

Cifre PhD Proposal: “Fairness in multi-stakeholder recommendation platforms”

NAVER LABS Europe / LIG (within the MIAI 3IA institute)

Keywords: fairness, transparency, machine learning, recommendation, ranking algorithm

Labs:

- NAVER LABS Europe, Meylan (near Grenoble), France
- Laboratoire d’Informatique de Grenoble (LIG), Grenoble, France (POLARIS/SLIDE teams)

Supervisors:

- Sihem Amer-Yahia (CNRS/LIG) – <http://lig-membres.imag.fr/amery/> – Sihem.Amer-Yahia@imag.fr
- Patrick Loiseau (Inria/LIG) – <http://lig-membres.imag.fr/loiseapa/> – patrick.loiseau@inria.fr
- Jean-Michel Renders (NAVER LABS Europe) – https://europe.naverlabs.com/people_user/Jean-Michel-Renders/ – jean-michel.renders@naverlabs.com

Background and PhD topic description

Recommendation is a prominent machine learning task, used in a variety of platforms ranging from news aggregators to webtoons providers, ad publishers, online dating application, job marketplace, etc. At the heart of recommendation lies a ranking algorithm that ranks contents presented to a user. As recommendation platforms affect users in many important ways, it is crucial to make them *fair*, but what is a fair ranking remains very unclear.

Algorithmic fairness has recently received great attention from the machine learning and data mining communities. A number of mathematical definitions of fairness have been proposed (demographic parity, equal opportunity, etc.) and researchers have proposed various solutions to build learning algorithms that respect those constraints [4, 8, 6, 5]. However, this line of work is currently limited in two directions. First, most of it considers classification whereas very little exists for ranking/recommendation (where it is arguably more complex to define/satisfy fairness). Second, it always considers one-sided fairness notions from the point of view of either *content producers* (e.g., news providers) or *content consumers* (e.g., users) in isolation [1, 7, 2, 3]. Recommendation platforms on the other hand act as mediators between these two actors and need to consider fairness notions from both points of view simultaneously. Naturally, whether a ranking is fair or not depends on the stakeholder’s perspective: intuitively, producers expect fairness in the exposure of their content objects while consumers expect fairness in the variety of items they are exposed to. These (possibly contradictory) objectives raise the crucial question of how to define fairness in multi-stakeholder recommendation settings and how to build algorithms that satisfy the defined notion.

The PhD student will conduct research on fairness in multi-stakeholder recommendation platforms, with three main objectives. *First*, we will empirically study one such platform. We will do that on the example of the news and webtoons recommendation platforms of Naver. We will work in particular on empirically quantifying unfairness. That will help us better understand the multi-stakeholder fairness issue from a data-driven perspective and to formalize the notions for this setting. *Second*, we will work on designing ranking algorithms that provide fair recommendation by design. This will involve theoretical work to prove that the designed algorithm satisfies the fairness properties identified. We will also work on characterizing the

trade-off between the fairness of the different stakeholders. *Finally*, we will test the algorithm in practice and design methods to audit the result so as to prove in practice to a third party that the algorithm respects the fairness properties. That involves in particular questions such as how to measure fairness, which data is needed to show that fairness is respected on a particular run, for how long, etc.

Throughout the PhD, we will work with the case of Naver news recommendation platform in mind for motivation and for algorithms testing and empirical studies, but we will provide solutions that can readily apply to other similar platforms.

Required skills

Candidates should hold (or be about to get) a MSc degree in computer science, applied mathematics, or a related field and have:

- a strong background in mathematics (at least in probability/statistics) and some background in machine learning;
- programming capabilities to perform data-driven empirical studies;
- interest in the societal impact of machine learning and the research area of algorithmic bias (no prior experience working in this area is required).

Additional information and application instructions

The position will be open until filled, **interested candidates are invited to send their application as soon as possible**. The start of the PhD is expected in Fall 2019. Interested candidates are invited to send the following documents to the supervisors mentioned above:

- a detailed CV,
- a list of courses and grades during the MSc (and if possible earlier years),
- a list of 2-3 references willing to support their application,
- a short statement of interest and any other information useful to evaluate the application.

The PhD student will be an employee of Naver during the whole duration of the PhD and receive competitive salary and benefits package. He/she will be register at the MSTII doctoral school of Univ. Grenoble Alpes and be a member of the Laboratoire d’Informatique de Grenoble (LIG). NAVER LABS Europe is the biggest research lab on artificial intelligence in Europe, located in Meylan (10 min from Grenoble) and LIG is one of the biggest lab for computer science in France, located on the campus of Saint Martin d’Hères (10 min from Grenoble).

The PhD student will be a member of the new MIAI institute (one of the four interdisciplinary institute on artificial intelligence in France created by the government in June 2019), as part of the “Explainable and Responsible AI” chair. As such, he/she will benefit from a lively research environment as well as a broad training offer on all aspects of AI.

NAVER LABS Europe is an equal opportunity employer and the MIAI institute is committed to promote the diversity of researchers working in the institute.

Interested candidates are encourage to contact directly the supervisors if they have any question about the position.

References

- [1] Asia J. Biega, Krishna P. Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR)*, pages 405–414, 2018.
- [2] Abhijnan Chakraborty, Gourab K Patro, Niloy Ganguly, Krishna P. Gummadi, and Patrick Loiseau. Equality of Voice: Towards Fair Representation in Crowdsourced Top-K Recommendations. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, January 2019.

- [3] Shady Elbassuoni, Sihem Amer-Yahia, Ahmad Ghizzawi, and Christine El Atie. Exploring fairness of ranking in online job marketplaces. In *The 22nd International Conference on Extending Database Technology (EDBT)*, page to appear, 2019.
- [4] Sara Hajian, Francesco Bonchi, and Carlos Castillo. Algorithmic bias: From discrimination discovery to fairness-aware data mining. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 2125–2126, 2016.
- [5] Moritz Hardt, Eric Price, and Nathan Srebro. Equality of opportunity in supervised learning. In *Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS)*, pages 3323–3331, 2016.
- [6] Michael J. Kearns, Seth Neel, Aaron Roth, and Zhiwei Steven Wu. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pages 2569–2577, 2018.
- [7] Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, pages 2219–2228, 2018.
- [8] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, and Krishna P. Gummadi. Fairness constraints: Mechanisms for fair classification. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 962–970, 2017.