

MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERJADINYA PLAGIARISM MENGGUNAKAN JACCARD MEASURE

¹Aditya Anggraini Yoga Kresna Wijayanti (1110651172),

²Ratih Ayuninghemi S.ST, M.Kom, ³Deni Arifianto, S.Kom

Jurusan Teknik Informatika Fakultas Teknik Univertas Muhammadiyah Jember

E-mail : aditya.taur05.ict@gmail.com

ABSTRAK

Dokumen laporan praktek kerja lapangan dibuat setelah melakukan praktek kerja lapangan pada suatu instansi pemerintah maupun swasta. pada laporan masih dapat ditemukan kemiripan atau plagiasi pada isi laporan tanpa mencantumkan sumber referensi. Hal ini diperlukan agar dapat meningkatkan kualitas dokumen laporan praktek kerja lapangan. Penelitian ini membahas kemiripan beberapa dokumen laporan untuk mendeteksi kesamaan dari isi dokumen tersebut. Untuk itu diperlukan suatu mekanisme mendeteksi plagiasi dokumen. Tahapan dalam metodologi penelitian terdiri dari: pengumpulan dokumen, isi laporan praktek kerja lapangan sebagai dataset, melakukan proses *text mining* yaitu : *casefolding*, *tokenizing*, *stemming* dan *filtering*, pembobotan dengan *TF-IDF*, perhitungan dengan metode *jaccard measure* sebagai alat untuk menukan apakah dokumen termasuk dinyatakan sebagai dokumen plagiat atau tidak plagiat dengan melihat dari isi laporan praktek kerja lapangan tersebut. Sebelum di lakukan penghitungan dengan *jaccard measure* , di lakukan pra proses *TF-IDF* terlebih dahulu, yaitu tingkat kemunculan kata dari suatu kalimat dengan koleksi dokumen yaitu jumlah yang akan di hitung bobotnya, dari bobot tersebut akan dijadikan masukkan untuk *jaccard measure*. Hasilnya berupa pengkalsifikasian tersebut telah diberikan 3 nilai ambang batas yaitu 0.5, 0.6, dan 0.7 . Data uji coba diambil dari beberapa laporan praktek kerja lapangan sebanyak 50 dokumen. Hasil pengujian menunjukan bahwa metode *jaccard measure*, mampu mengklasifikasikan dokumen plagiat maupun tidak plagiat dengan nilai ambang batas yang diberikan. *threshold* 0.5 sebesar 88%, *threshold* 0.6 sebesar 87%, dan *threshold* 0.7 sebesar 88% . Dengan hasil rata-rata akurasi sebesar 88%.

Kata kunci : *text mining*, *tf – idf*, *laporan praktek kerja lapangan*, *koleksi dokumen*, *jaccard measure*.

CLASSIFYING FIELD WORK PRACTICE REPORT TO PREVENT PLAGIARISM USE JACCARD MEASURE

¹*Aditya Anggraini Yoga Kresna Wijayanti (1110651172),*

²*Ratih Ayuninghemi S.ST, M.Kom,* ³*Deni Arifianto, S.Kom*

Department of Informatics Faculty of Engineering

University of Muhammadiyah Jember

E-mail : aditya.taur05.ict@gmail.com

ABSTRACT

Document field practice reports made after doing field work practice in a government agency or private. the report can still be found similarities or plagiarism on the contents of the report without including the source reference. This is necessary in order to improve the quality of the report documents the practice of field work. This research report discusses the similarity of some of the documents to detect the similarity from the contents of the document. For that required an mechanism to detect plagiarism document. Stages in research methodology consists of: collection of documents, the contents of the report as a dataset field practice, perform text mining process that is: casefolding, tokenizing, stemming and filtering, with TF-IDF the weighting, Jaccard measure calculation method as a tool for mentukan whether documents including expressed as a document plagiarism or not plagiarism by looking at the contents of the report of the field practice. Before doing the calculation with the Jaccard measure, in doing the pre-process of TF-IDF advance, that is the level of occurrence of the word of a sentence with a collection of documents that amount to be calculated weight, of the weight will be used as fill for the Jaccard measure. The result is pengkalsifikasian has been awarded three threshold values of 0.5, 0.6, and 0.7. Trial data taken from some field practice reports as many as 50 documents. Test results showed that the method Jaccard measure, able to classify documents plagiarism or not plagiarism with a given threshold value. The threshold 0.5 of 88%, the threshold 0.6 of 87%, and the threshold 0.7 of 88%. With an average accuracy result of 88%.

Keywords: *text mining, tf - idf, field practice reports, collections of documents, Jaccard measure.*