

Sujet PLDAC2021

Cotation d'information textuelle

Marie-Jeanne Lesot

La cotation d'information a pour but de quantifier la confiance que l'on peut accorder à une information, en fonction de son contenu, sa source ou son contexte par exemple défini par les autres informations fournies. De multiples critères peuvent être considérés pour capturer les effets de ces composantes, comme la fiabilité, la compétence, la plausibilité ou la crédibilité, ils sont ensuite combinés pour donner une note globale.

L'objectif du projet est d'implémenter des modèles de cotation pour des données de type texte, en particulier des revues sur le web participatif. Il s'agira de définir une architecture modulaire, permettant d'intégrer de multiples critères et de faire varier leur combinaison, et de proposer des définitions pour ces critères pour les données disponibles. Des expérimentations sur des revues examineront la correspondance entre les revues affectées d'une confiance faible et les fausses revues.

Les développements se feront en python et utiliseront `sklearn`.

Quelques références

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on twitter. *Proc. of the 20th Int. Conf. on World Wide Web, WWW'11* (pp. 675–684).

Pichon, F., Labreuche, C., Duqueroie, B., & Delavallade, T. (2014). Multidimensional approach to reliability evaluation of information sources. In P. Capet and T. Delavallade (Eds.), *Information evaluation*, chapter 5, 129–160. Wiley.

Revault d'Allonnes, A. (2014). An architecture for the evolution of trust : Definition and impact of the necessary dimensions of opinion making. In P. Capet and T. Delavallade (Eds.), *Information evaluation*, chapter 9, 261–294. Wiley.