

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pada saat ini data memiliki peran yang sangat penting dalam memenuhi sebuah tujuan dalam organisasi. Data yang dihasilkan oleh suatu instansi pasti akan terus-menerus semakin banyak sehingga pengelolaan data yang masih menggunakan cara manual akan menyebabkan aktivitas tidak efektif dan efisien. Salah satu data yang akan menjadi sangat besar dan tidak terorganisir seiring berjalannya waktu adalah data dokumen berita (Februariyanti, 2012). Berdasarkan peristiwa yang terus terjadi, dokumen berita dalam masa yang akan datang pasti akan semakin banyak dan akan terus-menerus diperbaharui. Dalam sebuah dokumen berita, terdapat jenis-jenis atau kelas-kelas yang mengkategorikan berita tersebut sesuai dengan bidangnya. Sangat penting dilakukannya pengklasifikasian berita, dimana ini berfungsi agar dokumen yang tidak terstruktur dikelompokkan kedalam kelas atau bidang yang menggambarkan isi dari dokumen (Februariyanti, 2012).

Pengklasifikasian dengan cara manual membuat aktifitas tidak efektif dan efisien di jaman digital seperti sekarang ini. Oleh karena itu, penggunaan sistem klasifikasi berbasis komputer sangat diperlukan agar aktivitas dalam pengklasifikasian dokumen berita lebih mudah dan tertata dengan baik serta dapat meminimalisir terjadinya kesalahan *input* kelas berita. Salah satu instansi yang memiliki permasalahan dalam aktivitas pengklasifikasian yaitu Balai Diklat Industri Denpasar.

Balai Diklat Industri Denpasar atau disingkat BDI Denpasar adalah sebuah instansi pemerintahan di bawah naungan Menteri Perindustrian yang memiliki tugas melaksanakan pendidikan dan pelatihan bagi sumber daya manusia industri dalam bidang keahlian animasi, kerajinan dan barang seni yang beralamat di Tohpati kota Denpasar. BDI Denpasar memiliki beberapa layanan didalamnya, antara lain: Diklat 3 in 1, Inkubator Bisnis, BIKIN *Makerspace*, Lembaga Sertifikasi Profesi (LSP) dan *E-Learning*. Diklat 3 in 1 adalah program pendidikan dan pelatihan yang meliputi pelatihan, sertifikasi, dan penempatan kerja bagi para animator muda di Indonesia, Inkubator Bisnis merupakan sebuah organisasi yang fokus dalam membantu perintis startup atau bisnis agar bisa berkembang dengan cepat dan menjadi sukses. BIKIN *Makerspace* adalah sebuah organisasi yang menyediakan fasilitas bagi pengrajin kriya, dimana para pengrajin dapat bereksperimen, berkarya, dan saling berkolaborasi. Lembaga Sertifikasi Profesi (LSP) merupakan sebuah lembaga yang melaksanakan kegiatan sertifikasi kepada tenaga profesional, dimana sertifikat profesi memperoleh lisensi dari Badan Nasional Sertifikasi Profesi (BNSP) berdasarkan Standar Kompetensi Kerja Nasional Indonesia (SKKNI) bidang animasi Nomor: 400 tahun 2014. *E-Learning* adalah wadah untuk mempelajari modul-modul yang diakses melalui web, dimana modul – modul tersebut terkait dengan ujian sertifikasi pada LSP.

BDI Denpasar memiliki layanan bernama “Kabar Insan Oke” di dalam websitenya sendiri yang dibuat menggunakan *CMS WordPress*, dimana layanan ini baru didirikan pada tahun 2021. Kabar Insan Oke berisikan informasi berita-berita tentang kegiatan dari BDI Denpasar. Berdasarkan hasil wawancara yang dilakukan, BDI Denpasar menginputkan berita sekitar 20 berita setiap bulannya.

