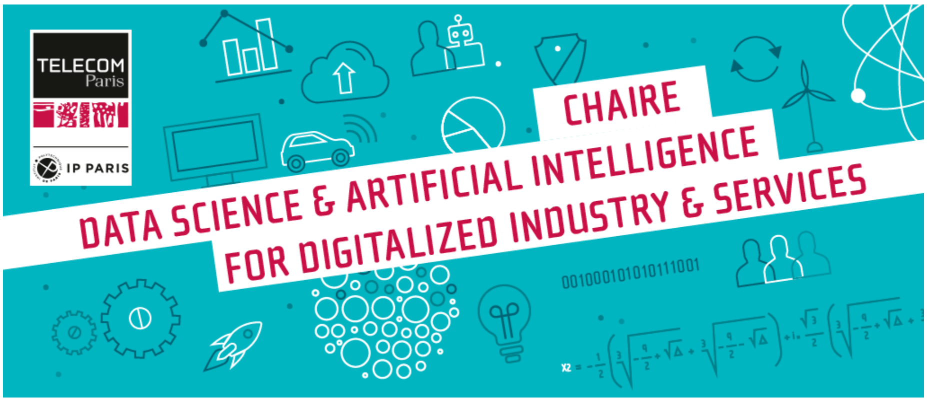

Offre de stage

Sujet : probabilistic optimization of expensive problems: getting the most of few samples

Possibilité de poursuivre sur une thèse

La Chaire Data Science and Artificial Intelligence for Digitized Industry and Services

Portée par Florence d'Alché-Buc, enseignante-chercheur dans le département Image, Données, Signal de Télécom ParisTech, la chaire DSAI réunit cinq partenaires industriels : Airbus Defence & Space, Engie, Idemia, Safran et Valeo Finance. Son objectif général est de développer, en liaison étroite entre les Parties, une formation et une recherche de niveau international.

Ses quatre principaux axes de recherche sont :

1. Analyse et prévision de séries temporelles (Predictive Analytics on Time Series) ;

2. Exploitation de données hétérogènes, massives et partiellement étiquetées (Exploiting Large Scale and Heterogeneous, Partially Labelled Data) ;

3. Apprentissage pour une prise de décision robuste et fiable (Learning for Trusted and Robust Decision) ;

4. Apprentissage dans un environnement dynamique (Learning through Interactions with a Changing Environment).

Description du stage

Encadrement

Ekhine Irurozki (Telecom Paris), Manuel Lopez-Ibanez (Manchester University, Universidad de Malaga)

Lieu et dates du stage

Telecom Paris, 19 Place Marguerite Perey, 91120 Palaiseau

Date de début du stage : 2021/2022

Équipe(s) d'accueil de la thèse

Equipe Signal, Statistique et Apprentissage (S^2A).

Mots clés

learning with few samples, combinatorial optimization,

Sujet détaillé

Background and motivation: In many learning contexts it is assumed that the number of available datasets is large. The best example is deep learning, where the dataset is usually is huge [7]. However, this assumption is unrealistic in many practical applications, where data samples are expensive to get. This internship proposal considers this scenario in which few data points are going to be available and therefore, we need to make the most of the few samples available.

The setting considered in this proposal is online learning, meaning that at each iteration one data point is observed and the probability model of the sample needs to be updated according to this new observation. Moreover, it is an active setting meaning that at each iteration we need to decide which will be the next data point to observe in such a way that the algorithm converge as fast as possible [12].

The data samples have the form of rankings (or permutations) [5]. This data type appears naturally in several contexts such as preference elicitation in recommender systems, matching problems, or routing problems in logistics. The trivial approach when dealing with complex data (such as permutations, time series, or graphs to name a few) is to use standard statistical techniques for multivariate real data and adapt them for the complex data. However, this approach is usually far from optimal, and new, ad-hoc techniques specific for the particular type of complex data offer in general better results. In this internship, we will propose ad-hoc algorithms for permutations and analyze their theoretical and practical performance.

This proposal is divided into several tasks, which will cover short-term goals and long-term goals. The first task will be to implement a learning algorithm for online, active setting on permutation data. This algorithm will extend previous methods [4,14]. An interesting question from a theoretical point of view is to characterize which are the conditions that guarantee convergence of the learning algorithm.

