moving block model developed by SIRTI, which is an industry leader in the design, production, installation and maintenance of railway signalling systems.

ASTRail will run until August 2019 and is coordinated by Riccardo Scopigno from the Istituto Superiore Mario Boella sulle Tecnologie dell'Informazione e delle Telecomunicazioni (ISMB, Italy). Other partners are SIRTI S.p.A. (Italy), Ardanuy Ingeniería S.A. (Spain), École Nationale de l'Aviation Civile (ENAC, France), Union des Industries Ferroviaires Européennes (UNIFE, Belgium) and ISTI-CNR (Italy).

Link:
ASTRail: http://www.astrail.eu/

References:
[1] F. Rispoli, et al.: "Recent Progress in Application of GNSS and Advanced Communications for Railway Signaling", in 23rd Intl. Conf. Radioelektronika, IEEE, 2013, 13-22. DOI: 10.1109/RadioElek.2013.6530882
[2] A. Fantechi, W. Fokkink, and A. Morzenti: "Some Trends in Formal Methods Applications to Railway Signaling", Chapter 4 in Formal Methods for Industrial Critical Systems: A Survey of Applications (S. Gnesi and T. Margaria, eds.). John Wiley & Sons, 2013, 61-84. DOI: 10.1002/9781118459898.ch4
[3] A. Fantechi: "Twenty-Five Years of Formal Methods and Railways: What Next?", in Software Engineering and Formal Methods, LNCS, vol. 8368. Springer, 2013, 167-183. DOI: 10.1007/978-3-319-05032-4_13

Please contact:
Stefania Gnesi, ISTI-CNR, Italy
stefania.gnesi@isti.cnr.it

The Impacts of Low-Quality Training Data on Information Extraction from Clinical Reports

by Diego Marcheggiani (University of Amsterdam) and Fabrizio Sebastiani (CNR)

In a joint effort between the University of Amsterdam and ISTI-CNR, researchers have studied the negative impact that low-quality training data (i.e., training data annotated by non-authoritative assessors) has on information extraction (IE) accuracy.

Information Extraction (IE) is the task of designing software artifacts capable of extracting, from informal and unstructured texts, mentions of particular concepts, such as the names of people, organisations and locations – where the task usually goes by the name of "named entity extraction". Domain-specific concepts, such as drug names, or drug dosages, or complex descriptions of prognoses, are also examples.

Many IE systems are based on supervised learning, i.e., rely on training an information extractor with texts where mentions of the concepts of interest have been manually "labelled" (i.e., annotated via a markup language); in other words, the IE system learns to identify mentions of concepts by analysing what manually identified mentions of the same concepts look like.

It seems intuitive that the quality of the manually assigned labels (i.e., whether the manually identified portions of text are indeed mentions of the concept of interest, and whether their starting points and ending points have been identified precisely) has a direct impact on the accuracy of the extractor that results from the training. In other words, one would expect that learning from inaccurate manual labels will lead

Figure 1: Screenshot displaying a radiological report annotated according to the concepts of interest.
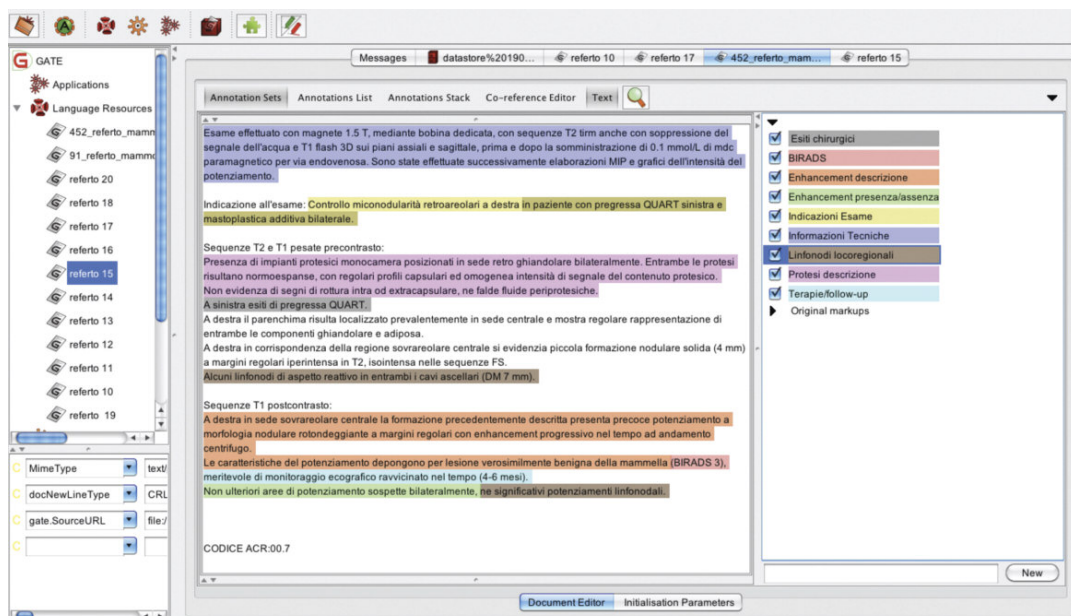
to inaccurate automatic extraction, according to the familiar "garbage in, garbage out" principle.