Pengkategorisasian berita dilakukan oleh BDI Denpasar agar berita lebih tertata dan terstruktur sesuai dengan bidang beritanya dan juga agar lebih mudah dalam mencari sebuah informasi berita. Dalam pengelolaan kelas berita di sistem layanan Kabar Insan Oke ini masih menggunakan cara konvensional seperti memilih kelas berita dengan cara manual yang menyebabkan aktivitas kurang maksimal apalagi berita pasti akan terus dilakukan pembaruan. Dalam sistem Kabar Insan Oke juga terdapat kesalahan input berita, dimana berita yang seharusnya masuk ke kategori Pengumuman tetapi berada pada kategori Diklat 3 in 1. Hal tersebut akan membuat pengunjung website merasa kebingungan mengenai berita yang ditampilkan tidak konsisten. Oleh karena itu diperlukan sebuah metode pengklasifikasian dokumen berita agar dalam mengelola kelas dapat dilakukan secara otomatis.

Menentukan kelas berita didalam sistem dapat dilakukan secara otomatis dengan menggunakan metode-metode klasifikasi. Salah satu metode yang dapat digunakan dalam pengklasifikasian berita secara otomatis yaitu *K-Nearest Neighbors*. Metode ini termasuk kedalam 10 metode *Data Mining* terbaik yang sering digunakan (Wu and Kumar, 2013). Penelitian ini menggunakan metode K-NN dikarenakan pada penelitian yang dilakukan oleh (Indriani, 2020) menganalisa perbandingan metode K-NN dengan NBC, dimana didapatkan hasil bahwa metode K-NN lebih baik tingkat akurasiya daripada metode NBC. Hal ini dibuktikan dengan tingkat akurasi untuk metode K-NN sebesar 80% sedangkan untuk NBC sebesar 73% yang dihitung dengan menggunakan metode *confusion matrix*. Sedangkan penelitian lainnya yang dilakukan (Utami, 2018) melakukan penelitian membandingkan metode klasifikasi *Naïve Bayes* (NB), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (K-NN). Dimana hasil yang

diperoleh yaitu tingkat akurasi dari metode *K-Nearest Neighbor* (K-NN) lebih tinggi sebesar 75,00% dibandingkan dengan metode *Naïve Bayes* (NB) sebesar 71,50% dan *Support Vector Machine* (SVM) sebesar 72,50%.

Melakukan pengklasifikasian dokumen berita dengan algoritma K-NN dapat menggunakan sebuah metode *term weighting* agar dalam proses pengklasifikasian tingkat akurasi yang dihasilkan lebih akurat dan lebih maksimal. Salah satu metode *term weighting* atau pembobotan kata yang digunakan dalam pengklasifikasian yaitu TF-RF (*Term Frequency-Relevance Frequency*) dan TF-ABS (*Term Frequency-Absolute*). TF-RF digunakan dalam pengklasifikasian pada penelitian ini dikarenakan pada penelitian yang dilakukan oleh (Deolika, Kusri dan Luthfi, 2019) menganalisis perbandingan pembobotan kata klasifikasi metode TF-IDF, TF-RF dan WIDF. Berdasarkan analisa pada hasil klasifikasi yang telah dilakukan, dapat disimpulkan bahwa pembobotan TF-RF dengan klasifikasi *Naïve bayes* lebih baik dari pembobotan TF-IDF dan WIDF dengan nilai *Accuracy* 98,67%, *Precision* 93,81%, dan *Recall* 96,67%. Dan TF-ABS digunakan juga berdasarkan penelitian sebelumnya, dimana Penelitian yang dilakukan oleh (Supono, Riza Adrianti dan Muhammad Azis, 2021) menganalisis perbandingan metode TF-ABS dengan TF-IDF pada klasifikasi teks menggunakan K-NN, dimana hasil yang didapatkan menunjukkan bahwa metode pembobotan TF-ABS lebih baik dari TF-IDF dengan nilai akurasi tertinggi 90,04% pada jumlah fitur 15% dan $k=3$. Penelitian ini berupaya meneliti sebuah metode klasifikasi yang menggunakan metode pembobotan antara TF-RF dengan TF-ABS. Analisis diperlukan untuk menentukan perbandingan metode pembobotan mana yang lebih baik digunakan didalam sistem klasifikasi otomatis nantinya dengan cara

menghitung *Recall*, *Precision* dan *Accuracy* menggunakan metode *Confusion Matrix*. Dimana pada metode yang terbaik nantinya yang akan diusulkan digunakan dalam sistem klasifikasi berita pada Balai Diklat Industri Denpasar.

