A Review on Massive E-Learning (MOOC) Design, Delivery and Assessment

doi: 10.1109/3PGCIC.2013.37

Abstract: MOOCs or Massive Online Open Courses based on Open Educational Resources (OER) might be one of the most versatile ways to offer access to quality education, especially for those residing in far or disadvantaged areas. This article analyzes the state of the art on MOOCs, exploring open research questions and setting interesting topics and goals for further research. Finally, it proposes a framework that includes the use of software agents with the aim to improve and personalize management, delivery, efficiency and evaluation of massive online courses on an individual level basis.

keywords: {courseware;distance learning;educational courses;software agents;MOOC assessment;MOOC delivery improvement;MOOC delivery personalization;MOOC design;MOOC efficiency improvement;MOOC efficiency personalization;MOOC evaluation improvement;MOOC evaluation personalization;MOOC management improvement;MOOC management personalization;OER;massive e-learning;massive online open courses;open educational resources;software agents;Educational institutions;Electronic learning;Market research;Monitoring;Software agents;MOOC;OER;agents;e-learning;virtual learning environment (VLE)},

URL:
An Integrated Framework for the Grading of Freeform Responses.


Abstract: Massive open online classrooms (MOOCs) have the potential to educate millions of people around the world. Initial MOOC courses were in science and engineering disciplines, where the problems involve constrained choices and can easily be graded automatically. MOOCs must still find ways to deal with essays and short answers, which are required for classes in humanities and the social sciences, and are useful to a variety of other disciplines. Three of the general techniques for evaluating freeform content are self assessment, peer assessment, and AI assessment. We describe how these approaches are implemented in the edX platform, and we present an approach which integrates scoring and feedback from the three techniques in order to maximize accuracy and minimize student and instructor effort. This combined approach has the potential to offer greater accuracy and better feedback with less overhead than any technique in isolation. We present a preliminary implementation of the integrated approach, as built into the edX platform, as well as results from pilot experiments with self-assessment and peer grading.


Abstract: This study examined the feasibility of employing Project Essay Grade (PEG) software to evaluate web-based student essays that serve as placement tests at a large Mid-Western university. The results of two experiments are reported. In the rst experiment, the essays of 1293 high school and college students were used to create a statistical model for the PEG software. PEG identi ed 30 proxes (observed variables) that could be incorporated into an evaluation of the written work. In the second experiment, the ratings from a separate sample of 617 essays were used to compare the ratings of six human judges against those generated by the computer. The inter-judge correlation of the human raters was $r = 0.62$ and was $r = 0.71$ for the computer. Finally, the PEG software was an ef cient means for grading the essays with a capacity for approximately three documents graded every second. Cycle time from the web-submission of the document to producing a report score was about 2 minutes. Although PEG would appear to be a cost-effective means of grading written work of this type, several cautionary notes are included.
Assessing writing in MOOCS: Automated essay scoring and Calibrated Peer Review.


abstract:
Two of the largest Massive Open Online Course (MOOC) organizations have chosen different methods for the way they will score and provide feedback on essays students submit. EdX, MIT and Harvard’s non-profit MOOC federation, recently announced that they will use a machine-based Automated Essay Scoring (AES) application to assess written work in their MOOCs. Coursera, a Stanford startup for MOOCs, has been skeptical of AES applications and therefore has held that it will use some form of human-based “calibrated peer review” to score and provide feedback on student writing. This essay reviews the relevant literature on AES and UCLA’s Calibrated Peer Review™ (CPR) product at a high level, outlines the capabilities and limitations of both AES and CPR, and provides a table and framework for comparing these forms of assessment of student writing in MOOCs. Stephen Balfour is an instructional associate professor of psychology and the Director of Information Technology for the College of Liberal Arts at Texas A&M University.
Finding the WRITE stuff: automatic identification of discourse structure in student essays

