MESE 1

Measurement in STEM Education (MESE1)
Naples, 30-31 January, 1 February

DISP Dipartimento di Scienze Politiche Università degli studi di Napoli Federico II   FISICA "ETTORE PANCINI"

Contribution ID : **41**   Type : **Invited talk**

Automated Assessment Of Open-Ended Question Of Invalsi Tests

martedì 31 gennaio 2023 16:45 (45)

This work describes the new procedures of automated corrections of freeform answers given by the 8th, 10th and 13th grade students to open-ended questions in CBT Computer Based Test) INVALSI tests. INVALSI team, composed of statistical and computer scientists, responsible of open-ended question correction, has implemented an algorithm to process text strings of different complexity. Before survey distribution, the correction team and the items authors group discuss to define the correction criteria, that is a set of rules to determine the correct or incorrect classification for each answer given by the students for a specific item. The discussion produced, moreover, the indications to remove useless elements for the classification, then translated in operations of the algorithm on the textual data such as punctuation detection and removal, special characters, articles, conjunctions, word lemmatisation, etc. The answer strings were subsequently processed by a "data cleaning" operation, that was focused on the automated correction of spelling and typing errors, by detection and substitution of "out-of-vocabulary" words (OOV words). After the "data cleaning" phase, the correction criteria fixed by the experts have been translated in logical IT patterns, aiming to uniquely defining the set of admissible ways to give a correct answer. The last test phases of the algorithm were characterized by a constant exchange of information about the encoding, among the authors' team and the correction team, this passage being critical to refine the logical rules used for correction and to get more consistency and precision between the encoding produced by the algorithm and the authors' indications. The final test of the algorithm ends with a comparison between the manual encoding by video correction and the one processed by the algorithm on a set of items already processed in a former test: the algorithm is accounted as accurate enough and aligned to the indications of authors' team when the complete accordance of the two encoding was achieved. The methodological approach, countable as a method of supervised automated correction, represents a valid compromise between a manual encoding and a totally automated one, typical of the machine learning algorithms. This method has indeed the benefit of considerably reduce the hours/man needed to correct the open-ended answer items, when compared to a manual procedure, and get a better accuracy reducing the wrong encoding matches, when compared to a non-supervised automated procedure. A comparison between supervised and non-supervised automated procedure has been eventually done to evaluate the distance between the two methodological approaches.

Research Strand

Presenter(s) :   Dr. MARSILI, Michele (Invalsi)

Session Classification : Keynote 7