

Grading of Evidence

Binod Gyawali (bgyawali@cis.uab.edu), Thamar Solorio (solorio@cis.uab.edu)

University of Alabama at Birmingham
Alabama, USA

UAB THE UNIVERSITY OF ALABAMA AT BIRMINGHAM
Department of Computer and Information Sciences

Introduction

“Evidence Based Medicine (EBM) is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of the individual patient. It means integrating individual clinical expertise with the best available external clinical evidence from systematic research” (Sackett et. al., 1996).

The goal of the automatic grading system is to grade the evidence given the abstracts from which the evidences are extracted.

The purpose of our experiment is to label each evidence automatically based on the **Strength of Recommendation Technology (SORT)** grading scale (Ebell et. Al., 2004). Each evidence is given one of the following three labels.

A: Strong B: Moderate C: Weak

Problem Description

The dataset provided by Australasian Language Technology Association (ALTA) Shared Task 2011 contains three sets of documents. They are:

- Training set
- Development test set
- Test set

Each dataset has a text file containing evidence IDs, a grade for each evidence and one or more abstract IDs for each evidence.

The abstracts are in XML format used by Pubmed containing title and body. They also contain information like publication types, MeSH headings etc.

Our Approach

We cast the problem of evidence grading as a classification task. We apply a supervised learning approach to train our model using the training dataset. Each evidence is represented as a feature vector by using the publication types and MeSH headings as the features.

Document Representation

Documents are represented by publication types and MeSH headings extracted from the abstracts. The steps of features extraction are as follows:

- Extract publication types.
- Extract MeSH headings having majortypicyn=Y.
- Convert publication types and Mesh headings to lowercase.
- Remove special characters like space, ‘:’, ‘,’ ‘.’ ‘”’ by underscore(_).

Experiments

We performed two set of experiments on the data set. The first experiment was done before the submission of the result. After the submission, we performed second experiment with a new approach. They are as follows:

First Experiment

- Each feature was weighted using term frequency inverse document frequency (tf-idf).
- Support Vector Machine (SVM) light multiclass classifier was used to classify the instances.

Second Experiment

- Each feature was weighted using binary value i.e. 0 or 1.
- Five classifiers were used to classify the instances.
 - ❖ Logistic Regression
 - ❖ K Nearest Neighbors
 - ❖ SVMs
 - ❖ Decision Trees
 - ❖ Bayes Net