doi: 10.1109/MIS.2003.1179191

Abstract: An essay-based discourse analysis system can help students improve their writing by identifying relevant essay-based discourse elements in their essays. Our discourse analysis software, which is embedded in Criterion, an online essay evaluation application, uses machine learning to identify discourse elements in student essays. The system makes decisions that exemplify how teachers perform this task. For instance, when grading student essays, teachers comment on the discourse structure. Teachers might explicitly state that the essay lacks a thesis statement or that an essay's single main idea has insufficient support. Training the systems to model this behavior requires human judges to annotate a data sample of student essays. The annotation schema reflects the highly structured discourse of genres such as persuasive writing. Our discourse analysis system uses a voting algorithm that takes into account the discourse labeling decisions of three independent systems.

keywords: {intelligent tutoring systems;learning (artificial intelligence);linguistics;natural languages;Criterion;annotation schema;essay-based discourse analysis system;intelligent tutor;machine learning;online essay evaluation application;student essays;voting algorithm;writing;Algorithm design and analysis;Application software;Feedback;Humans;Large-scale systems;Machine learning;Protocols;Testing;Voting;Writing},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1179191&isnumber=26484
Quality evaluation of online courses through rubrics and ADDIE model

doi: 10.1109/ICeLeTE.2013.6644380

Abstract: During the period 1998-2008, there was a 150% increase in the number of students who choose the online courses to carry out part of their credits. In 2013 it is expected that nearly 20% of college students will receive more then 80% of their credits by online courses. This scenario of online courses popularization requires attention to quality issues, as well as the establishment of mechanisms to their continuous improvement. The present paper proposes an analysis of online courses evaluation through the use of rubrics considering the use of the QualityMatters Program and its associated rubric QMRubric. The case studies presented on the paper indicates that the uncertainty regarding the phase at which the rubric could be used makes its use be a prescriptive use instead of promoting continuous monitoring. The reduced number of standards associated to QMRubric prevents a comprehensive assessment, especially related to didactic-pedagogical aspects, relevant for online course development. In this sense, it was made an alignment between the QMRubric and the steps planned in ADDIE Model. From the above alignment was also possible to perform a transposition of QMRubric as holistic rubric to analytical one, detailing the aspects included in the proposal for its effective implementation.

keywords: {continuous improvement; courseware; ADDIE Model; QMRubric; QualityMatters Program; continuous improvement; continuous monitoring; didactic-pedagogical aspects; online course development; online course evaluation; quality evaluation; Computer aided instruction; Educational institutions; Electrical engineering; Materials; Proposals; Standards; ADDIE; online courses; quality evaluation; rubrics},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6644380&isnumber=6644337
Outcome based education performance evaluation of capstone project using assessment rubrics and matrix

doi: 10.1109/MITE.2013.6756295

Abstract: Capstone project (CP) is an important part of every engineering and computer science discipline at undergraduate level. The main purpose of these capstone projects is to encourage students to apply the knowledge acquired during their studies in previous semesters, in particular a capacity to personally manage and conduct a complex, open-ended project to show how proficient they are in solving real world or the research based problems. The Outcome Based Education (OBE) is an education system that emphasis on outcomes measurement rather than inputs of curriculum covered. Outcomes may include a range of knowledge, skills and attitudes. In order to obtain the desired outcomes, teaching components and activities should be well organized, planned and continuously improved. We adopted OBE in our curriculum and for the course on project work we have written course learning objectives (CLO), i.e. at the end of the course the student should be able satisfy these objectives and they are mapped with the program outcomes (PO). This paper presents a novel method for the outcome-based assessment of engineering capstone project carried by Final Year Students. The assessment is done based on the rubrics written for each phase of the process. The outcome of the each phase is assessed by evaluation team and the guide using the assessment matrix which is based on assessment rubrics. Assessment matrix covers all the attributes/parameters for the assessment of each phase. Mapping of CLO-PO is done based on these attributes. Percentage of attainment of each objective and outcome are calculated.

