Large-scale Machine Learning for Healthcare

Paola Lapadula[1], Giansalvatore Mecca[1], Donatello Santoro[1], Luisa Solimando[2], and Enzo Veltri[1]

[1] University of Basilicata, Potenza, Italy
[2] Svelto! Big Data Cleaning and Analytics, Potenza, Italy

1 Introduction

Greg, ML [4] is a machine learning system that generates automatic diagnostic suggestions based on patient profiles. In summary, Greg takes as input a digital profile of a patient, and suggests one or more diagnosis that, according to its internal learned models, fit the profile with a given probability. We assume that a doctor inspects these diagnostic suggestions, and takes informed actions about the patients and the related medical treatments.

We do not reinvent the wheel, the idea of using machine learning for the purpose of examining medical data is not new [3, 6, 5]. In fact, several efforts that have been taken in this direction [1, 2]. However, to the best of our knowledge, all of the existing tools concentrate on rather specific learning tasks, for example identifying a single pathology – like heart disease [9, 7], or pneumonia [8], or cancer, where results of remarkable quality have been reported [10]. On the other hand, Greg has the distinguishing feature of being a broad-scope diagnostic-suggestion tool. In fact, at the core of the tool stands a multi-label learning model that allows to suggest large numbers of pathologies, currently about fifty, and in perspective several hundreds.

Greg is a research project developed by Svelto!, a spin-off of the data-management group at University of Basilicata.

2 Architecture of Greg

The architecture and the overall flow of Greg is depicted in Figure 1.

At the core of Greg there is a classifier for patient profiles that provides doctors with diagnostic suggestions. Patients profiles are entirely anonymous, i.e., Greg does not store nor requires any sensitive data, and are composed of three main blocks:

- anonymous biographical data, mainly *age* and *gender*, and medical history of the patient, i.e., past *medical events*, the past and the current *medical therapy* and *pathologies*, especially the chronic ones;
- result of lab exams, in numerical format;
- textual reports from instrumental exams, like RX, ultrasounds etc.

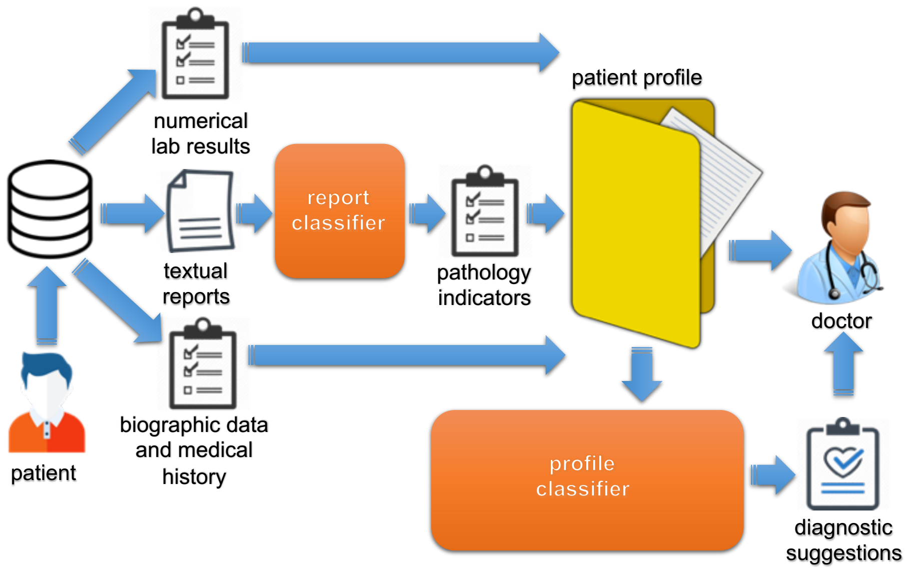

Fig. 1: Architecture of Greg.

These items compose what we called the patient profile that is fed to the profile classifier in order to propose diagnostic suggestions to doctors. Notice that, while biographic data, medical history and lab exam results are essentially structured data, and therefore can be easily integrated into the profile, reports of instrumental exams are essentially text, and consequently, they are in the form of unstructured data. As a consequence, Greg relies on a second learning module to extract what we call *pathology indicators*, i.e., structured labels indicating anomalies in the report that may suggest the presence of a pathology.

Actually, the report classifier is a natural-language processing module. It identifies pathology indicators from the text of the report in natural language. Then, the pathology indicators can be integrated within the patient profile.

The crucial module for the generation of the patient profile is the report classifier. In fact, reports of instrumental exams often carry crucial information for the goal of identifying the correct diagnostic suggestions. At the same time, their treatment is language-dependent, and learning is labor-intensive step, since it requires to manually label a large set of reports in order to train the classifier.

Once the profile for a new patient has been generated, it is fed to the profile classifier that outputs diagnostic suggestions to the physician. There are a few important aspects to be observed here.

– First, Greg is trained to predict only a finite set of diagnoses. This means that it is intended mainly as a tool to gain positive evidence about pathologies that might be present, rather than as a tool to exclude pathologies that are not present. In other terms, the fact that Greg does not predict a specific diagnosis does not mean that that can be ignored or excluded, since it might only be the case that Greg has not be trained for that particular pathology. It can be seen that handling a large number of diagnoses is crucial, in this respect.

– Second, with each diagnostic suggestion Greg associated a degree of probability, i.e., it ranks suggestions with a confidence measure. This is important, since the tool may provide several different suggestions for a given profile, and not all of them are to be considered as equally relevant.

It can be seen that a tool like Greg is as effective as seamless its integration with the everyday procedures of a medical institution is. To foster this kind of adoption, Greg can be used as a stand-alone tool, with its own user-interface, but it has been developed primarily as an engine-backed API, that can be easily integrated with any medical information system that is already deployed in medical units and wards. Ideally, with this kind of integration, accessing medical suggestions provided by Greg should cost no more than clicking a button, in addition of the standard procedure for patient-data gathering and medical-record compilation. Finally, we also developed a stand-alone Web app called the Greg Playground. The app is available at URL `https://demo.svelto-spinoff.it/Greg-ML-Playground/`.

References

1. Deo, R.C.: Machine learning in medicine. Circulation **132**(20), 1920–1930 (2015)
2. Holzinger, A.: Machine learning for health informatics. In: Machine Learning for Health Informatics, pp. 1–24. Springer (2016)
3. Kononenko, I.: Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine **23**(1), 89–109 (2001)
4. Lapadula, P., Mecca, G., Santoro, D., Solimando, L., Veltri, E.: Humanity Is Overrated. or Not. Automatic Diagnostic Suggestions by Greg, ML. In: New Trends in Databases and Information Systems. pp. 305–313. Springer International Publishing (2018)
5. Mohammed, O., Benlamri, R.: Developing a semantic web model for medical differential diagnosis recommendation. Journal of medical systems **38**(10), 79 (2014)
6. Peek, N., Combi, C., Marin, R., Bellazzi, R.: Thirty years of artificial intelligence in medicine (aime) conferences: A review of research themes. Artificial Intelligence in Medicine **65**(1), 61–73 (2015)
7. Rajpurkar, P., Hannun, A.Y., Haghpanahi, M., Bourn, C., Ng, A.Y.: Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv preprint arXiv:1707.01836 (2017)
8. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)
9. Soni, J., Ansari, U., Sharma, D., Soni, S.: Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications **17**(8), 43–48 (2011)
10. Steadman, I.: IBM's Watson is better at diagnosing cancer than human doctors. WIRED (2013), http://www.wired.co.uk/article/ibm-watson-medical-doctor