However, the real extent to which inaccurately labelled training data impact on the accuracy of the resulting IE system, has seldom (if ever) been tested. Knowing how much accuracy we are going to lose by deploying low-quality labels is important, because low-quality labels are a reality in many real-world situations. Labels may be low quality, for example, when the manual labelling has been performed by "turkers" (i.e., annotators recruited via Mechanical Turk or other crowdsourcing platforms), or by junior staff or interns, or when it is old and outdated (so that the training data are no longer representative of the data that the information extractor will receive as input). What these situations share in common is that the training data were manually labelled by one or more "non-authoritative" annotators, i.e., by someone different from the ("authoritative") person who, in an ideal situation (i.e., one where there are no time / cost / availability constraints), would have annotated it.

In this work, the authors perform a systematic study of the extent to which low-quality training data negatively impacts on the accuracy of the resulting information extractors. The study is carried out by testing how accuracy deteriorates when training and test set have been annotated by two different assessors. Naturally enough, the assessor who annotates the test data is taken to be the "authoritative" annotator (since accuracy is tested according to her/his judgment), while the one who annotates the training data is taken to be the "non-authoritative" one. The study is carried out by applying widely used "sequence learning" algorithms (either Conditional Random Fields or Hidden Markov Support Vector Machines) on a doubly annotated dataset (i.e., a dataset in which each document has independently been annotated by the same two assessors) of radiological reports. Such reports, and clinical reports in general, are generated by clinicians during everyday practice, and are thus a challenging type of text, since they tend to be formulated in informal language and are usually fraught with typos, idiosyncratic abbreviations, and other types of deviations from linguistic orthodoxy. The fact that the dataset is doubly annotated allows the systematic comparison between high-quality settings (training set and test set annotated by annotator A) and low-quality settings (training set annotated by annotator NA and test set annotated by annotator A), thereby allowing precise quantification of the difference in accuracy between the two settings.

Link:
http://nmis.isti.cnr.it/sebastiani/Publications/JDIQ2017.pdf

References:
[1] A. Esuli, D. Marcheggiani, F. Sebastiani: "An Enhanced CRFs-based System for Information Extraction from Radiology Reports", Journal of Biomedical Informatics 46, 3 (2013), 425–435.
[2] W. Webber, J. Pickens: "Assessor disagreement and text classifier accuracy", in Proc. of SIGIR 2013,929–932, 2013.

Please contact:
Fabrizio Sebastiani, ISTI-CNR, Italy
+39 050 6212892 fabrizio.sebastiani@isti.cnr.it

4th International Indoor Positioning and Indoor Navigation Competition

The fourth international Indoor Positioning and Indoor Navigation (IPIN) competition, seventh in the EvAAL series, was hosted by the international IPIN conference in Sapporo, Japan, on 17 September 2017.

The aim of the competition is to measure the performance of indoor localisation systems, that are usable in offices, hospitals or other big buildings like warehouses. The 2017 edition has attracted 28 teams and allowed participants to test their localisation solutions with rigorous procedures inside the two-floor structure of the Conference Hall of Hokkaido University.
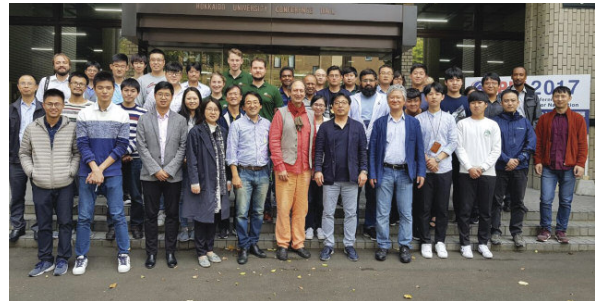
IPIN participants.

The competition ended with the awarding of four 150.000¥ (1.100€) prizes:
- smartphone-based (Chan Gook Park, Seoul National University, Corea)
- dead reckoning (Chuanhua Lu, Kyushu University, Japan)
- offline smartphone-based (Adriano Moreira, University of Minho, Portugal)
- offline PDR warehouse picking (Yoshihiro Ito, KDDI R&D Laboratories Inc., Japan).

Prizes were awarded by the official sponsors of the competition KICS , ETRI, TOPCON, and PDR Benchmark, respectively.

Francesco Potortì, Antonino Crivello and Filippo Palumbo, researchers at the WNLab group of CNR-ISTI at Pisa, were among the main organisers.

More information on the international competition, including complete results, photos and comments at
http://evaal.aaloa.org