

January 2011

Web-based Tool for Automatic Assignment Evaluation

Dr. Rama Gautam

Professor & Head, Department of Communication School of Humanities International Institute of Information Technology Pune, India, ramagoutam@gmail.com

Nikhil Bhatt

Software Engineer Motorola India Pvt. Ltd. Bangalore, India, nikhilbhat@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijcct>

Recommended Citation

Gautam, Dr. Rama and Bhatt, Nikhil (2011) "Web-based Tool for Automatic Assignment Evaluation," *International Journal of Computer and Communication Technology*. Vol. 2 : Iss. 1 , Article 1.

DOI: 10.47893/IJCCT.2011.1061

Available at: <https://www.interscience.in/ijcct/vol2/iss1/1>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Computer and Communication Technology by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

Web-based Tool for Automatic Assignment Evaluation

Dr. Rama Gautam

Professor & Head, Department of Communication
School of Humanities
International Institute of Information Technology
Pune, India

Nikhil Bhatt

Software Engineer
Motorola India Pvt. Ltd.
Bangalore, India

Abstract— The demands for teachers in most educational institutions are continually increasing. Teachers have to teach an increasing number of students every year while simultaneously providing meaningful and constructive feedback and correcting assignments in a minimal amount of time. Consequently, results and assessment of student's progress can be considerably delayed. We need easy to use technology for grading and assessing test results and communicating information to students. Changes can be brought about to help teachers become more efficient and effective. This paper examines the role of technology and Web-based software in the classroom for evaluating a students' assignment. The paper has used both primary and secondary sources of data.

Keywords- Automation, Assignment Evaluation, AutoEval.

I. INTRODUCTION

In the current debate about nation wide educational restructuring, perhaps no issue is more central to the concerns of equity than that of student assessment. We have a long history of using questionably relevant tests to sort children for differential educational opportunities. Awareness of how standardized testing shapes curriculum and teaching highlights the link between assessment and educational quality. Yet, there is no consensus about how educational reform is to be achieved or what the role of student assessment should be. Politically powerful advocates of "outcome based" education argue that high standards and a national system of testing will accomplish needed educational improvement. There is a wide interest in the academic community in how technology can be used to support assessment. Today, most people working in the sectors of IT and Education are aware of the immense potential that Information Communication Technologies (ICT) has in the sphere of education, especially in assessments. However, insufficient attention has been paid to the tools and software that can be used to make assessments easier, valid and relevant; with the result that even the simplest form of ICT based assessment tools are unknown or unused.

Assessments refer to the techniques such as tests and projects from which grades are derived and lead to evaluation. Some of the problems with student evaluation and assessment are as follows:

- Evaluation is a highly inconsistent process. Teachers give different numbers and types of assessments and weight them differently.
- There is disagreement on issues like the role and value of homework. Some teachers assign homework frequently and weigh it heavily, while some don't assign it at all.
- Some teachers will allow retakes of tests and quizzes, others do not.
- Different policies exist for work turned in late.
- Different final grades as a passing mark
- The validity and reliability of student assessments vary.

Teachers strive for grading objectively, but human judgment is always involved. Most teachers dread that time of the year: assessment time! Most teachers are burdened with stacks of answer books, many of a mediocre quality which they have to wade through within a short stipulated time! Teachers in Humanities who have essays, comprehensions, précis and grammar exercises to correct are greatly affected. However, this is not to say that the teachers of math's or the sciences have an easy time, though since objective questions constitute a large part of their question paper, it is comparatively easier for them.

Notwithstanding this, there is no doubt that teachers do consider assessments as an unavoidable and painful process of evaluation and as a result, the quality of the question papers can be suspect. There might be a temptation to set an easier, shorter, 'easier to correct' question paper. Will such a paper be able to test the objectives of teaching a particular subject? Will the test be reliable or valid? So many questions can be raised about the results of the test. Teachers will sincerely assert that they have been "objective" in assessment and evaluation, but when administrators compare student grades they are often baffled. The temptation is to label some teachers as "good" and some as "bad." The "good" classification will likely be applied to the teachers whose students achieve the best grades; but without a clear understanding— it is not possible to judge teachers by their students' grades.

