

Analisa unjuk kerja sistem penilai esai otomatis berbasis algoritma generalized latent semantic analysis (GLSA) menggunakan metode spektral laplacian eigenmaps embedding = Performance analysis of generalized latent semantic analysis (GLSA) based automated essay grading system using laplacian eigenmaps embedding spectral method

Henry Artajaya, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20348930&lokasi=lokal>

Abstrak

Metode spektral Laplacian Eigenmaps Embedding (LEM) dapat memelihara kemiripan dokumen dengan baik dibandingkan dengan metode reduksi dimensi lainnya. Hal ini terlihat dari unjuk kerja sistem berbasis GLSALEM yang lebih baik jika dibandingkan dengan sistem lainnya pada percobaan. Peningkatan unjuk kerja tidak hanya ditunjukkan dengan berkurangnya rata-rata selisih nilai yang dihasilkan oleh sistem dengan nilai yang dihasilkan oleh human rater tetapi juga jumlah percobaan dimana GLSA-LEM menghasilkan nilai yang paling mendekati dengan nilai yang dihasilkan oleh human rater. Kekurangan dari implementasi metode LEM adalah bahwa LEM hanya dapat diterapkan pada matriks jawaban referensi dan mahasiswa dengan dimensi yang lebih besar atau sama dengan enam. Oleh karena itu jawaban referensi dan jawaban mahasiswa yang terlalu pendek tidak akan dapat diproses oleh LEM. Hal ini dapat ditanggulangi dengan mengimplementasikan batas minimal kata jawab pada sistem berbasis GLSA-LEM sehingga semua jawaban dapat diproses oleh LEM. Pada percobaan ini didapatkan rata-rata selisih antara nilai yang dihasilkan oleh sistem dengan nilai yang dihasilkan oleh human rater pada sistem berbasis LSA adalah 44,49; pada sistem berbasis GLSA adalah 23,41; dan pada sistem berbasis GLSA-LEM adalah 11,67. Hasil tersebut menunjukkan bahwa GLSA-LEM paling unggul karena menghasilkan rata-rata selisih yang paling kecil antara nilai yang dihasilkan oleh sistem dengan nilai yang dihasilkan oleh human rater. Hal ini didukung oleh jumlah percobaan dimana sistem berbasis GLSA-LEM bekerja paling baik yakni dari sejumlah 245 percobaan yang dapat diterapkan LEM didapatkan bahwa pada 82 percobaan sistem GLSA-LEM menghasilkan selisih nilai yang paling kecil dibandingkan dengan sistem GLSA yang unggul pada 40 percobaan dan sistem LSA yang unggul pada 10 percobaan saja. Dengan demikian hipotesis yang diajukan terbukti benar bahwa implementasi LEM pada sistem GLSA akan meningkatkan akurasi sistem. Selisih nilai yang lebih kecil menandakan sistem dapat menghasilkan nilai yang lebih mendekati nilai yang dihasilkan oleh human rater. Hal ini sesuai dengan tujuan dari sistem penilai esai otomatis yang diciptakan untuk menggantikan kerja human rater dimana nilai yang dihasilkan harus dapat mendekati nilai yang dihasilkan oleh human rater. Rata-rata waktu proses LSA adalah 0,164 detik, GLSA sebesar 0,521 detik, dan GLSA-LEM sebesar 4,982 detik.

.....Laplacian Eigenmaps Embedding preserve semantic proximity better than other dimension reduction methods. GLSA performance may be improved further by implementing LEM. Experiment conducted has shown that GLSA-LEM based system has outperform on this experiment. Performance improvement not only shown from average delta between the grades calculated using the system and the grades resulted from human rater but also the number of the tests that outperformed by GLSA-LEM. The disadvantage of LEM implementation is that LEM only can be applied to answer matrices with minimum dimension of six. Therefore answers that are too short may not be processed using LEM. This can be mitigated by

implementing minimum threshold to the answers so it can't be submitted if less than required length. This experiment show that LSA average delta between grades resulted from the system and grades resulted from human rater is 44,49; GLSA's average delta is 23,41 and GLSA-LEM's average delta is 11,67.

These results show GLSA-LEM is the best because generate grades with the least average delta between the grades calculated using the system and the grades resulted from human rater. These results also supported by the number of essays from total of 245 essays that can be applied GLSA-LEM graded best with least delta by GLSA-LEM that is 82; compared to GLSA that is 40; and LSA that is 10. Therefore the hypotesis is proven to be correct that LEM implementation on GLSA based system improves system's accuracy. Least delta indicates system generate better grades that is closer to human rater. These results is in accordance with the purpose of automated essay grading system that created to replace human raters in which the grades resulted by the system should be close to the grade generated by human raters. LSA's average processing time is 0,164 seconds, GLSA's is 0,521 seconds, and GLSA-LEM's is 4,982 seconds.