An Interactive Analysis of the Costs of Technologically Assisted Medical Reviews

Giorgio Maria Di Nunzio
Dept. of Information Engineering, University of Padua
Padova, Italy
giorgiomaria.dinunzio@unipd.it

ABSTRACT

The understanding of medical data by experts with highly specific skills, as for example clinicians (a general term that encompasses every medical position involving patients), is necessary for diagnosing illnesses and administering treatment to patients through either medicine or protocols. However, the huge quantity of data produced by digital health technologies make the life of clinicians harder in terms of keeping up with such amount of information. In this demo, we present an interactive application which monitors the amount of relevant literature found by a clinician in a Continuous Active Learning (CAL) framework. The application allows to study the actual costs of completing a systematic review within a 95% confidence interval by alternating random samples of documents with examples selected by a probabilistic machine learning approach.

CCS CONCEPTS

• **Information systems → Clustering and classification**; *Probabilistic retrieval models*; • **Applied computing → Health care information systems**; **Health informatics**;

KEYWORDS

Systematic Reviews, Probabilistic Retrieval Models, Interactive Machine Learning

1 BACKGROUND

A fundamental difficulty with putting the findings of Precision Medicine (PM) into practice is the huge space of treatment options [3]. These options can easily overwhelm clinicians attempting to stay up-to-date with the latest findings and, at the same time, can inhibit a clinician's attempt to determine the best possible treatment for a particular patient [6]. Because PM is also a fast-moving field, keeping up-to-date with the latest literature can be challenging due to both the volume and velocity of scientific advances. According to the Cochrane Handbook for Systematic Reviews of Interventions:[1] "it is unlikely that [healthcare providers, researchers, and policy makers] will have the time, skills and resources to find, appraise and interpret all this evidence and to incorporate it into healthcare decisions." In the last few years, international evaluation campaigns have organized labs in order to study this problem in terms of the evaluation of methods designed to achieve very high recall through controlled simulation [4] and, in particular, for technology assisted reviews in empirical medicine [5].

In this demo, we present an extension of a work which studies a variable threshold approach to tackle the problem of a systematic review task based on the active learning framework [2]. This approach defines a stopping strategy based on the geometry of the two-dimensional space of documents [1] that uses the explicit relevance feedback information given by the expert to automatically estimate the number of documents (medical publications or clinical trials) that need to be read. This approach has demonstrated to outperform all the other systems by reaching a recall greater than 95% with 25,000 documents less than the best performing systems.

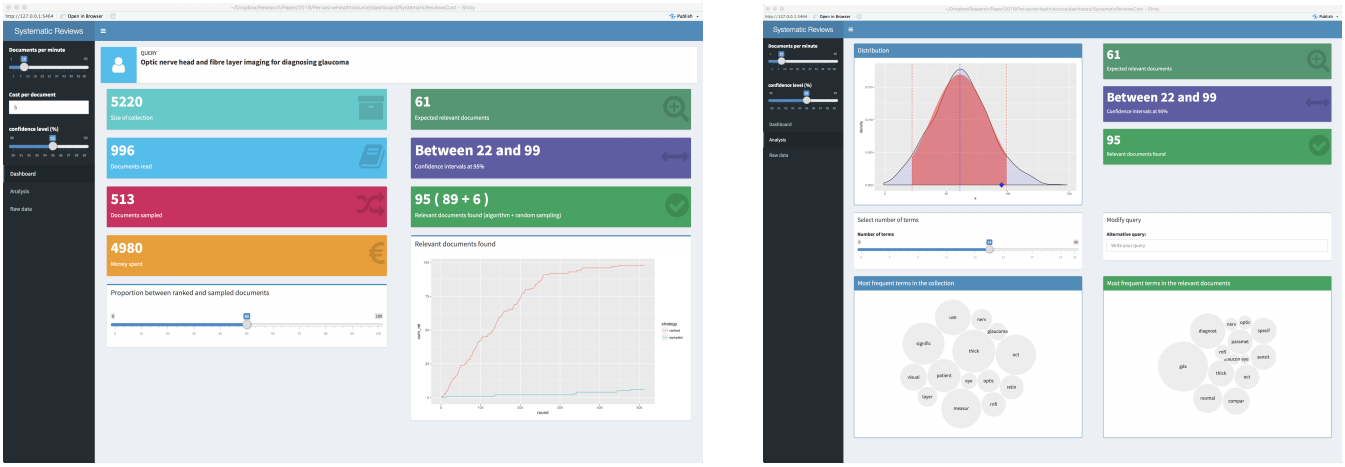
2 DEMO

In this work, we focus on two aspects: i) keeping track of the economical aspects for completing a systematic review, ii) proposing a method for computing the confidence interval for the expected number of relevant documents in the

