
Automation of a Ranking System by Computing the Correctness of its First Outcome

Malherbe Emmanuel^{*1}, Yves Vanrompay², and Marie-Aude Aufaure¹

¹Mathématiques Appliquées aux Systèmes - EA 4037 (MAS) – Ecole Centrale Paris – Grande Voie des Vignes 92295 Châtenay-Malabry, France

²Transportation Research Institute, Hasselt University (IMOB) – Hasselt University, Wetenschapspark 5 bus 6, 3590 Diepenbeek, Belgium, Belgique

Résumé

Many systems rank outcomes before suggesting them to a user, such as Recommender Systems or Information Retrieval Algorithms. These systems require manual validation, which is time consuming and costly in industrial context. As it is the case in our industrial applications, we assume that the user's needs can be fulfilled by only one relevant outcome. We thus consider an algorithm that systematically selects the top ranked outcome. This approach requires to compute a correctness, estimating the confidence of the automatic decision, or equivalently how likely the first outcome of the ranking system is to be correct. Based on this estimation, we can apply a threshold on the correctness, above which no manual action is required ; the system avoids human validation in many cases.

This paper proposes a novel method to estimate this correctness based on a supervised classification

approach using the manual validations available in the base coupled with a representation of the system's scores. We conducted experiments on Multiposting real-world datasets generated by algorithms used in the industry ; the first algorithm categorizes a job offer, the second recommends semantic equivalents for a given expression in a nomenclature. Our approach has thereby been evaluated and compared, and showed good results on our datasets, even with a limited training base. Moreover, in our experiments, for a given threshold, the better is the correctness estimation, the more performant is the semi-automatic system, showing that the correctness estimation leads thus to crucial efficiency gain.

Mots-Clés: Recommender Systems, Information Retrieval, Recommendation Confidence, Semi, Automatic Classification, E, Recruitment

^{*}Intervenant