The second task is to study the quality of this algorithm to learn distributions with few samples. We will carry out an experimental evaluation on a scientific computational cluster over benchmark permutation problems and analyze the results. In particular, we will use the algorithm proposed in the previous step to learn a surrogate distribution of different combinatorial optimization problems. This project needs theoretical as well as experimental advances. These advances will follow the path open in previous studies in related domains, which suggests that the proposed learning algorithm converges while minimizes the number of samples. We base our argument on two facts: (1) similar techniques proposed by the advisors have been successful in related contexts and (2) the unweighted version has been shown to converge to the true matching [14].

In a third stage, as a long term goal, will be extended to a fully Bayesian framework [3]. Current Bayesian surrogate-models for permutations rely on Gaussian Processes (GP) applied to a distance metric that converts the permutation space into a continuous (discretized) space. This has several drawbacks related to the ruggedness of combinatorial spaces and the huge amount of computation time required. We aim at replacing the GP model by a model more adequate for permutation spaces and developing appropriate acquisition functions (or infill criteria) that guide the choice of the next observation.

Profil du candidat

Student holding a Master 2 researchM

- Statistical learning, interest in combinatorics
- Good level in programming (Java, C/C++, Python)
- Good level of English

Candidatures

Send to irurozki@telecom-paris.fr

- Curriculum Vitae
- Personalized cover letter explaining the candidate's interest in the subject (directly in the body of the email)
- Transcripts from previous years
- Contact a reference person

Incomplete applications will not be considered.

Références

[1] Martin Zaefferer, J. Stork, M. Friese, A. Fischbach, Boris Naujoks, and Thomas Bartz-Beielstein. Efficient Global Optimization for Combinatorial Problems. In C. Igel and D. V. Arnold, editors, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2014, pages 871–878. ACM Press, New York, NY, 2014.

[2] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical Bayesian Optimization of Machine Learning Algorithms. In P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems (NIPS 25), pages 2960–2968. Curran Associates, Red Hook, NY, 2012.

[2] Alexander I. J. Forrester and Andy J. Keane. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences, 45(1-3):50–79, 2009.

[3] Eriksson, D., Pearce, M., Gardner, J., Turner, R. D., & Poloczek, M. (2019). Scalable global optimization via local bayesian optimization. In Advances in Neural Information Processing Systems (pp. 5496-5507).

[4] Irurozki, E., Calvo, B., & Lozano, J. A. (2019). Mallows and generalized Mallows model for matchings. Bernoulli, 25(2), 1160–1188. https://doi.org/10.3150/17-BEJ1017

[5] Maystre, L., & Grossglauser, M. (2017). Just Sort It! A Simple and Effective Approach to Active Preference Learning. Proceedings of the 34th International Conference on Machine Learning, 70, 2344–2353. http://proceedings.mlr.press/v70/maystre17a.html

[6] Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. D. (2016). Handbook of computational social choice. In Handbook of Computational Social Choice. https://doi.org/10.1017/CBO9781107446984

[7] Bengio, Y., Lodi, A., & Prouvost, A. (2020). Machine learning for combinatorial optimization: A methodological tour d'horizon. European Journal of Operational Research. https://doi.org/10.1016/j.ejor.2020.07.063

[8] Meila, M., & Chen, H. (2016). Bayesian non-parametric clustering of ranking data. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2016.2515599

[9] V. Jensen, F. (2001). Bayesian Networks and Decision Graphs. Springer Verlag.

[10] Asfaw, D., Vitelli, V., Sørensen, Ø., Arjas, E., & Frigessi, A. (2017). Time-varying rankings with the Bayesian Mallows model. Stat, 6(1), 14–30.

[11] Hüllermeier, E., Fürnkranz, J., Cheng, W., & Brinker, K. (2008). Label ranking by learning pairwise preferences. Artificial Intelligence, 172(16–17), 1897–1916.

[12] Lu, T., & Boutilier, C. (2011). Learning Mallows models with pairwise preferences. Proceedings of the 28th International Conference on Machine Learning, ICML 2011.

[13] Vitelli, V., Sørensen, Ø., Crispino, M., Frigessi, A., & Arjas, E. (2018). Probabilistic preference learning with the Mallows rank model. Journal of Machine Learning Research, 18(1).

[14] Irurozki, E., Lopez-Ibanez, M., (2020). Unbalanced Mallows Models for Optimizing Expensive Black-Box Permutation Problems, In proceedings of European Conference on Evolutionary Computation in Combinatorial Optimisation (EvoCOP) (Submitted).