

# Automated Essay Scoring Feedback (AESF): An Innovative Writing Solution to the Malaysian University English Test (MUET)

Ng Sing Yii<sup>1\*</sup>, Bong Chih How<sup>2</sup>, Lee Nung Kion<sup>3</sup>, Hong Kian Sam<sup>4</sup>

<sup>1</sup>Faculty of Cognitive Sciences and Human Development, UNIMAS,  
93400 Kota Samarahan, Sarawak, Malaysia  
ngsingyii@gmail.com

<sup>2</sup>Faculty of Computer Science and Information Technology, UNIMAS,  
93400 Kota Samarahan, Sarawak, Malaysia  
chbong@fit.unimas.my

<sup>3</sup>Faculty of Cognitive Sciences and Human Development, UNIMAS,  
93400 Kota Samarahan, Sarawak, Malaysia  
nkleefcs@unimas.my

<sup>4</sup>Faculty of Cognitive Sciences and Human Development, UNIMAS,  
93400 Kota Samarahan, Sarawak, Malaysia  
hksam1964@gmail.com

\*Corresponding Author

**Abstract:** Recent advances in information communication technology (ICT) infrastructure can be harnessed to support and improve the quality of teaching and learning of English writing skills especially for second language context where rule based support is necessary. Essay writing is indeed the most demanding task to both teachers and students. From conducting the class to assigning of task as well as marking and providing feedback by teachers whereas from drafting essays to final submission and resubmission of essays by students require on-going iterative cycles to facilitate improvement. However, a common scenario is that the iterative process takes too much time, thus resulting in limited practise. An innovative solution to imitate such process is via the Automated Essay Scoring Feedback (AESF). AESF is a networked tool that has the ability to score and provide feedback to students' essays instantaneously. With the speed that exceeds human ability and accuracy of a human scorer, it is hoped that AESF can increase the frequency of essay writing in the class that eventually result in improvement in students' performance. This paper aims to highlight the novelty and rationale of having AESF, its design and features as well as how this tool can be blended into the writing classroom, particularly for the Malaysian University English Test (MUET) extended essay writing.

**Keywords:** automated essay scorer, paragraph scoring, feedback, networked

## 1. Introduction

Harnessing computational methods in essay marking is no longer a new issue and is being greatly expanded to large scale assessment including SAT, GRE, TOFEL and GMAT (Attali, Burstein, Russell, & Hoffmann, 2006). Automated Essay Scorers (AES) dated as early as 1966 by Professor Ellis Page with Project Essay Grade (PEG) (Page, 2003) and thereafter, Criterion (J. Burstein, Chodorow, & Leacock, 2004), Intelligent Essay Assessor (IEA), IntelliMetric ("IntelliMetric® | Vantage Learning," n.d.) and many more probably unpublished or in the process of being published. AESF is an example that slowly makes its way to be noticed and introduced in to the Malaysian classroom. Automated Essay Scoring Feedback (AESF) is a web-based instructional writing tool that can score and provide feedback to submitted essays instantaneously as demanded by users targeting at students at the pre University stage. AESF is trained based on the Malaysian University English Test and is hoped to simplify teachers' task and improve students writing ability by on-going and sufficient practice as they needed. The novelty lies in teacher autonomy, student autonomy, cultural sensitivity, and paragraph level grading. It employs Natural Language Processing (NLP) and Machine Learning (ML) for automation.

## 2. Background

Automation of essay marking employs sophisticated language processing technologies and statistical methods to analyse a wide range of text features with its corresponding values that are being internalised or learned by the system to score unknown essays (Li, Link, & Hegelheimer, 2015). The automation process is generally similar with human holistic scoring but with huge samples. Human evaluation of essays is usually based on marking schemes that outline rubric that delineates specific expectation on essays responses. A moderation process is based on small samples that serve as benchmark and eventual agreement in marking between 2 or more graders to fine tune the marking score (Attali et al., 2006).

Unlike humans, they can read and internalise the scoring rubric with their background knowledge and language processing skills, system on the other hand, requires huge pool of data for learning and training before it can score accordingly (Dikli, 2006). Once, the system has internalised the text features, it can score as accurate as human scorers and more reliable than human with great speed that excludes all human weaknesses of being bias, inconsistent and individual preferences (J. Burstein & Chodorow, 1999).

Currently, Malaysia lacks home-grown AES that is tailor-made for the Malaysian context especially for marking extended English language essays. There are some local systems that only cater for short answer subject related response where predetermined finite answer keys can be possible (Omar, Razali, & Darus, 2009). As for extended general English writing skills improvement system, this technology is not available as most research result published about AES effectiveness with the Malaysian students are based on commercially available system like Criterion and My Access!. The greatest drawback in such system is that the grading may not be valid because the training model is based on essays written by native language users (L1) while the marking criteria/scheme that is not necessarily the same as how the Malaysia teacher may have graded their students' essays. Therefore, it is unfair to grade second language users (L2) against L1 where essays may also be culturally different than the L2. Thus, if essays are not measured on the same yardstick, the scoring cannot be valid.

Therefore, there is an urgent need for a tailor-made tool that can help score essays reliably and validly. AESF targets the Malaysia University English Test (MUET) for prototyping because MUET students are at a stage just before varsity. This is a good platform to train students to use ICT for independent learning at tertiary education because they are required to use ICT for producing reports, assignments and thesis extensively. Writing via computer is a must and hence, utilising MUET students for AESF development and usage can be more fulfilling for students who see the need to use ICT apart from being more critical and mature in providing feedback on the usage of the system. With this, the AESF prototype can be further polished and transferable to other level of education in school.

### 2.1 The Development of AES

PEG was one of the earliest automated essay scorer that was devised by Ellis Page in 1966 that uses proxy measures to determine grade of the essays (Page, 2003; Rudner & Gagne, 2001). The features includes average word length, essay length, the use of commas and semicolon (Rudner & Gagne, 2001). This system lacks semantic aspects that lacks the consideration human ability to organise and make meaningful transactions (Kukick quoted in (Hearst, 2000).

Subsequently, Intelligent Essay Assessor (IEA), a system which considers the semantic value of essays was introduced (Lemaire & Dessus, 2001). This is achieved using Latent Semantic Analysis (LSA) technique to assess essays. This scoring technique assumes that "there is a hidden semantic space in each text which is the accumulation of all words meaning" (Jiang & Wei, 2012, p. 58). With the application of matrixes, unique words are extracted and associated with its importance through frequency count. The latent semantic space created gives essay its meaning (Landauer & Dumains, 1997). The meaning is dependent on the co-occurrence of words in the corpus used (Lemarie & Desus, 2001). Therefore, it can only be reliable if the corpus is reliable at the first place. The weakness of this technique is that it cannot represent the actual knowledge of the students because word order, syntax, logic & other information are being ignored (Landauer, Ladam, & Folts, 2001).