

A Model-Based Reinforcement Learning Approach for a Rare Disease Diagnostic Task [†]

Rémi Besson ¹, Erwan Le Pennec ^{2,3} and Stéphanie Allassonnière ¹

¹ Centre de Recherche des Cordeliers, Université de Paris, INSERM, Sorbonne Université, Paris, France

² CMAP Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France

³ XPop, Inria Saclay, Palaiseau, France

[†] Presented at the Entropy 2021: The Scientific Tool of the 21st Century, 5–7 May 2021; Available online: <https://sciforum.net/conference/Entropy2021/>.

Published: 5 May 2021

In this work, we study the problem of inferring a discrete probability distribution using both expert knowledge and empirical data.

This is an important issue for many applications where the scarcity of data prevents a purely empirical approach. In this context, it is common to rely first on an a priori from initial domain knowledge before proceeding to an online data acquisition. We are particularly interested in the intermediate regime, where we do not have enough data to do without the initial a priori of the experts, but enough to correct it if necessary.

We formalize expert knowledge as a set of priors, e.g., on the marginals or on the support of distribution. The expert distribution is defined as the distribution of the maximum entropy that satisfies the constraints set by the experts. In turn, empirical data is used to construct the empirical distribution.

We present a new method for objectively choosing the weight to be given to the experts in relation to the data. We define our estimator as the projection of the experts on the confidence interval centered on the empirical distribution. This is the closest distribution from the experts which is consistent with the observed data. The confidence level is the unique parameter of this method.

We show, both empirically and theoretically, that our proposed estimator is always more efficient than the best of the two models (expert or data alone) within a constant.

Our estimator allows a bad a priori to be abandoned relatively quickly when the inconsistency of the data collected with the initial a priori is observed. At the same time, this same mixture makes it possible to keep the initial a priori if it is good. We prove empirically that our method outperforms a parametric Bayesian approach in such a task.



© 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).