



Potensi Keberhasilan Bakal Calon Legislatif Menggunakan Algoritma *K-Nearest Neighbors*

Cyntia Rivatunisa^{1✉}

¹Independent Researcher

cyntiarvtns@gmail.com

Abstract

A candidate for legislative member is someone who is sent to represent his political party in a legislative part which will be directly elected by the people through general elections. General election is a means of implementation in a democratic system to elect representatives of the people in the people's representative institutions directly, publicly, freely, confidentially, honestly and fairly. Legislative elections are a five-year democratic agenda that is held simultaneously, one of which is to elect members of the Provincial DPRD. Provincial DPRD is a regional people's representative institution that is domiciled as an element of provincial government administration. The purpose of this study is to understand the application of the K-Nearest Neighbor Algorithm, to design a system that can determine the success of legislative candidates in the general election in the Jambi provincial constituency using Data Mining with the K-Nearest Neighbor Algorithm, to test using Rapidminer supporting software, to test the system. Data Mining that has been built uses test data, and analyzes the output generated from the Data Mining system. The appropriate method for this research is Data Mining. Data mining is a data processing method to find patterns in data, the k-nearest neighbor algorithm is one of the classification algorithms on objects based on learning data that is closest to the object. The data used and processed in this research is the data of the elected DPRD members for the previous period, namely the data of the 2009-2014 DPRD members, the 2014-2019 DPRD members' data, and the 2019-2024 DPRD members' data. The results of this study are a system that can provide recommendations for prospective Provincial DPRD candidates for the 2024-2029 period in choosing electoral districts that match the profile and have a greater potential for victory during the General Election based on profile data of elected DPRD members in the previous period.

Keywords: Legislative Candidates, Legislative Elections, Electoral District, Data Mining, K-Nearest Neighbors.

Abstrak

Calon anggota legislatif ialah seseorang yang diutus untuk mewakili partainya dalam lembaga legislatif yang akan dipilih secara langsung oleh rakyat melalui pemilihan umum. Pemilihan umum adalah sarana pelaksanaan dalam sistem demokrasi untuk memilih wakil rakyat di lembaga perwakilan rakyat secara langsung, umum, bebas, rahasia, jujur, dan adil. Pemilu legislatif merupakan agenda demokrasi lima tahunan yang dilaksanakan serempak, salah satunya untuk memilih anggota DPRD Provinsi. DPRD Provinsi merupakan lembaga perwakilan rakyat daerah yang berkedudukan sebagai unsur penyelenggara pemerintahan daerah provinsi. Tujuan dari penelitian ini adalah memahami penerapan *Algoritma K-Nearest Neighbor*, 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*, melakukan pengujian menggunakan perangkat lunak pendukung *Rapidminer*, menguji sistem *Data Mining* yang telah dibangun menggunakan data uji, dan menganalisa keluaran yang dihasilkan dari sistem *Data Mining*. Metode yang sesuai untuk penelitian ini ialah *Data Mining*. *Data mining* ialah sebuah metode pengolahan data guna menemukan pola dalam data, algoritma k-nearest neighbors ialah salah satu algoritma klasifikasi pada objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek. Data yang digunakan dan diolah pada penelitian ini adalah data anggota DPRD terpilih periode sebelumnya yaitu data anggota DPRD 2009-2014, data anggota DPRD 2014-2019, dan data anggota DPRD 2019-2024. Hasil dari penelitian ini berupa sistem yang dapat memberikan rekomendasi untuk bakal calon DPRD Provinsi periode 2024-2029 dalam memilih daerah pemilihan yang sesuai profil dan memiliki potensi kemenangan yang lebih besar saat Pemilihan Umum berdasarkan data profil anggota DPRD terpilih periode sebelumnya.

Kata kunci: Calon Legislatif, Pemilihan Umum Legislatif, Daerah Pemilihan, *Data Mining*, *K-Nearest Neighbors*.

JSISFOTEK is licensed under a Creative Commons 4.0 International License.



1. Pendahuluan

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* [1]. 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 [2].

