

Extension des mesures textuelles d’informativité à l’évaluation de l’intérêt potentiel d’un passage

Carlos E. González-Gallardo*, Éric SanJuan-Ibekwe*, Juan Manuel Torres Moreno*

*Laboratoire d’Informatique d’Avignon
339 chemin des Meinajaries, 84911 Avignon cedex 9, FRANCE
eric.sanjuan@univ-avignon.fr
<http://lia.univ-avignon.fr/>

1 Introduction

Les mesures d’informativités utilisées pour évaluer le résumé automatique de textes s’appuient pour la plupart sur des mesures de recouvrement entre les n -grammes présents dans le résumé produit automatiquement et ceux apparaissant dans un résumé de référence (Torres-Moreno, 2014) généralement produit par un expert. Ces mesures diffèrent selon :

- la métrique utilisée (cosinus, Dice, Rouge, Kullback-Leibler, Similarité Logarithmique)
- le sac des termes considéré (mots simples, mots n -grammes, entités, pépites, etc.).

Récemment, les approches par plongement lexical de mots offrent une alternative numérique à ces approches discrètes basées sur la présence / absence d’une unité de texte (Ng et Abrecht, 2015).

Ces mesures ont été ensuite étendues à l’évaluation de la recherche ciblée d’information par des requêtes complexes (Bellot et al., 2016). En particulier, dans la tâche INEX de contextualisation de tweets, ce sont des contenus entiers de microblogs qui ont été considérés comme des requêtes.

2 Proposition

Nous définissons formellement l’évaluation de l’informativité sur de courts passages comme une relation ternaire entre un ensemble de sujets T , un sous-ensemble d’extraits courts P provenant d’une grande ressource documentaire et un ensemble S de notes normalisées tel que les passages classés en tête contiennent le plus d’informations les plus pertinentes vis à vis des thèmes explicitement mentionnées dans le passage ou sur des sujets connexes liés de manière implicite.

Nous définissons alors l’extension du concept d’intérêt (Koh et al., 2008) appliqué au texte comme une généralisation de la notion d’informativité où la requête traduisant un besoin d’information de l’utilisateur serait inconnue a priori mais serait susceptible d’être explicité à posteriori.

3 Données

Nous utilisons la collection de données TC@INEX 2012¹. Nous avons extrait l'ensemble des passages des résumés automatiques des participants que les évaluateurs humains ont marqués comme informatifs sur un sujet. Nous avons obtenu un ensemble de 30.000 extraits du Wikipedia évalués par au moins deux annotateurs relativement à 66 sujets complexes présentés sous forme de microblogs.

4 Résultats

Nous présentons le comportement des mesures d'informativité à l'état de l'art sur cette notion d'intérêt pouvant être suscité par un court passage textuel.

Nous observons que dans ce nouveau cadre, les plongements lexicaux ne surpassent les mesures discrètes que sur l'intérêt de mots isolés. Cependant la prise en compte des bi-grammes semble être un point clé de l'évaluation globale de l'intérêt d'un passage. Sur ce point nous constatons que la mesure de similarité logarithmique fournit les meilleurs résultats sur les données de CLEF-INEX 2012.

Références

- Bellot, P., V. Moriceau, J. Mothe, E. SanJuan, et X. Tannier (2016). INEX tweet contextualization task : Evaluation, results and lesson learned. *Inf. Process. Manage.* 52(5), 801–819.
- en
- Koh, Y. S., R. O'Keefe, et N. Rountree (2008). Interestingness Measures for Association Rules : What Do They Really Measure? <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-59904-960-1.ch002>, 36–58.
- Ng, J.-P. et V. Abrecht (2015). Better summarization evaluation with word embeddings for rouge. *arXiv preprint arXiv :1508.06034*.
- Torres-Moreno, J.-M. (2014). *Automatic Text Summarization*. London : John Wiley and Sons.

1. <http://tc.talne.eu>