

ABSTRACT

Assessing the essay answer is a unique one. If the document has different answer between the key answers, that is not mean incorrect, but if the answer has similarity with the keys, they will get the scores. In the development of Automated Essay Scoring (AES), weighting and measuring the similarity level of essay documents using the TF-IDF method resulted in relatively small scores, impressed rigid and distant from actual scores. Thus an effort is needed to make a more optimal score with an improvement method TF-IDF like as TF-IDF-DF.

The text data in the evaluation document is first through the word extraction process to obtain the essence (term), then the next step is weighting the term using TF-IDF-DF method. To produce an appropriate assessment, a similarity comparison strategy is used to measure the level of similarity of evaluation documents with expert documents using the Cosine similarity method. The Weighting results converted to the scores with range between (0-100).

Testing was conducted on 10 of student answers data in text form and an expert answer as keyword document for comparison purposes. The resulting value was increased around (7-19) points. The correlation between manual and system judgements is a moderate positive correlation, which means there is a tendency when manually gives a high score result, then the resulting score by method will be high too (and vice versa). Through the methods of TF-IDF-DF, the score seems more natural and proportionate.

Keywords : automated essay scoring, TF-IDF-DF, cosine similarity

INTISARI

Menilai jawaban esai memiliki keunikan tersendiri. Jika isi teks berbeda dengan kunci jawaban belum tentu akan dinilai salah, melainkan jawaban yang memiliki kedekatan makna dengan kunci jawaban yang akan mendapatkan poin penilaian. Dalam pengembangan bidang *Automated Essay Scoring* (AES), pembobotan dan pengukuran tingkat similaritas dokumen jawaban esai menggunakan metode TF-IDF menghasilkan kisaran skor yang relatif kecil, sehingga terkesan kaku dan jauh mewakili representasi skor sesungguhnya. Dengan demikian diperlukan suatu upaya agar skor yang dihasilkan lebih optimal dengan *improvement method* TF-IDF yakni TF-IDF-DF.

Data teks pada dokumen evaluasi terlebih dahulu melalui proses ekstraksi kata untuk mendapatkan intisari (*term*) yang selanjutnya dilakukan pembobotan menggunakan metode TF-IDF-DF. Untuk menghasilkan penilaian yang tepat, dilakukan mekanisme perbandingan similaritas untuk mengukur tingkat kemiripan dokumen evaluasi dengan dokumen ahli menggunakan metode *Cosine Similarity*. Bobot yang diperoleh kemudian dikonversi menjadi skor antara (0-100).

Pengujian yang dilakukan pada 10 data evaluasi siswa menghasilkan peningkatan skor yang berkisar antara (7-19) poin. Korelasi antara penilaian manual dengan penilaian yang dilakukan melalui metode memiliki kecenderungan positif, yang artinya jika penilaian secara manual memberikan hasil skor yang tinggi, maka melalui metode skor yang dihasilkan akan cenderung tinggi pula (demikian sebaliknya). Melalui metode TF-IDF-DF, skor penilaian terlihat lebih natural dan proporsional.

Kata kunci – temu kembali informasi, penilaian esai otomatis, TF-IDF-DF, cosine similarity