Sebagai salah satu metode klasifikasi untuk menyelesaikan masalah pada bidang *Data Mining* adalah Algoritma *K-Nearest Neighbor*. Algoritma ini 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 [3].

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% [4]. 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% [5].

Pemilu adalah panggung pertarungan untuk menempatkan kursi jabatan politik di pemerintahan yang berlandaskan pada pilihan legal dari masyarakat yang memenuhi syarat sebagai pemilih [6]. Prosedur istimewa untuk keberlangsungan demokrasi agar rakyat tetap berkuasa atas dirinya sendiri [7]. Ada banyak faktor lainnya yang dapat meningkatkan elektabilitas seorang caleg seperti mobilitas dan kampanye aktif [8].

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 [9]. Di Provinsi Jambi ada 6 daerah yang dikelompokkan menjadi Jambi I sampai dengan Jambi VI [10]. Maka dilakukan penelitian ini untuk mengklasifikasi terhadap objek terhadap jarak paling dekat dalam menentukan daerah pilih di daerah pemilihan Provinsi Jambi untuk membantu calon legislatif lebih efektif dan efisien dalam meningkatkan potensi keberhasilan.

2. Metodologi Penelitian

Kerangka kerja penelitian ini bertujuan untuk menguraikan prosedur sistematis yang akan dilakukan pada penelitian. Kerangka kerja pada penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Kerangka Penelitian

2.1. Mengidentifikasi Masalah

Saat mengidentifikasi masalah, dilakukan pengamatan objek yang ditujukan untuk merumuskan masalah yang ada pada objek penelitian menggunakan metode Algoritma *K-Nearest Neighbor*, sehingga dapat menentukan batasan dari permasalahan yang akan diteliti.

2.2. Menganalisa Masalah

Setelah mengidentifikasi atau mengenali masalah apa saja yang ada pada objek selanjutnya masalah tersebut harus melewati proses Analisa terlebih dahulu, agar permasalahan tersebut dapat lebih mudah dicerna dan dipahami, sehingga nantinya menghasilkan rumusan masalah.

2.3. Mengidentifikasi solusi

Saat hendak mengidentifikasi sebuah solusi, diharuskan untuk mempelajari dan memahami beberapa solusi yang telah ada yang berkaitan dengan penelitian yang diteliti, yang selanjutnya dapat menentukan solusi terbaik yang akan digunakan pada penelitian. Penelitian ini menggali

solusi yang berkaitan menggunakan media internet, dan karya ilmiah tentang Algoritma K-Nearest Neighbor.

2.4. Mengumpulkan Data

pada tahapan ini, peneliti mengumpulkan informasi data-data yang dibutuhkan pada penelitian ini. Dalam metode pengumpulan data peneliti menggunakan beberapa cara.

2.5. Menganalisa menggunakan Algoritma K-Nearest Neighbor

Tahapan-tahapan yang dilakukan untuk mengklasifikasikan data menggunakan KNN adalah sebagai berikut [11]. Menentukan parameter dari jumlah tetangga terdekat atau didefinisikan sebagai koefisien k.

- a. Menghitung jarak besaran jarak antara data uji dan semua data yang terdapat pada data latih. Besaran jarak dihitung menggunakan persamaan Euclidean Distance disajikan pada Rumus (1).

$$c(d, e) = \sqrt{\sum_{i=1}^f (d_i - e_i)^2} \quad (1)$$

Dimana c(d,e) adalah Jarak antara data latih dan data uji, d adalah nilai ciri pada data latih, di mana d = d_1, d_2,..., d_n. e ialah nilai ciri pada data uji, di mana e adalah e_1, e_2,..., e_n. dan f adalah jumlah data yang dihitung

- b. Nilai jarak yang telah dihitung kemudian diurutkan dan ditentukan nilai dari tetangga terdekat yang memiliki jarak terkecil berdasarkan nilai koefisien k yang ditentukan pada tahap pertama.
- c. Menentukan kategori dari nilai tetangga terdekat Menggunakan kategori mayoritas sederhana dari tetangga terdekat sebagai nilai prediksi dari data baru.

