A Representation-Learning Game for Classes of Prediction Tasks

Other Conference Item

Author(s):
Uzan, Neria; Weinberger, Nir

Publication date:
2024-03-06

Permanent link:
https://doi.org/10.3929/ethz-b-000664555

A Representation-Learning Game for Classes of Prediction Tasks

Neria Uzan and Nir Weinberger
The Viterbi Faculty of Electrical and Computer Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 3200004, Israel
neriauzan@gmail.com, nirwein@technion.ac.il

Abstract—**We propose a game-theoretic formulation for learning dimensionality-reducing representations of feature vectors, when a prior knowledge on future prediction tasks is available. We analytically find the value of the game and optimal mixed (randomized) strategies for the case of linear representations, tasks, and the mean squared error loss, and propose an algorithm for general classes of representations, tasks, and loss functions.**

Motivation: Data of unlabeled feature vectors $\{\boldsymbol{x}_i\} \subset \mathcal{X}$ is commonly collected without a knowledge of the *specific* downstream prediction task it will be used for. When a prediction task becomes of interest, responses $\boldsymbol{y}_i \in \mathcal{Y}$ are also collected, and a learning algorithm is trained on $\{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}$. Modern sources, such as high-definition images or genomic sequences, have high dimensionality, and this necessitates to reduce their dimensionality, either for better generalization, for storage/communication savings, or for interpretability. The goal is thus to find a low-dimensional *representation* $\boldsymbol{z} = R(\boldsymbol{x}) \in \mathbb{R}^r$, that preserves the relevant part of the features, for all possible downstream prediction tasks. Unsupervised methods for dimensionality reduction, such as principal component analysis (PCA), kernel PCA and auto-encoders [1], aim that the representation $\boldsymbol{z}$ will maximally preserve the *variation* in $\boldsymbol{x}$, and thus ignore any prior knowledge on future prediction tasks. Following a formulation proposed in [2] for the supervised learning setting, we propose a game-theoretic formulation for the case the downstream task is only known to belong to a given class.

Problem formulation: Assume that the response is drawn according to $\boldsymbol{y} \sim f(\cdot \mid \boldsymbol{x} = x)$, where $f \in \mathcal{F}$ for some known class $\mathcal{F}$. Let $\boldsymbol{z} := R(\boldsymbol{x}) \in \mathbb{R}^r$ be an r-dimensional representation of $\boldsymbol{x}$ where $R\colon \mathcal{X} \to \mathbb{R}^r$ is chosen from a class $\mathcal{R}$ of representation functions, and let $Q\colon \mathcal{X} \to \mathcal{Y}$ be a prediction rule from a class $\mathcal{Q}_{\mathcal{X}}$, with the loss function $\mathsf{loss}\colon \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$. The pointwise *regret* of (R, f) is

$$\mathsf{regret}(R, f \mid P_{\boldsymbol{x}}) := \min_{Q \in \mathcal{Q}_{\mathbb{R}^r}} \mathbb{E}\left[\mathsf{loss}(\boldsymbol{y}, Q(R(\boldsymbol{x})))\right] - \min_{Q \in \mathcal{Q}_{\mathcal{X}}} \mathbb{E}\left[\mathsf{loss}(\boldsymbol{y}, Q(\boldsymbol{x}))\right].$$

The *minimax regret in mixed strategies* is the worst case response function in $\mathcal{F}$ given by

$$\mathsf{regret}_{\mathsf{mix}}(\mathcal{R}, \mathcal{F} \mid P_{\boldsymbol{x}}) := \min_{\mathsf{L}(\boldsymbol{R}) \in \mathcal{P}(\mathcal{R})} \max_{f \in \mathcal{F}} \mathbb{E}\left[\mathsf{regret}(\boldsymbol{R}, f \mid P_{\boldsymbol{x}})\right], \tag{1}$$

where $\mathcal{P}(\mathcal{R})$ is a set of probability measures on the possible set of representations $\mathcal{R}$. The *minimax regret in pure strategies* restricts $\mathcal{P}(\mathcal{R})$ to degenerated measures (deterministic), and so the expectation in (1) is removed. Our main goal is to determine the optimal representation strategy, either in pure $R^* \in \mathcal{R}$ or mixed strategies $\mathsf{L}(\boldsymbol{R}^*) \in \mathcal{P}(\mathcal{R})$.

Theoretical contribution: We address the basic setting in which the representation, the response, and the prediction are all linear functions, under the mean squared error (MSE) loss, and the class is $\mathcal{F}_S = \{\|f\|_S \leq 1\}$ for a known symmetric matrix S. Combined with the covariance matrix of the features, S determines the relevant directions of the function in the feature space, in contrast to just the features variability, as in standard unsupervised learning. We establish the optimal representation and regret in pure strategies, which shows the utility of the prior information, and in mixed strategies, which shows that randomizing the representation yields *strictly lower* regret. We prove that randomizing between merely ℓ^* different representation rules suffices, where $r + 1 \leq \ell^* \leq d$ is a precisely characterized *effective dimension*.

Algorithmic contribution: We develop an algorithm for optimizing mixed representations for general representations/response/predictors and loss functions, based only on their gradients. The algorithm operates incrementally, and at each iteration it finds the response function in $\mathcal{F}$ that is most poorly predicted by the current mixture of representation rules. An additional representation rule is added to the mixture, based on this function and the ones from previous iterations. To optimize the weights of the representation, the algorithm solves a two-player game using the classic multiplicative weights update (MWU) algorithm [3].

Further details: A full version of the paper can be found in [4].

References

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
[2] Y. Dubois, D. Kiela, D. J. Schwab, and R. Vedantam, "Learning optimal representations with the decodable information bottleneck," *Advances in Neural Information Processing Systems*, vol. 33, pp. 18674–18690, 2020.
[3] Y. Freund and R. E. Schapire, "Adaptive game playing using multiplicative weights," *Games and Economic Behavior*, vol. 29, no. 1-2, pp. 79–103, 1999.
[4] N. Uzan and N. Weinberger, "A representation-learning game for classes of prediction tasks," *In preparation*. Available at https://drive.google.com/file/d/15SAcFDRJt6qUzsausG8C3W1L1v-PKpWh/view?usp=drive_link.