The difficulties have been mentioned, but let us not forget the opportunities. There has to be a way by which we can give back the teachers the time to teach and this can be achieved in a variety of ways. Many different types of technology can be used to support and enhance learning. Everything from

video content and digital moviemaking to laptop computing and handheld technologies (Marshall, 2002) have been used in classrooms, and new uses of technology such as podcasting are constantly emerging. Technologies available in classrooms today range from simple tool-based applications (such as word processors) to online repositories of scientific data and primary historical documents, to handheld computers, closed-circuit television channels, and two-way distance learning classrooms.

Even the cell phones that many students now carry with them can be used to learn (Prensky, 2005). Each technology is likely to play a different role in students' learning. Rather than trying to describe the impact of all technologies as if they were the same, researchers need to think about what kind of technologies are being used in the classroom and for what purpose. (Reeves, 1998 & Ringstaff & Kelley, 2002).

Without standard procedures for assessment and evaluation it is difficult if not impossible to ascertain the effectiveness of teaching and the efficiency of learning. Scores on various assessments and standardized tests vary greatly, but where does the accountability lie? With so much inconsistency it is hard to glean meaningful data from scores. Some of the advantages in the use of IT for Assessment are:

- More frequent formative and summative assessment.
- Staff can be alerted sooner to adapt their teaching.
- Can spend less time marking.
- Self-assessment; in the student's own time, at their own pace, when they are ready.
- Increased student confidence.
- Students like rapid results.

The objective of this paper is to suggest a method/system by which teachers will not be burdened with assessments as an ICT assessment tool can do this task easily, effectively and in time. This system can also provide immediate feedback on performance, which will be much appreciated by students. Staff can also use this for prompt diagnosis of any important areas of difficulty. Automatic marking, and comprehensive statistical analysis of results release staff to pursue other areas related to enhancing their students' learning. In this paper we have examined how technology can be used to assess a student's performance by suggesting a working model for assessment.

II. THE TECHNIQUE

There is a lot of literature on the subject of automated evaluation of short answers, essays and so on. While most of the effort is towards checking the correctness of English language in the responses, there is not much specifically

related to 'content'. There have been some seminal efforts on this front by the ELT group (E-rater) and it has proven to be quite successful. It has been successfully demonstrated and used in competitive exams like GMAT.

There are multiple approaches that can be taken to form a Automatic Assessment evaluation software. The tests involved are subjective in nature (as opposed to objective tests which are easier to grade), and therefore the results need to be analyzed in a different way. Most approaches use either a keyword driven approach to analyzing the results or by some form of Natural Language Processing (NLP). In this paper we aim to present a holistic view in the use of Automatic Assessment Evaluators using a keyword driven approach, backed with manual assessment.

We propose that a solution in this space, need to take a two pronged approach. One is to use the simplified keyword search approach, followed with the actual assessment as carried out traditionally. The aim is not to have a completely automated system, but for the system to be used more like an assistant. We recognize the importance of a completely automated system, but also like to point out the difficulties that are presently encountered in them such as reliability, efficiency and cost. The idea here is to utilize whatever tools are available for an easy, faster and less costlier solution which can give immediate benefits in terms of effectiveness and scalability if required in future.

An automated tool AutoEval, derived and delivered by the use of ICTs, is used to check for basic correctness in assessments and use it as an input for the manual corrections. In some form, it also tries to eliminate the bias that can come about when evaluating. The output of step one helps in prioritizing assessments and the focus is on faster evaluations. Based on a keyword search approach, the results are classified in 3 or 4 separate classes.

There is several other data that can be extracted when we use this approach. For example, the approximate pass percent, the median scores of all candidates etc can be used to determine many other useful data, quickly and with a fair degree of reliability.

Keeping the above in view, the tool we have developed tries to incorporate keyword search and an evaluator's dashboard for deriving other useful data.