2.6. Pengujian Hasil

Pada tahap pengujian hasil, peneliti melakukan perbandingan antara hasil perhitungan manual dengan hasil perhitungan yang dilakukan oleh Rapid Miner untuk mengetahui kesesuaian sistem yang dibuat dengan tujuan yang diharapkan. Pengujian dilakukan dengan cara sebagai berikut:

- a. Menentukan jumlah cluster, klasifikasi, dan nilai bobot yang dibutuhkan
- b. Mentransformasi nilai atribut dan nilai masing-masing variable
- c. Melakukan proses perhitungan manual menggunakan excel
- d. Menjalankan perangkat lunak Rapid Miner menggunakan data latih berupa excel yang sudah disediakan.

- e. Membandingkan hasil perhitungan antara hasil perhitungan menggunakan Rapid Miner dengan hasil perhitungan manual.

2.7. Hasil Penelitian

Pada tahapan hasil dan pembahasan ini akan menjabarkan keluaran akhir yang dihasilkan dari sistem mulai dari tahap penambangan data, perhitungan Data Mining menggunakan Algoritma K-Nearest Neighbor hingga dapat memberikan advice dalam memilih daerah pemilihan yang sesuai untuk bakal calon legislatif berdasarkan latar belakang anggota DPRD Provinsi terpilih periode sebelumnya.

2 Hasil dan Pembahasan

Pada penelitian ini menggunakan 12 variabel data yaitu nama, tempat lahir, tempat organisasi, tempat pendidikan, domisili, tempat berkarir, petahana, total tahun berkarir di legislatif, jabatan tertinggi, pendidikan terakhir, total tahun organisasi, total tahun berkarir, perjalanan karir, pengalaman organisasi, tokoh masyarakat dan daerah pemilihan. Pada variabel tempat lahir, tempat organisasi, tempat Pendidikan, tempat berkarir dan domisili memiliki atribut yang sama yaitu Kab. Batanghari, Kab. Bungo, Kab. Kerinci, Kab. Merangin, Kab. Muaro Jambi, Kab. Sarolangun, Kab. Tanjung Jabung Barat, Kab. Tanjung Jabung Timur, Kab. Tebo, Kota Jambi, Kota Sungai Penuh, luar Provinsi Jambi Nasional, luar Provinsi Jambi luar negeri. yang dimaksud dengan luar jambi nasional pada atribut ini ialah seluruh wilayah di negara Indonesia selain Provinsi Jambi akan di asumsikan menjadi luar provinsi jambi nasional. Begitu juga dengan maksud dari atribut luar Jambi luar negeri yaitu seluruh wilayah di dunia selain negara Indonesia akan di asumsikan menjadi luar jambi luar negeri.

Untuk variabel jabatan tertinggi terdapat 7 atribut yaitu ketua/dewan, wakil ketua, sekretaris, wakil sekretaris, bendahara, wakil bendahara dan anggota. Sama halnya dengan variable petahana/ incumbent dengan variabel tokoh masyarakat hanya ada 2 yaitu ya dan tidak. Variabel total tahun berkarir, total tahun berkarir di legislatif, total tahun ber organisasi, pengalaman organisasi dan perjalanan karir juga memiliki atribut yang sama yaitu tidak punya, 1, 2, 3, 4, 5, 6, 7, 8, 9, dan >= 10. Makna atribut terakhir adalah untuk semua angka yang lebih atau sama dengan 10 akan di kategorikan menjadi atribut >=10. Atribut untuk variable pendidikan terakhir ialah S3, S2, S1, D3, dan SMA. Atribut pendidikan disajikan pada Tabel 1.

Tabel 1. Pembobotan Pendidikan

Atribut	Bobot
S3	5
S2	4
S1	3
D3	2
SMA	1

Pada penelitian ini peneliti membuat klasifikasi daerah menjadi 6 cluster yaitu jambi 1 (kota jambi), jambi 2 (kab. Batanghari dan kab. Muaro jambi), jambi 3 (kab. Merangin dan kab. Sarolangun), jambi 4 (kab. Kerinci dan kota sungai penuh), jambi 5 (kab. Bungo dan kab. Tebo) dan jambi 6 (kab. Tanjung jabung barat dan kab. Tanjung jabung timur).

