

Bursting the Bubble with a Needle from the Haystack

Davide Ceolin, CWI, davide.ceolin@cwi.nl

Lesia Tkacz, VU Amsterdam

Robin Sharma, VU Amsterdam

Chantal van Son, VU Amsterdam, c.m.van.son@vu.nl

Keywords: Textual Entailment, Quality Assessment, Filter Bubble, HCI

Introduction - Purpose

Searching for information on the Web can be a challenging task. The Web is a democratic medium, since it allows several points of view to be voiced. However, not all of these voices deserve the same attention, because the quality of the information they present might not meet the user's requirements. For example, the spread of misinformation on the Web is a well-known issue, since anybody can say anything on the Web. This allows varying points-of-view to be launched from the foundations of low-quality information, which makes the task of finding relevant Web documents even harder for the user. Thus, quality assessment of Web information is an important step towards allowing Web users to consciously consume online information.

At the same time, quality assessments are not sufficient to handle the massive amount of information that users can retrieve from the Web. Such information tends also to repeat itself, and users should have the possibility to identify such repetitions. On the one hand, the sheer amount of potentially accessible documents pertaining to any given subject simply overwhelm the user's capacity to read them all. On the other hand, users should be aware of the fact that they might be exposed to a limited number of opinions and perspectives as a result of personalisation mechanisms (filter bubble).

In this abstract, we outline a work in progress that aims at combining the visualization of quality assessment with textual entailment (also known as Natural Language Inference) to overcome such issues and help online users increase their awareness regarding the information they consume online.

Methods

Perspective Detection through Textual Entailment

Textual entailment describes the relationship between two text samples, such as sentences, and indicates whether one hypothesis sentence can be inferred from the evidence sentence (positive entailment), or if they are contradictory (negative entailment). In recent years, research on recognizing textual entailment (RTE) has shifted away from relying on heavily engineered linguistic features and logic, to more recent models which are driven by end-to-end neural networks. Starting with such off-the-shelf LSTM (long short term memory) RTE systems, we leverage the possibilities that textual entailment offers for identifying potential points of agreement or controversy across web documents, thereby indicating the perspectives that their content may express on a given topic. We currently focus on testing the extent to which complimenting or contrasting perspectives can be represented through textual entailment by filtering for topically relevant cross-document sentence pairs within the QuPiD project vaccination corpus. The vaccination corpus is constructed out of a diverse selection of Web documents ranging from high quality sources to otherwise, which express both pro and con views with regards to the highly publicized and controversial vaccination debate.. We theorize that this combination of honing in on the most pivotal text fragments in a document, and further predicting their relationship to others, drastically reduces the amount of information a user must consume to satisfy their goals. At the same time, we predict that our unique approach to prospective identification could enable an indication of the diversity of opinion which may be expressed across a topic.

Quality Assessment

Whereas an informal definition of quality is ‘fit for purpose’, quality assessment of online information is a rather challenging task. The velocity, variety, and volume of online information require information quality assessment tools to address challenging technical issues in order to be efficient and effective to their users. In addition to these challenges shaped by the nature of online information, these tools need to deal with the complexity of quality itself: while ‘fit for purpose’ is a rather generic and apparently simple definition, the fitness function and the purpose are determined by users and the contexts where they operate. Therefore, in principle, quality assessment calls for personalised tools that allow users with different backgrounds and requirements to determine whether a given piece of information meets their requirements. To increase the feasibility of such a task, quality is often times broken down into ‘dimensions’, i.e., aspects that allow evaluating the quality of the information from a given point of view. Such dimensions, by limiting the boundaries of quality assessment, make their automatic estimation more feasible: for example, it is easier to determine whether a given piece of information is accurate, or whether it is of high quality in general. When information is assessed through different dimensions, users have the possibility to determine whether it fits their purpose, by considering those dimensions that are relevant for them. Of course, this all assumes that quality assessments are possible, and that the dimensions evaluated meet the user requirements. Previous work of ours [1,2] shed a light on the assessability of the quality of online documents. The focus of the current abstract, instead, is on how such assessments should be presented to the users. We are helping users to prevent information overload through quality assessments that help them identifying those documents that are more useful to them. But, if quality assessments come into the form of multiple quality dimensions scores per each document (e.g., neutrality, precision), then the user might still suffer from information overload. For this reason, our system adopts the radar graph visualization depicted in Figure 1 in order to provide the user with a glimpse on the quality of the documents he might be interested in, while still breaking down such assessments into multiple dimensions. While relying on an external tool based on human and automated computation of quality assessments [1,2], the focus of this abstract is how to better use the result of such computation.

Model Overview

Figure 1 presents an overview of the model we present in this abstract. First, we use textual entailment as a diversity measure among documents, and we feed such a measure into a clustering algorithm that allows us to discern the various perspectives implied in the documents. Then, we make use of an external quality assessment infrastructure to compute quality assessments of our documents from multiple dimensions, and we present then the clustered and assessed documents to the final users.

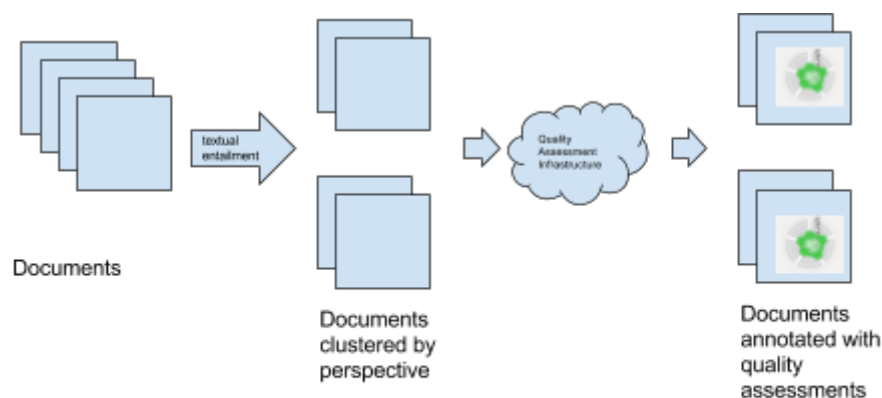


Figure 1: Model Overview

References

- [1] D. Ceolin, J. Noordegraaf, L. Aroyo, Capturing the Ineffable: Collecting, Analysing and Automating Web Document Quality Assessments. In Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2016), pages: 83-97. Springer. 2016.
- [2] D. Ceolin, J. Noordegraaf, L. Aroyo, and C. van Son. Towards web documents quality assessment for digital humanities scholars. In 8th ACM Conference on Web Science. ACM, 2016.