

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

KDD (*Knowledge Discover In Database*) ialah keseluruhan proses pengenalan pola, pengetahuan dan informasi tersembunyi dari sekumpulan data besar. Ada beberapa tahapan pada proses KDD yaitu *data, selection, pre-processing/cleaning, transformation, Data Mining, interpretation/evaluation* dan *knowledge* (Buulolo,2019). Menjadi bagian utama dari proses KDD, *Data Mining* didefinisikan sebagai proses menemukan struktur yang menarik dalam data di mana struktur dapat berupa seperangkat aturan, grafik atau jaringan, pohon, satu atau beberapa persamaan (Roiger, 2019). Sebagai salah satu metode klasifikasi untuk menyelesaikan masalah pada bidang data mining, *Algoritma K-Nearest Neighbor* memiliki ciri dengan pendekatan untuk mencari kasus dengan menghitung kedekatan masalah yang baru dengan masalah yang lama serta menggunakan Teknik pembobotan dari sejumlah objek kasus yang ada (Nofriansyah, 2019).

Penggunaan metode *Algoritma K-Nearest Neighbor* dalam pengklasifikasian telah banyak dilakukan oleh peneliti sebelumnya, salah satu penelitian yang telah dilakukan untuk klasifikasi kepuasan pelanggan terhadap pelayanan perusahaan jasa transportasi udara , objek penelitian ini adalah pelanggan yang menggunakan jasa transportasi udara. Dengan jumlah atribut sebanyak 23 dan data berjumlah 129.880 data , dengan nilai akurasi tertinggi senilai 97,28% (Yunitasari,2021).

Metode *Algoritma K-Nearest Neighbor* juga digunakan untuk mengidentifikasi citra jenis bunga dengan ekstraksi warna *HSV* dan tekstur *GLCM*, dikarenakan banyaknya keragaman jenis bunga dengan tingkat kemiripan yang tinggi sehingga sukar untuk dibedakan .Dengan menerapkan *Algoritma K-Nearest Neighbor* dapat membantu mengidentifikasi jenis bunga lebih cepat dengan ciri warna dan tekstur.

Terdapat 10 citra dan 100 *dataset*, dengan akurasi nilai tertinggi sebesar 71% (Salsabila,2021).

Pemilu adalah panggung pertarungan untuk menempatkan kursi jabatan politik di pemerintahan yang berlandaskan pada pilihan legal dari masyarakat yang memenuhi syarat sebagai pemilih Serta merupakan prosedur teristimewa untuk keberlangsungan demokrasi agar rakyat tetap berkuasa atas dirinya sendiri (Hikmadia, 2019). Ada banyak faktor lainnya yang dapat meningkatkan elektabilitas seorang caleg seperti mobilitas dan kampanye aktif (Putra, 2020).

Pada daerah pemilihan di Provinsi Jambi, ada 6 daerah pemilihan yang dikelompokkan menjadi Jambi I sampai dengan Jambi VI. Pada penentuan daerah pilih (dapil), calon legislatif dapat menentukan dapilnya sendiri. Namun kerap pada saat menentukan daerah pemilihan terjadi ketidakselarasan daerah antara calon legislatif dengan daerah pemilihan, seperti latar belakang calon legislatif yang tidak cocok untuk ditempatkan pada daerah pemilihan yang dipilihnya, yang nantinya menjadi kendala pada saat pendekatan ke masyarakat serta meningkatkan potensi kekalahan dengan calon legislatif lainnya. Maka pada penelitian ini data yang digunakan adalah data latar belakang anggota legislatif terpilih sebelumnya yang telah berhasil di daerah pemilihan pilihannya guna dijadikan tolok ukur, nantinya dapat menemukan kesamaan pola saat menentukan daerah pemilihan di Provinsi Jambi. Algoritma yang sesuai untuk penelitian ini adalah *Algoritma K-Nearest Neighbor* (KNN).