Figure 1. AutoEval

We propose the use of a MySQL backend database which acts as the store house of all questions and their respective

answers. This does take some time and effort on part of the assessors, but we believe it is a one-time activity and the benefits quite surpass the efforts involved in this process. Moreover, the system is not required to do a one-to-one mapping of the answers received with that in the system. Since it is a keyword search approach, it only requires that those keywords are mentioned in the answers given, which also means that the entries made in the database need only mention the keywords.

If all the keywords expected are present, it would mean that the student is fairly familiar with the subject matter as he/she knows what is involved in constructing the answer. Of course this poses a lot of other problems in assessment like accuracy, predictability of the system and so on, and that is why the follow up with a manual method is required. It is for these reasons that the tool is useful when the answers are fairly objective and does not involve any subjective assessment or analysis. It is also a requirement that the keywords are not spelt incorrectly. There is a provision made in the database to consider the noun form of the keywords and all other related forms of the same word (with the use of wildcards). In case that itself is incorrect then the system will treat it as a fault.

Ultimately, the tool can help in giving only a rough idea of the overall metrics of the results. There are many operational issues such as maintaining the servers, updating the database and having a fairly good network in place that are the pre-requisites to derive the best benefits that it can offer.

Currently the tool is being further developed to extract useful metrics and data. It is not too difficult to customize it and it can also be developed in-house. A basic knowledge of programming and scripting along with some experience in constructing database tables is enough to develop a fairly good model.

III. RESEARCH METHODOLOGY

Objective Test

A test based on a specific chapter was administered to the First Year students of an undergraduate college. This test was objective in nature and had 20 questions; all related to the chapter under study. There were 20 questions in all. The students were 30 in all, and their answer books were given to their teacher to correct. The same questions and the correct answers were also fed into the computer, using the software tool AutoEval. The time taken by the teacher and the time taken by the computer were noted down. The test was for 20 marks.

Subjective Test

A test comprising of two essay type questions were given to the same batch of 30 students. Each question carried 10 marks, cumulative total being 20 marks. The computer was fed the questions, a model answer, and some key words. Then the answers of the students were fed and the result tabulated.

The test answer books had been already assessed by the same teacher and the time taken for correcting the answer books also noted.

IV. FINDINGS

In the first case, the objective test, the time taken by the computer for assessing all 30 assessments was less than a minute. The time taken by the teacher who did the assessment manually was 150 minutes. That is a significant reduction in the time.

This shows that automatic evaluation was definitely faster than the manual evaluation. Greater the number of assessments more will be the impact. There was no difference in the marks allotted by the computer and the teacher.

In the second case, for the subjective test, it was found that in terms of time AutoEval again did a great job, clocking in less than one minute while manual assessment took 300 minutes.

However, it was also found that if the student had used synonyms of the keywords, it was not accepted by the computer. Hence, the total-synonyms too need to be fed into the computer as possible variations. Some more options need to be given.

Secondly, the computer was unable to take care of ambiguous statements which did not contain the key words. Some of the students had written the correct answers but in a roundabout manner, ambiguously. So ambiguities too should be considered. This can be done with the help of NLP (Natural Language Programming). This can be one area where further research can be conducted.

The differences in marks between the evaluation of the computer and the evaluation of the teacher were noted. It was found that in some cases the difference between the marks was quite large. The computer in most instances gave lesser marks than the teacher. The researchers attributed it to various reasons. Objectively, it could be because of the ambiguities and the synonyms used.

A few language and grammar errors too were missed out by the computer but these were minimal. The researchers believe that once the computer has done its job of evaluating the paper and giving the suggested scores, the subject teacher can just take a closer look at it, on the lines of a moderator and confirm or change the evaluation or assessment to make it more perfect.

On the basis of the tests and the results it can be said that objective tests can very affectively be assessed by AutoEval. It can lead to reduction of time and labor.

The subjective tests assessment also is quite impressive, but not uniform. A little more research needs to be done on this, and surely with the help of NLP, these tests also can be administered and assessed more effectively.

An effective, efficient and appropriate assessment tool can assess the knowledge of the trainees or students and also check to see if they are able to apply it in certain given situations.