Selanjutnya dilakukan transformasi jabatan yang datanya disajikan pada Tabel 2, patahan pada Tabel 3, lama tahun berkarir pada Tabel 4, banyak organisasi yang diikuti pada Tabel 5, banyak tempat berkarir pada Tabel 6 dan klaster daerah pemilihan pada Tabel 7.

Tabel 2. Transformasi Variabel Jabatan

Atribut	Bobot
Ketua/Dewan	7
Wakil Ketua	6
Sekretaris	5
Wakil Sekretaris	4
Bendahara	3
Wakil Bendahara	2
Anggota	1

Tabel 3. Transformasi Petahana

Atribut	Bobot
Ya	2
Tidak	1

Tabel 4. Transformasi Tahun Berkarir

Atribut	Bobot
>=10 Tahun	11
9 Tahun	10
8 Tahun	9
7 Tahun	8
6 Tahun	7
5 Tahun	6
4 Tahun	5
3 Tahun	4
2 Tahun	3
1 Tahun	2
Tidak Punya	1

Tabel 5. Banyak Organisasi

Atribut	Bobot
>=10 Karir	11
9 Karir	10
8 Karir	9
7 Karir	8
6 Karir	7
5 Karir	6
4 Karir	5
3 Karir	4
2 Karir	3
1 Karir	2
Tidak Punya	1

Tabel 6. Daerah

Atribut	Bobot
Kab. Batanghari	1
Kab. Bungo	2
Kab. Kerinci	3
Kab. Merangin	4
Kab. Muaro Jambi	5
Kab. Sarolangun	6
Kab. Tanjung Jabung Barat	7
Kab. Tanjung Jabung Timur	8
Kab. Tebo	9
Kota Jambi	10

Kota Sungai Penuh	11
Provinsi Jambi	12
Luar Provinsi Jambi Nasional	13
Luar Provinsi Jambi Luar Negeri	14

Tabel 7. Klasifikasi Cluster

Daerah Pemilihan	Klasifikasi	Cluster
Kota Jambi	Jambi 1	C1
Kab. Batanghari - Kab. Muaro Jambi	Jambi 2	C2
Kab. Merangin - Kab. Sarolangun	Jambi 3	C3
Kab. Kerinci - Kota Sungai Penuh	Jambi 4	C4
Kab. Bungo - Kab. Tebo	Jambi 5	C5
Kab. Tanjung Jabung Barat dan Timur	Jambi 6	C6

Daerah pemilihan yang sesuai berdasarkan profil data anggota Legislatif 3 dengan 135 data latih dan 20 data uji dengan jumlah cluster 6 yang diolah dengan Algoritma K-nearest Neighbor didapatkan, yaitu:

Nama: Fahrurrozi, A.Md; tempat lahir: Kabupaten Sarolangun; tempat organisasi Kabupaten Sarolangun; tempat pendidikan di Sumatra Barat, domisili kabupaten Sarolangun; tidak petahana; total tahun berkarir di legislatif belum punya; jabatan tertinggi Wakil Ketua; pendidikan terakhir S1; total lama berorganisasi 10 tahun; total berkarir non legislatif 12 tahun; dengan banyak pengalaman 3 organisasi; banyaknya perjalanan karir yaitu 4 karir; dan termasuk dalam tokoh masyarakat/ terkenal. Data ini berdekatan atau kesamaan dengan data latih 1, data latih 2 dan data latih 3 sehingga menghasilkan nilai similaritas disajikan pada Tabel 8.

Tabel 8. Hasil Perhitungan

Kasus	Similarity
Kedekatan Data Latih 1 dengan Data Uji	0,6186
Kedekatan Data Latih 2 dengan Data Uji	0,7201
Kedekatan Data Latih 3 dengan Data Uji	0,4514

Berdasarkan Tabel 8 dengan similarity tertinggi terdapat pada Data latih 2 dengan data uji dengan nilai similarity 0.7201, sehingga direkomendasikan daerah pemilihan yang sesuai untuk Fahrurrozi, A.Md adalah daerah pemilihan Jambi III (Tiga) yaitu Kabupaten Merangin – Kabupaten Sarolangun.