Berdasarkan persoalan yang dibahas sebelumnya, diharapkan dengan dilakukannya penelitian yang berjudul “**Analisis Perbandingan Metode Pembobotan TF-RF dan TF-ABS pada Kategorisasi Dokumen Berita di Balai Diklat Industri Denpasar Menggunakan Klasifikasi K-Nearest Neighbors**” dapat menjadi referensi atau pedoman jika instansi ingin membangun sebuah sistem klasifikasi berita secara otomatis sehingga dapat memaksimalkan kinerja pegawai BDI Denpasar. Penelitian ini melakukan analisis perbandingan metode, dimana hasil dari penelitian ini berupa modul program klasifikasi yang menggunakan metode terbaik antara TF-RF dan TF-ABS yang akan diberikan ke BDI Denpasar.

1.2 Rumusan Masalah

Adapun persoalan yang terdapat pada latar belakang yaitu pengklasifikasian berita pada BDI Denpasar masih dilakukan secara manual sehingga dapat menyebabkan resiko terjadinya kesalahan *input* kelas. Dari permasalahan tersebut, dapat diambil sebuah rumusan masalah yaitu:

1. Bagaimana hasil implementasi klasifikasi dokumen berita menggunakan K-*Nearest Neighbors* dengan metode pembobotan kata terbaik antara TF-RF dan TF-ABS?

2. Bagaimana hasil perbandingan metode pembobotan kata antara TF-RF dengan TF-ABS menggunakan *K-Nearest Neighbors* pada dokumen berita di Balai Diklat Industri Denpasar?

1.3 Tujuan Penelitian

Berikut tujuan dibuatnya penelitian ini berdasarkan rumusan masalah yang didapatkan, yaitu:

1. Untuk mengetahui hasil klasifikasi dokumen berita menggunakan *K-Nearest Neighbors* dengan metode pembobotan kata terbaik antara TF-RF dan TF-ABS.
2. Untuk mengetahui hasil perbandingan metode pembobotan kata TF-RF dan TF-ABS dengan menggunakan *K-Nearest Neighbors* pada dokumen berita di Balai Diklat Industri Denpasar.

1.4 Ruang Lingkup Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, adapun ruang lingkup dari penelitian ini agar pembahasan lebih fokus dan spesifik pada tujuan dari penelitian ini. Ruang lingkup tersebut antara lain sebagai berikut:

1. Analisis ini berfokus pada kategorisasi dokumen berita pada layanan “Kabar Insan Oke” yang ada di website <https://bdiidenpasar.kemenperin.go.id> pada instansi Balai Diklat Industri Denpasar.

2. Kategori dokumen yang digunakan dibagi menjadi 7 yaitu Pengumuman, Agenda, Kabar Insan Oke, Bikin Maker Space, Diklat 3 in 1, Teaching Factory dan Inbis Tohpati.
3. Berita yang digunakan dalam pengklasifikasian menggunakan Bahasa Indonesia.
4. Berita yang digunakan dari rentang waktu bulan Maret 2021 sampai bulan Februari 2022.
5. Hasil penelitian berupa modul klasifikasi berita yang akan diberikan ke BDI Denpasar.

1.5 Manfaat Penelitian

Adapun manfaat yang didapatkan untuk beberapa pihak dari penelitian ini, antara lain sebagai berikut:

1. Memudahkan instansi dalam menginputkan kategori berita sehingga dapat memaksimalkan kinerja pegawai BDI Denpasar
2. Pembaca mengetahui cara menganalisis dan hasil perbandingan metode antara pembobotan kata TF-RF dengan TF-ABS menggunakan algoritma *K-Nearest Neighbors*.
3. Menjadikan penelitian ini sebagai pedoman referensi untuk melengkapi penelitian sejenis yang ingin dibuat oleh peneliti lainnya.
4. Meningkatkan kemampuan berfikir penulis dan menambah wawasan baru terkait penerapan teori yang didapat dari mata kuliah.