

## DAFTAR PUSTAKA

- [1] N. Hunt, J. Hughes, and G. Rowe, "Formative automated computer testing (fact)," *British Journal of Education Technology*, p. 5, 2002.
- [2] M. Mohler, R. Bunesu, and R. Mihalcea, "Learning to grade short answer questions using semantic similarity measures and dependency graph alignments," *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies ACL-HLT*, pp. 752–762, 2011.
- [3] D. Bennouar, "An automatic grading system based on dynamic corpora," *International Arab Journal of Information Technology*, pp. 552–564, 2017.
- [4] A. Lajis and N. A. Aziz, "NI scoring and bloom competency test; an experimental result," *International Conference on Ubiquitous Information Management and Communication*, 2012.
- [5] N. Iadad, G. A. Dafoulas, E. Kalaitzakis, and L. A. Macaulay, "Evaluation of online assessment: The role of feedback in learner-centered e-learning," *Proceedings of the 37th International Conference on System Science*, 2004.
- [6] M. Mohler and R. Mihalcea, "Text-to-text semantic similarity for automatic short answer grading," *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics on - EACL*, pp. 567–575, 2009.
- [7] F. Rodrigues and P. Oliveira, "A system for formative assessment and monitoring of students's progress," *Computer and Education*, pp. 30–41, 2014.
- [8] N. Pachler, C. Daly, Y. Mor, and H. Mellar, "Formative e-assessment:practionary cases," *Computer and Education*, vol. 54, 2010.
- [9] G. Conole and B. Warburton, "Review of computer-assisted assessment," *Alt-F Research in Learning Technology*, vol. 13, 2010.
- [10] M. Limniou and M. Smith, "The role of feedback in e-assessments for engineering education," *Education and Information Technologies*, pp. 209–225, 2014.
- [11] T. Deutsch, K. Herrmann, T. Frese, and H. Sandholzer, "Implementing computer-based assessment â€" a web-based mock examination changes attitudes," *Computer and Education*, pp. 1068–1075, 2012.
- [12] S. Jayashankar and R. Sridaran, "Superlative model using word cloud for short answers evaluation in elearning," *Education and Information Technologies*, pp. 2383–2402, 2017.
- [13] S. A. Livingston, "Constructed-response test questions: Why we use them; how we score them," *R and D Connections: Educational Testing Service (ETS)*, pp. 1–8, 2009.

- [14] M. A. Sultan, C. Salazar, and T. Sumner, "Fast and easy short answer grading with high accuracy," *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 1070–1075, 2016.
- [15] M. J. A. Aziz, F. D. Ahmad, A. A. A. Ghani, and R. Mahmud, "Automated marking system for short answer examination (ams-sae)," *IEEE Symposium on Industrial Electronics and Applications, ISIEA*, pp. 47–51, 2009.
- [16] D. Perez, A. Gliozzo, and C. Strapparava, "Automatic assessment of students' free-text answers underpinned by the combination of a bleu -inspired algorithm and latent semantic analysis," *FLAIRS conference*, pp. 358–363, 2005.
- [17] S. Burrows, I. Gurevych, and B. Stein, "The eras and trends of automatic short answer grading," *International Journal of Artificial Intelligence in Education*, pp. 60–117, 2015.
- [18] Y. Tian, "Onile automated assessment of student english-chinese translation," *World Congress on Computer Science and Information Engineering*, 2009.
- [19] M. D. Shermis, "Contrasting state-of-the-art in the machine scoring of short-form constructed responses," *Educational Assessment*, pp. 46–65, 2015.
- [20] R. Siddiqi, C. J. Harrison, and R. Siddiqi, "Improving teaching and learning through automated short-answer marking," *IEEE Transactions on Learning Technologies*, pp. 237–249, 2010.
- [21] J. Z. Sukkarieh, S. G. Pulman, and N. Raikes, "Auto-marking: using computational linguistics to score short, free text responses," *The annual conference of the International Association for Educational Assessment (IAEA), Manchester, UK*, pp. 1–15, 2003.
- [22] A. Lajis and N. A. Aziz, "Automated scoring in context: Rapid assessment for placed students," *Assessing Writing*, pp. 62–84, 2013.
- [23] M. D. Shermis, "State-of-the-art automated essay scoring: Competition, results, and future directions from a united states demonstration," *Assessing Writing*, pp. 57–76, 2014.
- [24] J.-E. Kim, K. Park, J.-M. Chae, H.-J. Jang, B.-W. Kim, and S.-Y. Jung, "Automatic scoring system for short descriptive answer written in korean using lexico-semantic pattern," *Soft Computing*, pp. 4241–4249, 2018.
- [25] A. Stunden and D. Jefferies, "The effectiveness of short answers test papers in evaluating academic nursing programs: A review of the literature," *Nurse Education in Practice journal*, pp. 94–101, 2018.
- [26] Y. Attali, B. Bridgeman, and C. Trapani, "Performance of a generic approach in automated essay scoring," *Journal of Technology, Learning, and Assessment*, pp. 5–16, 2010.