[1] http://training.cochrane.org/handbook

PervasiveHealth (Demos, Posters, Doctoral Colloquium) 2018, April 20-21

DOI 10.4108/eai.20-4-2018.2276346



(a) Main dahsboard with relevant information found

(b) Probabilistic and text analysis

Figure 1: The two main panels of the demo. On the left, the relevant information found at a particular instant of time; on the right, a detailed analysis on the confidence intervals and the distribution of words in the documents judged by clinicians.

collection by using a mixed strategy for the sampling of documents.

The demo has been implemented using the Shiny Dashboard R package[2] and it is open source on GitHub.[3] The application has two main parts: a sidebar for the main interaction with the system and the body for monitoring the results in real time. We use the data provided by the CLEF 2017 Technologically Assisted Medical Reviews Task [5] to simulate the query (50 in total) and the relevance assessment of the clinicians (for a total of more than 125,000 documents).

The sidebar of the application contains the main inputs: the rate of documents assessed per minute by the clinicians, the cost to assess each single document, the confidence level we want for the expected number of relevant documents.

The body of the application consists of two main panels: the dashboard and the analysis panel (plus a raw data panel that can be used to further analyze the data which is not presented here for space reasons). The dashboard (figure on the left) shows the main pieces of information we need in order to estimate the costs of the systematic review and the amount of relevant information found which is updated every time a clinician reviews a document. The documents are selected according to two strategies: randomly sampled, and selected by the automatic stopping strategy [1]. The system allows to adjust the proportion of documents that are sampled against those that are selected by the automatic system. In fact, we can balance the amount of money we want to spend to estimate the confidence intervals more accurate or get the most relevant information as quick as possible. The second panel (figure on the right) shows a detailed analysis on the probability distribution of the relevant documents in the collection as well as two word clouds that show the most frequent terms found in the whole collection and in the relevant documents. These suggested words can be used to write on the fly alternative queries which can be used by the system to find new relevant documents that were initially missed.

[2]https://rstudio.github.io/shinydashboard/

[3]https://github.com/gmdn/PervasiveHealth2018

REFERENCES

[1] G. M. Di Nunzio. 2018. A Study of an Automatic Stopping Strategy for Technologically Assisted Medical Reviews. In *Proc. of ECIR 2018*. Springer, 672–677. https://doi.org/10.1007/978-3-319-76941-7_61

[2] G. M. Di Nunzio, F. Beghini, F. Vezzani, and G. Henrot. 2017. An Interactive Two-Dimensional Approach to Query Aspects Rewriting in Systematic Reviews. IMS Unipd At CLEF eHealth Task 2. In *Working Notes of CLEF 2017*, Vol. 1866. CEUR-WS. http://ceur-ws.org/Vol-1866/paper_119.pdf

[3] L. J. Frey, E. V. Bernstam, and J. C. Denny. 2016. Precision Medicine Informatics. *J. of the Am. Med. Inform. Assoc.* 23, 4 (2016), 668–670. https://doi.org/10.1093/jamia/ocw053

[4] Maura R. Grossman, Gordon V. Cormack, and Adam Roegiest. 2016. TREC 2016 Total Recall Track Overview. In *In Proc. of TREC 2016*. http://trec.nist.gov/pubs/trec25/papers/Overview-TR.pdf

[5] E. Kanoulas, D. Li, L. Azzopardi, and R. Spijker (Eds.). 2017. *CLEF 2017 Technologically Assisted Reviews in Empirical Medicine Overview. In Working Notes of CLEF 2017. http://ceur-ws.org/Vol-1866/invited_paper_12.pdf*. Number 1866 in CEUR Workshop Proceedings. CEUR-WS.org.

[6] A. O'Mara-Eves, J. Thomas, J. McNaught, M. Miwa, and S. Ananiadou. 2015. Using Text Mining for Study Identification in Systematic Reviews: A Systematic Review of Current Approaches. *Syst. Rev.* 4, 1 (2015), 5. https://doi.org/10.1186/2046-4053-4-5