keywords: {engineering education; further education; matrix algebra; teaching; Capstone project; assessment matrix; assessment rubrics; course learning objectives; outcome based education performance evaluation; program outcomes; Computer science; Conferences; Education; Object recognition; Software; Technological innovation; Testing; CLO; Capstone Project; OBE; PO; Rubrics},

URL:
Performance of students across assessment methods and courses using Correlation analysis


Abstract: Professional education aims to provide students with requisite knowledge and skills to face today's competitive environment. To achieve this currently practices like systematically assessing student learning and measuring learning outcomes are presently in place. The primary goal achieved by this process is the continued improvement of academic quality for the institution. This paper investigates the application of correlation analysis for finding relationships between skills acquired is one test with the other components like assignments, individual marks with the final result and the correlation of performances of students' across courses. The results obtained for the first study indicates that there exists moderately strong relationship. The second study shows that correlation coefficients of related courses are better while that of some unrelated courses have moderate value. This means that in case of related courses the skills acquired in a prerequisite course can help them to do better in the current course. This exploratory analysis when properly interpreted can help plan and retrospect and re-adjust the mode of assessment or course delivery to enhance student learning.

keywords: {correlation methods;data mining;educational administrative data processing;educational courses;statistical analysis;correlation analysis;education data mining;educational course;multivariate statistics;professional education;student learning;students assessment method;Conferences;Correlation;Correlation coefficient;Distance measurement;Educational institutions;Correlation analysis;Education Assessments;Evaluation;student performance},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6756359&isnumber=6756292
Anomalies detection in evaluation of essay type answers by multiple evaluators

doi: 10.1109/TALE.2013.6654413

Abstract: In this paper, we present a methodology to detect suspected anomalies in evaluation of essay type answer books based on relationship between marks awarded and symbols used during evaluation. We identified dominant symbols used by evaluators and their relationships with marks awarded based on analysis of evaluated essay type answer copies. Further, for each evaluator relationship between symbols used and marks awarded is distinct. We identified outliers for each evaluators and each subject based on relationship between marks awarded and tick symbol and termed them as suspected anomaly. All grievances of students related to evaluation were in the set of identified outliers.

keywords: {educational administrative data processing;security of data;anomaly detection;dominant symbol;essay type answer book;multiple evaluators;Computer science;Conferences;Correlation;Correlation coefficient;Educational institutions;Linear regression;anomaly;assessment methodology;essay type answer;evaluation;grievances;outlier},

URL:
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6654413&isnumber=6654377
Correlation between human assessment of essays and ROUGE evaluation of essays' summaries


Abstract: In this paper we have addressed the qualitative (human evaluation) and quantitative (ROUGE) evaluation of computer generated summaries of the students' essays. The experimental results show that there is a positive high correlation between ROUGE scores and human assessment of the essays (human assigned marks). We have also found out that human evaluation of the automatic summaries positively correlates with the human assessment of the essays. These correlations can be used to classify students' essays into broad bands of quality.

keywords: {classification; educational administrative data processing; information retrieval; text analysis; ROUGE evaluation; automatic classification; automatic summarization; classify student essay; computer generated summary; experimental result; human assessment; human assigned mark; human evaluation; information retrieval; positive correlation; qualitative evaluation; quantitative evaluation; source document; text length; Computer science; Current measurement; Gold; Humans; Measurement standards; Natural language processing; System testing},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5340933&isnumber=5340906
An unsupervised approach to automated selection of good essays

doi: 10.1109/RAICS.2011.6069393

Abstract: Evaluating essays automatically has been an area of active research for some time. In this paper, we propose an unsupervised technique to select a set of good essays from a large selection of essays written on the same topic. We use a 'bag of words' approach which does not require deep parsing. The approach is based on the content of individual essays and the divergence of the individual essay from the collection when the collection is considered as one large essay. The approach is unsupervised and does not require any reference text to build computational learning model. We evaluate our approach on a set of essays, written by different people, on a single topic submitted to a competition internally within our organization. The approach enables selection of good essays which have a good correlation with the human based selection.

