(S)$ which is called the *worth* of S. For the question of how to distribute the worth $\nu(N)$ of the *grand coalition* N to the individual n many players, the *Shapley value* [3] forms a payoff distribution allocating to each player p_i the value

$$\phi_i = \sum_{S \subseteq N \setminus \{p_i\}} \frac{1}{n\binom{n-1}{|S|}} \cdot (\nu(S \cup \{p_i\}) - \nu(S)) \,.$$

The difference in worth $m(i|S) = \nu(S \cup \{p_i\}) - \nu(S)$ is called p_i's *marginal contribution* given S. The *Shapley value* is the most popular solution concept, as it is the only one satisfying a number of desired requirements for fair payoff distributions [3].

An inherent drawback of the Shapley value is the huge computational effort caused by the exponentially (in the number of players) growing number of marginal contributions to be averaged over. Several approximation methods have been proposed [4, 5] to tackle this difficulty all of them sharing the same idea of calculating mean estimates for randomly sampled marginal contributions uniformly for all players. In many applications the true objective is not to obtain precise Shapley value estimates for all players, but to identify a certain number of k players with the highest Shapley values.

patrick.kolpaczki@upb.de (P. Kolpaczki); viktor.bengs@upb.de (V. Bengs); eyke@ifi.lmu.de (E. Hüllermeier)

2. Shapley Bandits

The *Top-k Shapley* problem is given by a cooperative game (N, ν) in which accesses to the value function ν are costly. Although ν is known (in the sense that we can access $\nu(S)$ for all $S \subseteq N$), the Shapley values remain unknown, since it is practically infeasible for a sufficiently large number of players to compute them. The players in N can be ordered (not necessarily uniquely) such that $\phi_{(1)} \geq \ldots \geq \phi_{(n)}$. For sake of simplicity, we assume that the there are no ties at the top-k-th position. Given a number $k \in [n]$, the learner's performance is measured in the *fixed budget* scenario by its probability to successfully identify the top-k players $p_{(1)}, \ldots, p_{(k)}$ with highest Shapley values, after a given number T of accesses to ν that the learner is allowed to make.

The problem can be reduced to the problem of *multiple arms identification* (MAI), being specified by a set of *arms* $\{a_1, \ldots, a_n\}$ each arm a_i of which is endowed with an unknown *reward distribution* ζ_i having *mean reward* μ_i. The learning process takes place in successive rounds, where in each round t the learner can *pull* an arm a_i of its choice, meaning that it retrieves a random sample $X_i^t \sim \zeta_i$ drawn independently conditioned on the history of the previous rounds. The arms can be ordered (not necessarily uniquely) such that $\mu_{(1)} \geq \ldots \geq \mu_{(n)}$. Given a number $k \in [n]$, the learner's performance is measured in the fixed budget scenario of the MAI problem by its probability to successfully identify the top-k arms $a_{(1)}, \ldots, a_{(k)}$ with highest mean rewards, after a given number T of pulls the learner is allowed to make.

Given a cooperative game (N, ν), the marginal contribution $m(i|S)$ of each player p_i can be viewed as a discrete random variable X_i if S is drawn randomly from $\mathcal{P}(N \setminus \{p_i\})$. Further, by drawing any S with probability $1/n\binom{n-1}{|S|}$, X_i has mean $\mathbb{E}[X_i] = \phi_i$. Thus, by interpreting a player p_i as an arm a_i within a MAI problem, where retrieving a sample of the arm's distribution corresponds to drawing a (independent) sample of X_i, we obtain that the arm's mean μ_i equals the player's Shapley value ϕ_i. Together with the Shapley values, the corresponding arms' means remain unknown to us. Hence, the objective of identifying the top-k players $p_{(1)}, \ldots, p_{(k)}$ with highest Shapley values is equivalent to the task of finding the corresponding k arms $a_{(1)}, \ldots, a_{(k)}$ having highest means. We denote the resulting bandit problem as *Shapley bandits*. This general reduction scheme allows leveraging any algorithm for multiple arms identification to the Top-k Shapley problem without affecting its internal mechanisms. Finally, it should be emphasized that each pull of an arm a_i involves two accesses to the value function ν, one for $\nu(S)$ and the other for $\nu(S \cup \{p_i\})$.

3. Learning Algorithms

Uniform Random Sampling As an illustrative example of how the approach can be applied we present the *Uniform Random Sampling* algorithm (URS) textually. It is a modification of the *ApproShapley* algorithm in [4], which samples in round-robin fashion a marginal contribution for each player by drawing random coalitions. For each player p_i a mean estimate $\hat{\phi}_i$ of ϕ_i is kept by URS and at termination the k players with highest estimates are returned. Utilizing the techniques presented in [5], we derive performance guarantees for the fixed budget setting depending on the variances or ranges of the marginal contributions of each player assuming that T is a multiple of $2n$.