3 Kesimpulan

Hasil penelitian ini dapat menjadi rekomendasi daerah pemilihan bagi bakal calon legislatif. Data masukan berdasarkan 12 yaitu, tempat lahir, tempat berorganisasi, tempat pendidikan, kota domisili, tempat berkarir, petahana, total tahun berkarir di legislatif, jabatan tertinggi, pendidikan terakhir, total tahun berorganisasi, total tahun berkarir, banyaknya pengalaman organisasi, dan banyaknya perjalanan karir. Serta mengacu pada nilai similarity dengan anggota legislatif periode sebelumnya.

Daftar Rujukan

- [1]. Buulolo, E. (2020). *Data Mining Untuk Perguruan Tinggi*. Deepublish. https://books.google.co.id/books?id=-K%5C_SDwAAQBAJ
- [2]. Roiger, R. J. (2017). *Data mining: a tutorial-based primer*.

- Chapman and Hall/CRC.
- [3]. Sivalenka, V., & Bai, A. (2021). Materials Today : Proceedings An analysis on prediction of breast cancer using radius Nearest Neighbor algorithm over other classification algorithms. *Materials Today: Proceedings*, xxx, 10–12. <https://doi.org/10.1016/j.matpr.2021.03.344>
- [4]. Hopipah, H. S., Mayasari, R., Karawang, U. S., Timur, T., & Barat, J. (2021). *Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-Nearest Neighbor (k-NN) dan Naïve Bayes*. 6(1), 99–110.
- [5]. Puteri, A., Salsabila, B., Yunita, R. D., Rozikin, C., Studi, P., Informatika, T., Karawang, U. S., & Karawang, K. (2021). *Identifikasi Citra Jenis Bunga menggunakan Algoritma KNN dengan Ekstrasi Warna HSV dan Tekstur GLCM*. 6(1), 124–137
- [6]. Haj, A. S. A., & Amrizal, V. (2020). *Analisis Sentimen Kinerja KPU Pemilu 2019 Menggunakan Algoritma K-Means Dengan Algoritma Confix Stripping Stemmer Sentiment Analysis KPU Performance 2019 Elections Using K-Means Algorithm with Confix Stripping Stemmer*. 2(01), 9–18.
- [7]. Hikmania, Y. S. (2018). Pergulatan Konsep Contested Election dan Uncontested Election (Studi Kasus Pilkada Kabupaten Pati Tahun 2017) (Doctoral dissertation, Tesis. Tidak Diterbitkan. Fakultas Ilmu Sosial dan Ilmu Politik. Yogyakarta: Universitas Gadjah Mada).
- [8]. Mariani, A. (2003). *Pesohor Sebagai Komoditas Media (Analisis Ekonomi-Politik Terhadap Tabloid dan Infotainment di Televisi)*. *Jurnal Penelitian Edisi Ilmu-ilmu Sosial*, 15(2), 566-575.
- [9]. Ubaid, A. H. (2020). *Selebritis Menjadi Politisi: Studi tentang Bagaimana Selebritis Menang atau Kalah dalam Pemilu Legislatif*. 5. <https://doi.org/10.14710/jiip.v5i1.7311>
- [10]. Femiliona, F. (2020). *Usaha KPU Mempertahankan Prinsip Transparansi pada Proses Pencalonan melalui Sistem Informasi Pencalonan (SILON) Abstrak PENDAHULUAN KPU sebagai EMB (Electoral Management Body) atau penyelenggara pemilu di Indonesia mempunyai wewenang dan fungsi utama*. 2(2), 277–320.
- [11]. Nurmahaludin, & Cahyono, G. R. (2019). *Klasifikasi Kualitas Air PDAM Menggunakan Algoritma KNN Dan K-Means. Prosiding SNRT*, 5662(November), 1–7.