- [27] N. N. Karanikolas, "Computer assisted assessment (caa) of free-text," *Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 2010.
- [28] D. Perez and E. Alfonseca, "Application of the bleu algorithm for recognising textual entailments," *Proceedings of the Recognising Textual Entailment Pascal Challenge*, 2005.
- [29] A. M. B. Omran and M. J. A. Aziz, "Automatic essay grading system for short answers in english language," *Journal of Computer Science*, pp. 1369–1382, 2013.
- [30] R. Ziai, N. Ott, and D. Meurers, "Short answer assessment: Establishing links between research strands," *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pp. 190–200, 2012.
- [31] F. S. Pribadi, T. B. Adji, and A. E. Permanasari, "Automated short answer scoring using weighted cosine coefficient," *IC3e (International Conference of e-learning, e-assessment, and, e-services)*, pp. 70–74, 2016.
- [32] —, "Short answer scoring system using automatic reference answer generation and geometric average normalized-longest common subsequence (gan-lcs)," *Education and Information Technologies*, pp. 70–82, 2018.
- [33] W. H. Gomaa and A. A. Fahmy, "Automatic scoring for answer to arabic test questions," *Computer Speech and Language*, pp. 833–857, 2014.
- [34] Noorbehbahani and Kardan, "The automatic assessment of free text answers using a modified bleu algorithm," *computer and education*, pp. 337–345, 2011.
- [35] C. Leacock and M. Chodorow, "C-rater: Automated scoring of short-answer questions," *Computers and the Humanities*, pp. 389–405, 2003.
- [36] J. Z. Sukkarieh, S. G. Pulman, and N. Raikes, "c-rater: Automatic content scoring for short constructed responses," *Proceedings of the Twenty-Second International FLAIRS Conference*, pp. 290–295, 2009.
- [37] P. Kudi, A. Manekar, K. Daware, and T. Dhattrak, "Ieee global conference on wireless computing and networking (gcwcn)," *Online Examination with Short Text Matching*, pp. 57–60, 2014.
- [38] T. Mitchell, T. Russell, P. Broomhead, and N. Aldridge, "Towards robust computerised marking of free-text responses," *In Proceedings of the 6th International Computer Assisted Assessment (CAA) Conference, Loughborough, UK*, pp. 233–249, 2002.
- [39] D. Perez, E. Alfonseca, and P. Rodriguez, "Upper bounds of the bleu algorithm applied to assessing student essays," *Proceedings of the 30th International Association for Educational Assessment (IAEA) Conference*, 2004.
- [40] B. Furlan, V. Batanovic, and B. Nikolic, "Semantic similarity of short texts in languages with a deficient natural language processing support," *Decision Support Systems*, pp. 710–719, 2013.

- [41] W. H. Gomaa and A. A. Fahmy, "Short answer grading using string similarity and corpus-based similarity," *International Journal of Advanced Computer Science and Applications*, pp. 115–121, 2012.
- [42] A. Islam and D. Inkpen, "Semantic text similarity using corpus-based word similarity and string similarity," *ACM Transaction Knowledge Discovery Data*, pp. 10.1–10.25, 2008.
- [43] R. Klein, A. Kyrilov, and M. Tokman, "Automated assessment of short free-text responses in computer science using latent semantic analysis," *Proceedings of the 16th annual joint conference on Innovation and technology in computer science education*, 2011.
- [44] R. Adhitia and A. Purwarianti, "Penilaian esai jawaban bahasa indonesia menggunakan metode svm - lsa dengan fitur generik," *Jurnal Sistem Informasi*, pp. 33–41, 2012.
- [45] L. F. Bachman, N. Carr, G. Kamei, M. Kim, M. J. Pan, C. Salvador, and Y. Sawaki, "A reliable approach to automatic assessment of short answer free responses," *Proceedings of the 19th international conference on Computational linguistics*, pp. 1–4, 2002.
- [46] A. Valdestilhas, T. Soru, and A.-C. N. Ngomo, "A high-performance approach to string similarity using most frequent k characters," *Proceedings of The Ontology Matching Workshop*, 2017.
- [47] E. F. Aries, "Asesmen dan evaluasi," *Aditya Media*, 2011.
- [48] Z. Arifin, "Evaluasi pembelajaran," *Direktorat Jenderal Pendidikan Islam Kementerian Agama*, 2012.
- [49] A. Lajis and N. A. Aziz, "Simple: Sistem penilai esai otomatis untuk meniai ujian dalam bahasa indonesia," *Makara, Teknologi*.
- [50] Suladi, "Paragraf," *Pusat Pembinaan dan Pemasyarakatan Badan Pengembangan dan Pembinaan Bahasa Kementrian Pendidikan dan Kebudayaan, Jakarta*, 2014.
- [51] S. Roy, Y. Narahari, and O. D. Deshmukh, "A perspective on computer assisted assessment techniques for short free-text answers," *International Computer Assisted Assessment Conference*, pp. 1–14, 2015.
- [52] F. S. Pribadi, T. B. Adji, A. E. Permanasari, and T. Ninomiya, "Short answer scoring using w-bleu for regular assessment in vocational high school," *International Journal Innovation and Learning*, 2018.
- [53] S. G. Pulman and J. Z. Sukkarieh, "Automatic short answer marking," *Proceedings of the second workshop on Building Educational Applications Using NLP*, pp. 9–16.
- [54] S. Jordan and T. Mitchell, "e-assessment for learning? the potential of short-answer free-text questions with tailored feedback," *British Journal of Educational Technology*, pp. 371–385, 2009.