Berdasarkan latar belakang permasalahan yang ada, perlu dilakukan klasifikasi terhadap objek dari data pembelajaran yang jaraknya paling dekat dalam menentukan daerah pilih di daerah pemilihan Provinsi Jambi guna membantu calon legislatif bahkan partai politik untuk menentukan daerah pilih yang tepat sehingga meningkatkan potensi berhasil pada pemilihan umum. Sehingga saya tertarik untuk melakukan penelitian tesis dengan judul: “*Data Mining* dalam Menentukan Potensi Keberhasilan Bakal Calon Legislatif Di Daerah Pemilihan Provinsi Jambi Menggunakan *Algoritma K-Nearest Neighbor* ”.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah di atas, penulis merumuskan masalah sebagai berikut :

1. Bagaimana menentukan potensi keberhasilan bakal calon legislatif pada daerah pemilihan Provinsi Jambi menggunakan *Algoritma K-Nearest Neighbor* ?
2. Bagaimana menerapkan *Algoritma K-Nearest Neighbor* untuk menentukan potensi keberhasilan bakal calon legislatif pada daerah pemilihan Provinsi Jambi?

1.3 Batasan Masalah

Ruang lingkup dari permasalahan penelitian ini bertujuan supaya tidak meluasnya permasalahan mengingat masalahnya yang cukup kompleks. Adapun batasan dari permasalahannya sebagai berikut:

1. Data anggota legislatif yang dijadikan pembanding adalah anggota DPRD Provinsi Jambi terpilih periode 2009-2014 ,2014-2019 dan 2019-2024.
2. Daerah pemilihan di Provinsi Jambi, dengan 6 daerah pemilihan.
3. Pengujian metode *Algoritma K-Nearest Neighbors* pada penelitian ini menggunakan software *Data Mining* pendukung *RapidMiner*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Memahami penerapan *Algoritma K-Nearest Neighbor*
2. Merancang sistem yang dapat menentukan keberhasilan bakal calon legislatif pada pemilihan umum di daerah pemilihan Provinsi Jambi menggunakan *Data Mining* dengan *Algoritma K-Nearest Neighbor*
3. Melakukan Pengujian menggunakan perangkat lunak pendukung *RapidMiner*.
4. Menguji sistem *Data Mining* yang telah dibangun menggunakan data uji.
5. Menganalisa keluaran yang dihasilkan dari sistem *Data Mining*.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Sistem dapat menjadi salah satu cara alternatif dalam membantu bakal calon legislatif bahkan partai politik dalam menentukan daerah pemilihan sehingga dapat meningkatkan potensi keberhasilan pada pemilihan umum.
2. Mengetahui tingkat akurasi dari *Algoritma K-Nearest Neighbor* untuk menentukan potensi keberhasilan bakal calon legislatif pada daerah pemilihan Provinsi Jambi.

1.6 Sistematika Penulisan

Susunan penulisan tesis ini dirangkum dalam sistematika penulisan di bawah ini:

Bab I Pendahuluan

Pada bagian ini dijabarkan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian serta sistematika penyajian penelitian.

Bab II Landasan Teori

Bagian ini mengandung teori-teori yang berkaitan dengan penelitian yang didapat dari berbagai referensi, seperti buku dan jurnal.

Bab III Metodologi Penelitian

Bab ini memaparkan langkah-langkah yang digunakan untuk memecahkan permasalahan dalam menentukan potensi keberhasilan bakal calon legislatif di daerah Pemilihan Provinsi Jambi.

Bab IV Analisis Dan Perancangan

Bab ini menjabarkan cara menganalisa dan melakukan klasifikasi penentuan potensi keberhasilan bakal calon legislatif di daerah pemilihan Provinsi Jambi.

Bab V Implementasi Dan Hasil

Pada bab ini akan diuraikan implementasi dan hasil klasifikasi data yang telah dilakukan dengan menggunakan metode *Algoritma K-Nearest Neighbor* .

Bab VI Kesimpulan Dan Saran

Bagian akhir ini berisi tentang kesimpulan, saran, dan kritik yang berguna untuk pengembangan selanjutnya.