The changing environment and increasing complexity of the 21st century workplace, flatter organizations, a more diverse

employee base and greater use of teams have all made communication essential to organizational success. Globalization and increased international business are also important factors in the changing business environment. With new technologies the thought of globalization becomes a reality for even the smallest of companies. Various Information Communication Technologies like video conferencing, blogging, videophone, Internet chat and Internet meeting rooms, e-mails, collaborative technologies and innovative intranets are being used to conduct various transactions in companies. Students should be given opportunities to learn about various technologies that are being used to make things more efficient and accurate. Why not start with their assignment and tests assignments?

V. CONCLUSIONS

Computer-based assessment is clearly set to play an increasingly important role in education. The roles of computer-based learning and computer-based assessment are closely linked and progress in one depends on progress in the other. Support is also needed for teachers and lecturers as they learn how to take advantage of both roles. The advantages that can be derived by incorporating the ICT based tool for assessment are many and it is important to integrate computer based assessment with wider institutional assessment practice. There is much to be done to reshape educational procedures into a profession that makes use of data and research. Teachers and administrators must become more involved in staff development that is centered on deciding “what works, why it works, and how it is known that it works.”

Technology can be used for assessment purposes at various levels ranging from the management of the assessment information to a fully automated assessment system. Using technology for the management of assessment information can enable information to be presented in different ways to meet the needs of different audiences (such as teachers, students, course organizers and external examiners). In a fully automated assessment system all aspects of the system from the assessment which the student completes to the processing and administration of the marks, including the overall management of assessment information, can be technology-based. The overall balance of assessments in a course is of vital importance, and although computerized testing facilities can provide a rapid means of assessing and providing feedback to large numbers of students it is essential to consider their use as part of the overall course/unit strategy, especially as multiple choice/limited response type questions can lead to an emphasis on “shallow” learning. Electronic assessment tools are unlikely to reduce significantly the burden of assessment, but they can be used to promote deeper and more effective learning, by testing a range of skills, knowledge and understanding. A wide range of innovative assessment methods lend themselves to computer based implementation.

There needs to be a rebirth of instruction that emphasizes intellectual prowess and relies more on the science of learning and teaching. Neither excellence nor equity in education can be achieved as long as student assessment instruments, policies and practices limit opportunities to learn and narrow or dilute curricula and instruction. Both excellence and equity goals can, on the other hand, be served by assessments that help teachers to identify students' strengths as well as their needs and to determine the most appropriate and effective means of helping them to learn and grow. “We can't solve problems by using the same kind of thinking we used when we created them.” — A. Einstein

References

- [1] Crook, Charles (1994) 'Computers and the collaborative experience of learning'. Routledge, London.
- [2] CTISS (1995) "Computer-Assisted Assessment", Active Learning Issue No.1CTISS.
- [3] Darby J (1992) 'Computers in teaching and learning in U.K. higher education' Computers Educ, 19,1/2,1-8.
- [4] Edwards J D (1993) 'Benefits, Uses, Potential and Pitfalls of CBL', University of Bristol.
- [5] Kulik C.C. & Kulik J.A. (1991) 'Effectiveness of Computer Based Instruction: An updated analysis'; Computers in Human Behaviour, 7,75-94.
- [6] Kulik J.A, Kulik C.C. & Cohen P.A. (1980) 'Effectiveness of Computer Based College Teaching: A meta analysis of findings'; Review of Educational Research, 50, 4, 525-544.
- [7] Laurillard D (1993) 'Rethinking University Teaching - a framework for the effective use of educational technology', Routledge, London & New York.
- [8] MacFarlane, Alistair G J (1994) 'Future patterns of teaching and learning' NCET (1995) "Using IT for Assessment: Key Issues", NCET
- [9] Thorndike, Robert L (1971) "Educational Measurement" (2nd Edition), American Council for Education, Washington DC.
- [10] Whittlestone K, Howe G & Longstaffe A (1993) 'Getting Started with Computer Based Learning'.

Websites:

- [11] <http://trainingpd.suite101.com/article.cfm/educational-problems-with-grading-students#ixzz0nELCSfwk>
- [12] <http://www.maec.org/educate/11.html>