Experiment Work Flow

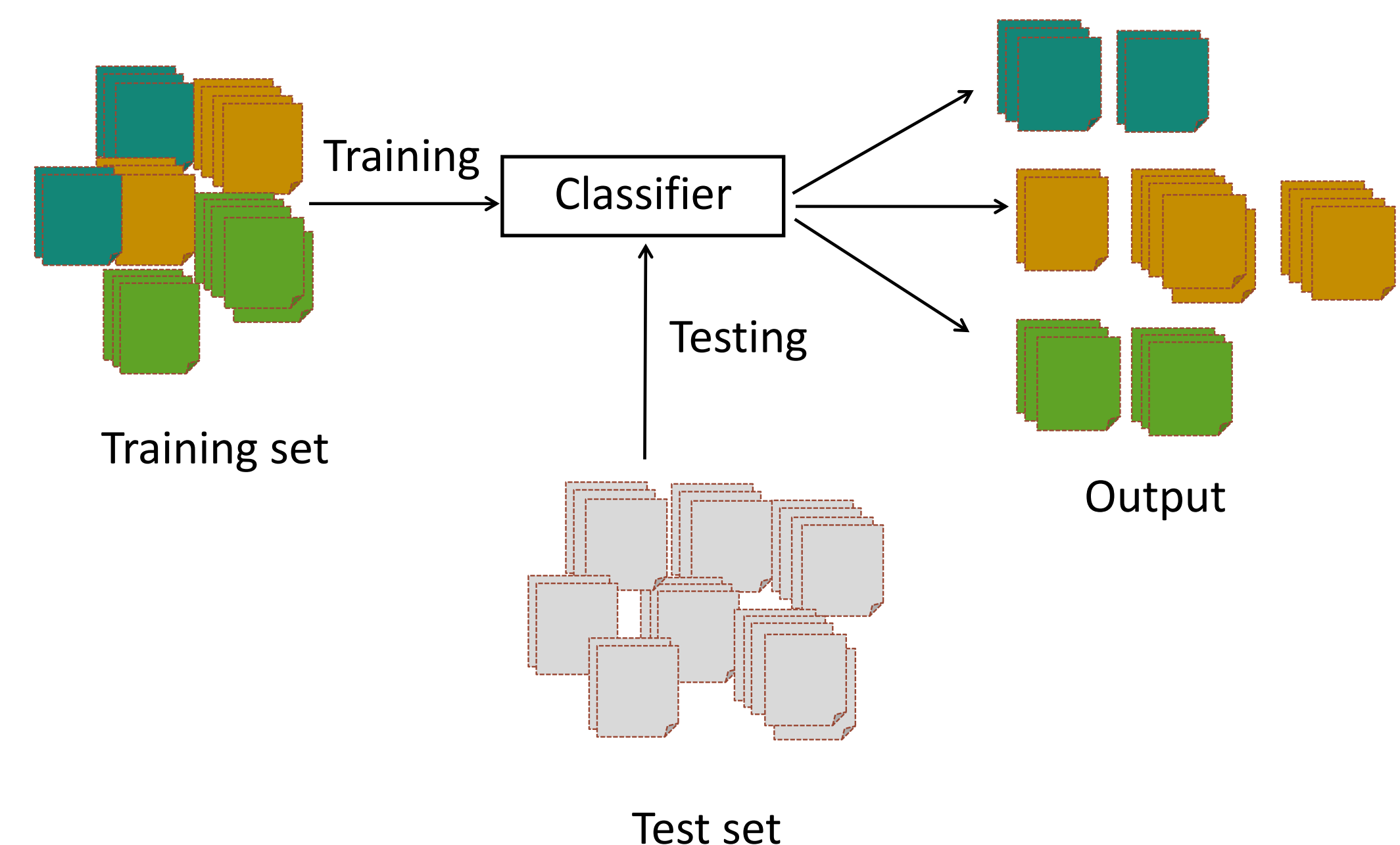


Fig 1: Classifying the evidence

Results

First Experiment

Classifier	Development Test Set	Test Set
SVM Light MultiClass	47.75%	43.71%

Table 1: % Accuracy in the first experiment

Second Experiment

Classifier	Development Test Set	Test Set
Bayes Net	54.49%	45.90%
Logistic Regression	49.44%	44.26%
SVMs	53.93%	52.45%
K Nearest Neighbors	50.00%	38.79%
Decision Trees	57.30%	49.18%

Table 2: % Accuracy in the second experiment

Conclusions

We propose a system of grading evidences using publication types and meSH headings as features. Best result is obtained in second experiment where features are weighted using binary values and SVMs classifier gives the consistent results in the development test set and test set.

The results show that NLP techniques can be applied to practice Evidence Based Medicine (EBM) and application of better approaches can help to achieve a good accuracy.

Future Work

Though we have proposed an approach for grading the evidences, our approach is not competitive enough and doesn't have a high accuracy. Thus, we propose some other techniques to improve the accuracy in our future work.

- Extract new features from the title and body of the abstract.
- Use semantic information.
- Test other classification approaches such as the Profile Based Approach.

References

1. Ebell M. H., Siwek J., Weiss B. D., et al. *Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. Journal of the American Board of Family Practice (JABFP), 2004*
2. Sackett D.L., Rosenberg W. M. C., Gray J. A. M., et al. *Evidence based medicine: what it is and what it isn't. British Medical Journal (BMJ), 1996*

Acknowledgements

We would like to thank ALTA 2011 for providing us an opportunity to participate in the competition.