Computers can assist in retrieving and assessing new information in the medical domain

Oana Frunza (in collaboration with: Diana Inkpen and Stan Matwin)

Evidence-Based Medicine - explicit and judicious use of current best evidence in making decisions about the care of individual patients.

Systematic Review - is a literature review focused on a single topic which tries to identify, appraise, select and synthesize all high-quality research evidence relevant to that topic.
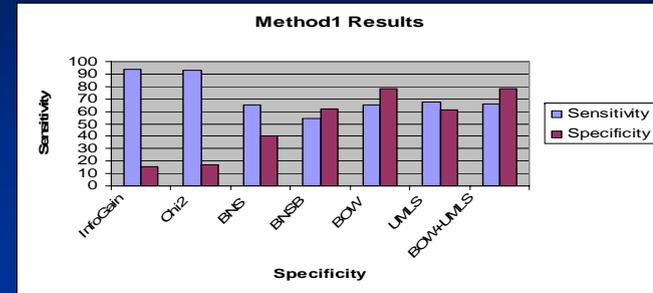
At the very beginning of the Systematic Review, documents / abstracts are classified as relevant or not relevant to the topic of review. This is a task that we are automating by using Machine Learning/ Text Classification.

Topic: dissemination strategy of health care services for elderly people

The purpose of this research was to predict the influence of socioecological factors, including social support, perceived health status, barriers to health promotion behaviors, and selected demographic variables, on the health promotion behaviors of rural women with heart failure (HF) … (Included)

As the number of people who are living long enough to grow old – and then living longer in old age – grows, there has been increasing attention on the roles of physicians, nurses, social workers, and health care administrators in working with older people. … (Excluded)

Method1: a single classifier is trained on instances containing an abstract and its class: Included or Excluded.

Method1 Results

(Bar chart: Sensitivity vs Specificity across methods InfoGain, Chi2, BNS, BNSB, BOW, UMLS, BOW+UMLS; y-axis Sensitivity 0–100, x-axis Specificity)

Method2: trains multiple classifiers, one for each question from the specific protocol of the systematic review.
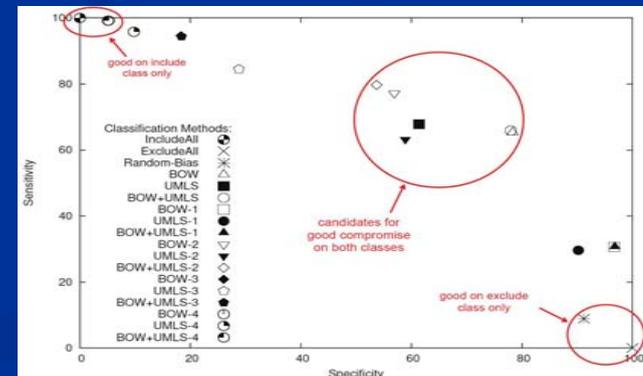
Q1. Is this article about a dissemination strategy or a behavioral intervention?
Q2. Is this a primary study?
Q3. Is the population in this article made of individuals 65-year old or older or does it comprise individuals who serve the elderly population needs (i.e. health care providers, policy makers, organizations, community)?
Q4. Is this a review?

Four voting schemes to identify the relevance of an abstract.

(Scatter plot: Sensitivity vs Specificity with Classification Methods: IncludeAll, ExcludeAll, Random-Bias, BOW, UMLS, BOW+UMLS, BOW-1, UMLS-1, BOW+UMLS-1, BOW-2, UMLS-2, BOW+UMLS-2, BOW-3, UMLS-3, BOW+UMLS-3, BOW-4, UMLS-4, BOW+UMLS-4; annotations: "good on include class only", "candidates for good compromise on both classes", "good on exclude class only")

Final Goal: replace/assist a human reviewer in the workflow of building a systematic review.