

Report 5

Prof. Dr. Jacinto Nascimento
Prof. Dr. Nuno Nunes

Francisco Maria Calisto

Date: 09/04/2018

1 Problem Statement

The hereby **Report 5** will state an essay for literature knowledge that might support our research. We evaluate several research work. This report will focus on the topic towards (i) **Gamification & Crowdsourcing** application on *Health Informatics*; (ii) **Gathering Annotated Data**; and (iii) **Interactive Learning**. All topics are of chief importance to our research work, therefore should be analysed.

2 Investigation/Research

In this section we describe our investigation and research over the proposal readings. Investigation throughout the examination of the reading facts, novel possibilities and results. Also, addressing the research reasoned conclusions of those readings.

2.1 Gamification & Crowdsourcing

A paper with the called "*Dr. Detective*": *combining gamification techniques and crowdsourcing to create a gold standard in medical text* authored by Dumitrache et al. [3] proposes a design for a gamified crowdsourcing workflow to extract annotation from medical data. The importance of this paper to our work is trivial; it will support our research questions regarding medical annotations. In a context of a general crowdsourcing platform, the authors developed a game with the purpose of engaging medical experts in solving annotation tasks. The authors have incorporated incentives such as learning features, to motivate a continuous involvement of the expert crowd. In our work, we are also interested in a technique to engage our group, the Radiologists and Medical Imaging experts. Therefore we will have more and better annotations. The game was designed to identify valuable data for training machines and interpret the data relation in the context of medical diagnosing.

Supported by this paper and the work done by Aroyo et al. [1], we can solve our main problem in gathering a ground truth from experts. By reporting their results, both authors showed that the quality of the annotations by the expert crowd is higher by allowing the game users to access each others' annotations. Therefore it will increase both quality and number of annotations. The authors are proposing to test how each of the gaming features performs individually, to find knowledge regarding the application influence on the data quality. Nevertheless, the volume of the result to adapt the best needs of the user is also important. From this work, we can explore how to integrate the gaming and task crowdsourcing workflows on our work, by using the output of the workflow or by asking the crowd to validate the output.

2.2 Gathering Annotated Data

From the above section we went deeper into gathering annotated data. That said, we analysed several research works regarding this topic. In a paper with the title *CrowdTruth: Machine-Human Computation Framework for Harnessing Disagreement in Gathering Annotated Data* from Inel et al. [4], the authors are introducing a *CrowdTruth* framework for machine-human computation. The implementation of this framework is a novel approach to gathering human annotation data in a wide range of annotation tasks and on a variety of media. Despite of this paper works on text, images and videos, we will focus just on images since we are working with medical imaging, only.

The *CrowdTruth* approach will be useful for us since it captures human semantics through a pipeline of three processes. The pipeline of processes are as follows on the next list.

List of the three processes:

1. Combining various machine processing of medical image in order to understand better the input content and optimise the task suitability, optimising the time and cost regarding the crowdsourcing process;
2. Providing reusable human-computing task rules to collect the maximum diversity in the human interpretation, thus collect richer human semantics;
3. Implementing *CrowdTruth* metrics to support a deep analysis of the quality and semantics of the crowdsourcing data;

The authors are in this paper demonstrating the applicability of the approach to similar works as ours. This paper also showed us the advantages of using their open standards and their extensibility of the framework with our medical imaging multimodalities and annotation tasks.

2.3 Interactive Learning

The work done by Guerra et al. titled as *Interactive Optimal Teaching with Unknown Learners* (not yet published) where the authors introduce a new approach for machine teaching, addressing a mismatch between what teacher assumes and the actual process about the learning achievements. The authors analysed several situations, by focusing their analysis on the case of a Bayesian Gaussian learning [6] process. Therefore, this work was useful for us since their study will help us as a guide towards a compelling analysis of an interactive learning process on a User Interface (UI). The authors are presenting in this paper two sets of results. First, by involving simulation data and, second, by learning the involving task. Their results compared interactive and non-interactive methods in different scenarios.

The user study of this work could be vital to us, while the authors present how to approximate the result information with theoretical results. The authors also performed a Mann-Whitney U tests [5] showing us how to compare the distributions obtained with the several methods. This comparison might be relevant to our work since we are also applying and analysing several interactive processes. Also, the authors performed the Kruskal-Wallis H tests [2] to found out the use of different values.

3 Final Evaluation

We evaluated three research works, starting by an introduction towards both **Gamification & Crowdsourcing** [3, 1] topics while it describes the induction of **Gamification** and **Crowdsourcing** into the *Health Informatics* field.

The second is related to **Gathering Annotated Data** [4] where authors showed the advantages of their open standards and framework extensibility with our medical imaging multimodalities and annotation tasks. Third and final, it was introduced a new approach for machine teaching [6], while their results compared interactive and non-interactive methods in different scenarios. The comparison might be relevant to our work is presenting the application and analysis of several interactive processes. All three jobs are necessary because of the above reasons. So that, it is essential to analyse the two works and authors' research further.

References

- [1] Lora Aroyo and Chris Welty. Crowd truth: Harnessing disagreement in crowdsourcing a relation extraction gold standard. *WebSci2013. ACM*, 2013, 2013.
- [2] Ton J Cleophas and Aeilko H Zwinderman. Non-parametric tests for three or more samples (friedman and kruskal-wallis). In *Clinical Data Analysis on a Pocket Calculator*, pages 193–197. Springer, 2016.
- [3] Anca Dumitrache, Lora Aroyo, Chris Welty, Robert-Jan Sips, Anthony Levas, et al. Dr. detective: combining gamification techniques and crowdsourcing to create a gold standard for the medical domain. 2013.
- [4] Oana Inel, Khalid Khamkham, Tatiana Cristea, Anca Dumitrache, Arne Rutjes, Jelle van der Ploeg, Lukasz Romaszko, Lora Aroyo, and Robert-Jan Sips. Crowdtruth: Machine-human computation framework for harnessing disagreement in gathering annotated data. In *International Semantic Web Conference*, pages 486–504. Springer, 2014.
- [5] Thomas W MacFarland and Jan M Yates. Mann–whitney u test. In *Introduction to Non-parametric Statistics for the Biological Sciences Using R*, pages 103–132. Springer, 2016.
- [6] Xiaojin Zhu. Machine teaching for bayesian learners in the exponential family. In *Advances in Neural Information Processing Systems*, pages 1905–1913, 2013.