Theorem 1.
Let $k \in [n]$, as well as $\varepsilon > 0$ with $\varepsilon \leq \phi_{\pi(k)} - \phi_{\pi(k+1)}$. Then, URS identifies the top-k players correctly after T many accesses to ν with probability at least

- $1 - {}^{8n^2\sigma^2}/_{\varepsilon^2 T}$ for $\sigma^2 \geq \mathbb{V}[X_i]$ for all $p_i \in N$;
- $1 - 2n \exp\left(-{}^{\varepsilon^2 T}/_{4nr^2}\right)$ for r being an upper bound for the range of X_i for all $p_i \in N$.

Border Uncertainty Sampling Next, we propose a new algorithm (cf. Algorithm 1) called *Border Uncertainty Sampling* (BUS). In similar fashion to Gap-E [6] a measure of (un-)certainty whether a player p_i belongs to the top-k players or not is at the heart of BUS. However, the gaps Δ_i in its measure are calculated in a slightly different manner, namely as the absolute distance to the average of the k-th and $(k+1)$-th highest mean estimates $\hat{\phi}_{\hat{\pi}(k)}$ and $\hat{\phi}_{\hat{\pi}(k+1)}$. Next, BUS draws a sample for the player p_i that minimizes $\Delta_i \cdot t_i$, i.e., the gap times the number of samples BUS has already drawn for it. The intuition behind this measure of certainty is that for players with larger gap Δ_i we are more certain to tell whether it belongs to the top-k players or not. Likewise, a larger number t_i of samples drawn indicates a higher precision of the estimate $\hat{\phi}_i$. BUS outperforms state-of-the art MAI algorithms Gap-E [6] and SAR [6] on synthetic data. In comparison, it stands out by not demanding any problem instance specific parameters like the reward gaps between arms.

Algorithm 1: Border Uncertainty Sampling

Input: N, ν, k
Initialize: $\hat{\phi}_i \leftarrow 0, t_i \leftarrow 1 \ \forall p_i \in N$
for $i = 1, \ldots, n$ **do**
  $\hat{\phi}_i = m(i|S)$ with $S \subseteq N \setminus \{p_i\}$ drawn with probability ${}^1/_{n\binom{n-1}{|S|}}$
end
for $t = n+1, \ldots$ **do**
  Let $\hat{\pi} : [n] \to [n]$ with $\hat{\phi}_{\hat{\pi}(1)} \geq \ldots \geq \hat{\phi}_{\hat{\pi}(n)}$
  $\hat{\phi}^* \leftarrow \left(\hat{\phi}_{\hat{\pi}(k)} + \hat{\phi}_{\hat{\pi}(k+1)}\right)/2$
  $\Delta_i \leftarrow |\hat{\phi}_i - \hat{\phi}^*| \ \forall p_i \in N$
  $i \leftarrow \arg\min_{j \in [n]} \Delta_j \cdot t_j$
  $t_i \leftarrow t_i + 1$
  $\phi_{i,t_i} = m(i|S)$ with $S \subseteq N \setminus \{p_i\}$ drawn with probability ${}^1/_{n\binom{n-1}{|S|}}$
  $\hat{\phi}_i \leftarrow \left((t_i - 1)\hat{\phi}_i + \phi_{i,t_i}\right)/t_i$
end
Output: $p_{\hat{\pi}(1)}, \ldots, p_{\hat{\pi}(k)}$ for $\hat{\pi} : [n] \to [n]$ with $\hat{\phi}_{\hat{\pi}(1)} \geq \ldots \geq \hat{\phi}_{\hat{\pi}(n)}$

Acknowledgments

This research was supported by the research training group Dataninja (Trustworthy AI for Seamless Problem Solving: Next Generation Intelligence Joins Robust Data Analysis) funded by the German federal state of North Rhine-Westphalia.

References

[1] S. B. Cohen, G. Dror, E. Ruppin, Bandit Algorithms, Neural Computation, 2007.

[2] K. Pfannschmidt, E. Hüllermeier, S. Held, R. Neiger, Evaluating Tests in Medical Diagnosis: Combining Machine Learning with Game-Theoretical Concepts, in: IPMU, 2016.

[3] L. S. Shapley, A value for n-person games, Contributions to the Theory of Games, 1953.

[4] J. Castro, D. Gómez, J. Tejada, Polynomial calculation of the Shapley value based on sampling, Comput. Oper. Res., 2009.

[5] S. Maleki, L. Tran-Thanh, G. Hines, T. Rahwan, A. Rogers, Bounding the Estimation Error of Sampling-based Shapley Value Approximation With/Without Stratifying, CoRR abs/1306.4265, 2013.

[6] S. Bubeck, T. Wang, N. Viswanathan, Multiple Identifications in Multi-Armed Bandits, in: ICML, 2013.