- [55] R. Siddiqi and C. Harrison, "A systematic approach to the automated marking of short-answer questions," *EEE International Multitopic Conference*, pp. 329–332, 2008.
- [56] P. G. Butcher and S. E. Jordan, "Acomparision of human and computer marking of short free-text student responses," *Computer and Education*, pp. 489–499, 2010.
- [57] P. Thomas, D. Haley, A. deRoeck, and M. P. Computing, "Automatic short answer marking," *eLearn '04 Proceedings of the Workshop on eLearning for Computational Linguistics and Computational Linguistics for eLearning*, pp. 38–44.
- [58] J. Kadupitiya, S. Ranathunga, and G. Dias, "Sinhala short sentence similarity measures using corpus-based simi- larity for short answer grading," *Proceedings ofthe 6th Workshop on South and Southeast Asian Natural Language Processing*, pp. 44–53, 2016.
- [59] D. Yustina, "Penilaian otomatis terhadap jawaban esai pada soal berbahasa indonesia menggunakan latent semantic analysis," *Proceedings - Inovasi dalam Desain dan Teknologi, IDEaTech*, pp. 123–130, 2015.
- [60] D. Callear, J. Jerrams-Smith, and V. Soh, "Bridging gaps in computerised assessment of texts," *Proceedings - IEEE International Conference on Advanced Learning Technologies, ICALT*, pp. 139–140, 2001.
- [61] W. H. Gomaa and A. A. Fahmy, "A survey of text similarity approaches," *International Journal of Computer Applications*, pp. 13–18, 2013.
- [62] F. S. Pribadi, T. B. Adji, A. E. P. A. Mulwinda, and A. B. Utomo, "Automatic short answer scoring using words overlapping methods," *Engineering International Conference (EIC): AIP Conference Procciding*, pp. 1–6, 2016.
- [63] J. Burstein, S. Wolff, and C. Lu, "Using lexical semantic techniques to classify free-responses," *Text, Speech and Language Technology*, pp. 227–244, 199.
- [64] D. Perez, E. Alfonseca, and P. Rodriguez, "Application of the sc bleu method for evaluating free-text answers in an e-learning environment," *Proceedings of the Language Resources and Evaluation Conference (LREC)*, 2004.
- [65] M. Wolska, A. Horbach, and A. Palmer, "Computer-assisted scoring of short responses: The efficiency of a clustering-based approach in a real-life task," *Advances in Natural Language Processing, NLP 2014*, p. 2014, 298-310.
- [66] V. S. Kumaran and A. Sankar, "Towards an automated system for short answer assessment using ontology mapping," *International Arab Journal of e-Technology*, 2015.
- [67] D. Meurers, R. Ziai, N. Ott, and S. Bailey, "Integrating parallel analysis modules to evaluate the meaning of answers to reading comprehension questions," *International Journal of Continuing Engineering Education and Life-Long Learning*, pp. 355–962, 2011.

- [68] D. Meurers, R. Ziai, N. Ott, and J. Kopp, "Evaluating answers to reading comprehension questions in context: Results for german and the role of information structure," *Proceedings of the TextInfer Workshop on Textual Entailment, EMNLP*, pp. 1–9, 2011.
- [69] J. Carbonell and J. Goldstein, "The use of mmr, diversity-based reranking for reordering documents and producing summaries," *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, 1998.
- [70] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 311–318, 2001.
- [71] J. D. Evan, "Straightforward statistics for the behavioral sciences," *Pacific Grove : Brooks/Cole Pub. Co.*, 1996.
- [72] Sugiyono, "Metode penelitian kuantitatif kualitatif dan  $r$  &  $d$ ," *Alfabeta Bandung*, 2007.
- [73] R. Mihalcea, "Graph-based ranking algorithms for sentence extraction, applied to text summarization," *Proceedings of the ACL on Interactive poster and demonstration sessions, Association for Computational Linguistics*, 2004.
- [74] V. Spiliotopoulou-Papantoniou, "Models of the universe: Children's experiences and evidence from the history of science," *Science and Education*, 2007.