keywords: {educational administrative data processing;natural language processing;unsupervised learning;Kullback-Leibler divergence;automated good essay selection;computational learning model;essay evaluation;human based selection;information retrieval;natural language processing;unsupervised technique;Equations;Feature extraction;Humans;Open wireless architecture;Pragmatics;Probability density function;Writing;Information Retrieval;Kullback-Leibler divergence;Natural Language Processing},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6069393&isnumber=6069258
Benchmarking quality in online teaching and learning: A rubric for course construction and evaluation.


Abstract:
Online courses have many components and dimensions. Both the form (structure) and the content (expression) are situated in an overall environment. The sum of these elements results in student outcomes and learning. In order to facilitate construction and evaluate the quality of an online course, a four-part rubric was designed to reflect:
Structure (Context, Organization, and Environment)
Content (Presentation of Information)
Processes (Relationships and Interactions)
Outcomes (Mastery of Content and Course Evaluation)

This rubric was designed to provide quantitative and qualitative standardized evaluation for faculty self-evaluation, peer evaluation, and administrator evaluation. The rubric was piloted at two universities and shown to be highly effective in eliciting effective and usable feedback for course instructors and program directors. It was concluded that a uniform rubric that can be applied to any discipline could facilitate evaluation of all online courses within a program to a set standard that can then be used for course enhancement and improvement with structured comprehensive evaluation from instructors, peers, or program directors. It was found that a well-designed course (structure), with relevant and credible information (content), as well as mechanisms for interaction and collaboration (processes), could result in enhanced student learning (outcomes).
Design of an Automated Essay Grading (AEG) system in Indian context

doi: 10.1109/TENCON.2008.4766677

Abstract: Automated Essay Grading (AEG) or scoring systems are not more a myth they are reality. These AI systems are blessings to the educational community where teachers normally face lots of problem while correcting students' essays. Valuation of huge amount of student essays within stipulated time frame, with feedback is a real challenge. But today, the human written (not hand written) essays easily evaluated by AEG systems easily. The TOEFL exam is one of the best examples of this application. The students' essays are evaluated both by human and automated essay grading system. Then the average is taken. AEG might provide precisely the platform we need to explicate many of the features those characterize good and bad writing and many of the linguistic, cognitive and other skills those underline the human capability for both reading and writing. They can also provide time-to-time feedback to the writers/students by using that the people can improve their writing skill. A meticulous research of last couple of years has helped us to understand the existing systems which are based on AI & Machine Learning techniques and finding the loopholes and at the end to propose a system, which will work under Indian context, presently for English language influenced by local languages. Currently most of the essay grading systems is used for grading pure English essays or essays written in pure European languages. In India we have almost 21 recognized languages and influence of these local languages, in English, is very much here. Due to the influence of local languages and English written by nonnative English speakers (ie. Indians) the result of TOEFL exams has shown lower scores against Indian students (also Asian students). This paper focuses on the existing automated essay grading systems, basic technologies behind them and proposes a new framework to overcome the problems of influence of local Indian languages in English essays while correcting and by providing proper feedback to the writers.

keywords: {artificial intelligence; educational computing; learning (artificial intelligence); English essays; artificial intelligence; automated essay grading; automated essay scoring; local Indian languages; machine learning; Artificial intelligence; Computer science; Cost accounting; Design engineering; Feedback; Feeds; Humans; Machine learning; Natural languages; Writing},

URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4766677&isnumber=4766377

An overview of current research on automated essay grading

Predicting the readability of short web summaries.

Automated Essay Scoring by Maximizing Human-Machine Agreement

Predicting Percent of Article Errors in ESL Student Essays