

ANALISIS PERBANDINGAN METODE PEMBOBOTAN TF-RF DAN TF-ABS PADA KATEGORISASI DOKUMEN BERITA DI BALAI DIKLAT INDUSTRI DENPASAR MENGGUNAKAN KLASIFIKASI *K-NEAREST NEIGHBORS*

Oleh

I Kadek Wahyu Dananjaya, NIM 1815091039

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Balai Diklat Industri Denpasar adalah sebuah instansi pemerintahan yang bertugas melaksanakan pendidikan dan pelatihan bagi sumber daya manusia industri di bidang keahlian animasi, kerajinan dan barang seni. BDI Denpasar dalam pengelolaan kelas berita di sistem layanan Kabar Insan Oke ini masih menggunakan cara konvensional yang menyebabkan kinerja pegawai kurang maksimal. Oleh sebab itu diperlukan sebuah modul klasifikasi berita otomatis. Penelitian ini dilakukan untuk membandingkan performa klasifikasi berita pada BDI Denpasar menggunakan klasifikasi *K-Nearest Neighbors* (K-NN) dengan metode pembobotan kata TF-RF dan TF-ABS. Metode yang memiliki tingkat performa yang tinggi akan diimplementasikan pada modul klasifikasi berita. Penelitian ini dilakukan dengan cara mengumpulkan dokumen berita, *preprocessing*, *term weighting*, klasifikasi, validasi model dan pengujian. Klasifikasi *K-Nearest Neighbors* (K-NN) menggunakan parameter *n_neighbor* (k) yaitu k=3, k=5, k=7 dan k=9 menggunakan dataset sebanyak 324 dokumen berisikan 7 kelas yang diambil dari website BDI Denpasar. Berdasarkan hasil pengujian yang dilakukan, metode pembobotan kata TF-RF memperoleh tingkat performansi yang lebih tinggi atau lebih baik dalam pengujian *confussion matrix* pada nilai k=5 dengan *accuracy* adalah 71% dengan *precision* 73% dan *recall* sebesar 71%. Metode pembobotan kata TF-ABS nilai k yang performansinya tertinggi terdapat pada k=9 yang memperoleh *accuracy* sebesar 70%, *precision* sebesar 63% dan *recall* sebesar 70%. Jadi metode yang akan diimplementasikan pada modul klasifikasi berita yaitu TF-RF.

Kata Kunci: Klasifikasi, *Term Weighting*, TF-RF, TF-ABS, *K-Nearest Neighbors*

**COMPARISONAL ANALYSIS OF TF-RF AND TF-ABS WEIGHTING
METHODS ON NEWS DOCUMENT CATEGORY AT BALAI DIKLAT
INDUSTRI DENPASAR USING K-NEAREST NEIGHBORS
CLASSIFICATION**

By

I Kadek Wahyu Dananjaya, NIM 1815091039

Informatics Engineering Department

Information Systems Study Program

ABSTRACT

Balai Diklat Industri Denpasar is a government agency tasked with carrying out education and training for industrial human resources in the fields of animation, crafts and art. BDI Denpasar in managing news classes in the Kabar Insan Oke service system still uses conventional methods which cause employee performance to be less than optimal. Therefore we need an automatic news classification module. This research was conducted to compare the performance of news classification at BDI Denpasar using the K-Nearest Neighbors (K-NN) classification with the TF-RF and TF-ABS term weighting methods. Methods that have a high level of performance will be implemented in the news classification module. This research was conducted by collecting news documents, preprocessing, term weighting, classification, model validation and testing. The K-Nearest Neighbors (K-NN) classification uses the $n_neighbor(k)$ parameter, namely $k=3$, $k=5$, $k=7$ and $k=9$ using a dataset of 324 documents containing 7 classes taken from the BDI Denpasar website. Based on the results of the tests carried out, the TF-RF term weighting method obtained a higher or better level of performance in the confusion matrix test at a value of $k=5$ with an accuracy of 71% with a precision of 73% and a recall of 71%. The TF-ABS term weighting method with the highest performance value is found at $k=9$ which obtains 70% accuracy, 63% precision and 70% recall. So the method that will be implemented in the news classification module is TF-RF.

Keyword: *Classification, Term Weighting, TF-RF, TF-ABS, K-Nearest Neighbors*