

Sistem Pendukung Keputusan Pemilihan Minat Calon Lulusan SMK Menggunakan Metode *K-Nearest Neighbor* (Studi Kasus: SMK Pembangunan Bogor)

Hilda Cinthyariani¹, Herfina², Boldson Herdianto³
^{1,2,3} Program Studi Ilmu Komputer FMIPA Universitas Pakuan Bogor
e-mail: hildacinthyaa@gmail.com

Abstrak—Sistem pendukung keputusan pemilihan minat bagi calon lulusan SMK merupakan salah satu sistem informasi yang dikembangkan untuk membantu calon lulusan SMK dalam menentukan pilihan akan melakukan apa setelah lulus dari SMK nanti. Sistem ini menggunakan sebuah metode klasifikasi yaitu metode *K-Nearest Neighbor* sebagai penentu hasil perhitungan kuesioner sehingga didapat hasil yang akan membantu dalam mengambil sebuah keputusan. Metode *K-Nearest Neighbor* ini menggunakan prinsip kerja mencari nilai atau jarak terdekat diantara banyak data yang memiliki banyak kemiripan namun tetap memiliki perbedaan tertentu. Perhitungan yang digunakan dalam metode ini adalah perhitungan *euclidean distance* yang akan menghitung jarak masing-masing variabel pada data latih terhadap data uji. Data latih yang digunakan dalam sistem ini adalah data yang didapat dari kuesioner yang disebar kepada para calon lulusan SMK, dan di dalam kuesioner itu terdapat banyak pertanyaan yang terbagi ke dalam lima kriteria yang didapat dari hasil wawancara dengan Guru BK SMK Pembangunan Bogor, yaitu kriteria ekonomi, lingkungan sosial, cita-cita, kemampuan akademik, dan minat. Dari hasil pengujian data kuesioner yang didapat hasil perhitungan keakurasiannya mencapai 76,67% pada data latih siswa. Kemudian dilakukan juga perhitungan menggunakan data latih alumni yang hasil perhitungan keakurasiannya mencapai 66,67%.

Kata Kunci—Euclidean Distance, K-Nearest Neighbor, Kuesioner, Pemilihan Minat

I. PENDAHULUAN

Keputusan yang diambil oleh siswa calon lulusan SMK merupakan salah satu keputusan yang cukup sulit diambil karena berbagai faktor yang mempengaruhinya. Tidak semua lulusan SMK dapat langsung bekerja, dan banyak diantara mereka yang ingin melanjutkan jenjang pendidikan ke perguruan tinggi namun belum tentu dapat diwujudkan karena banyak faktor dan pertimbangan [1]. Faktor dan pertimbangan ini tidak hanya timbul dari internal siswa itu sendiri melainkan juga timbul dari beberapa faktor lainnya. Para calon lulusan SMK ini perlu diberikan arahan dan dukungan dalam menentukan pilihan setelah lulus, apakah mau melanjutkan pendidikan ke perguruan tinggi, bekerja, berwirausaha, atau kuliah sambil bekerja [2].

Untuk membantu para calon lulusan SMK dalam menentukan pilihannya, dapat dibuat suatu Sistem Pendukung Keputusan dengan menggunakan metode klasifikasi yaitu metode KNN (*K-Nearest Neighbor*) sebagai penentu hasil perhitungan kuesioner sehingga didapat hasil yang akan membantu memberikan sebuah keputusan. Dan di dalam sistem ini digunakanlah beberapa kriteria yang diberikan oleh Guru BK SMK Pembangunan Bogor, yaitu kriteria ekonomi, lingkungan sosial, cita-cita, kemampuan akademik, dan minat [3].

II. METODE PENELITIAN

Metode klasifikasi yang digunakan dalam sistem ini adalah metode *K-Nearest Neighbor*, yang merupakan metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat atau memiliki persamaan ciri paling banyak dengan objek tersebut. Dekat atau jauhnya tetangga biasanya dihitung dengan jarak *Euclidean* [4]. Metode *K-Nearest Neighbor* dilakukan perhitungan kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap data sampel yang diberikan. Dalam perhitungan jarak ada beberapa rumus jarak contohnya seperti *Euclidean distance*, *Square euclidean distance*, dan *Manhattan distance*. Ada banyak cara mengukur jarak antara data baru dengan data lama (*data training*) di antaranya *euclidean distance* dan *manhattan distance* (*city block distance*), yang paling sering digunakan adalah *euclidean distance* [5]. Berikut menunjukkan rumus perhitungan untuk mencari jarak, dengan d adalah jarak [6].

$$d = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots (2)$$

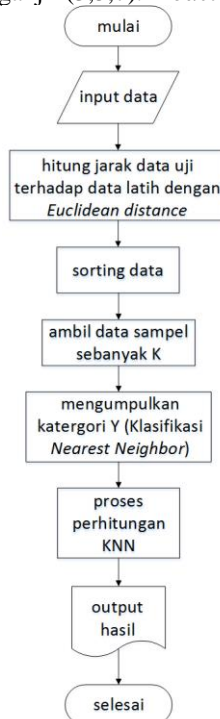
Dengan :

- x_i = Sampel Data
- y_i = Data Uji
- i = Variabel Data
- d = Jarak

Dalam metode K-NN, pendekatan sederhana dalam menentukan nilai K bisa dilakukan dengan cara berikut [7], yaitu :

$$K = \sqrt{N} \dots\dots\dots (3)$$

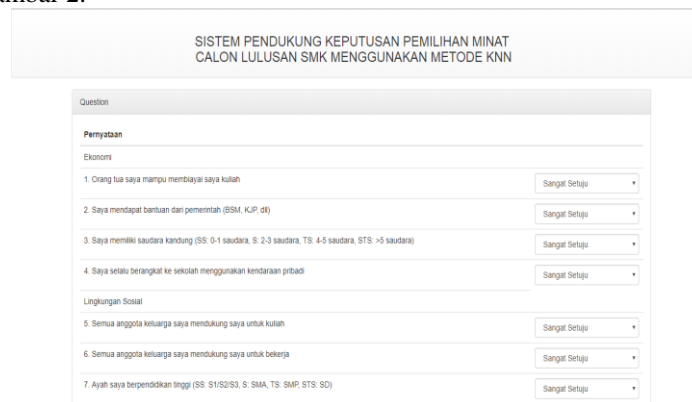
Dalam metode K-NN nilai K biasanya berjumlah ganjil (3,5,7). Model base penerapan metode KNN dilihat pada Gambar 1.



Gambar 1.
Model Base Penerapan Metode KNN

III. HASIL DAN PEMBAHASAN

Hasil pembuatan Sistem Pendukung Keputusan Pemilihan Minat Calon Lulusan SMK Menggunakan Metode *K-Nearest Neighbor* adalah sebuah sistem berbasis *website* yang dibangun dengan menggunakan bahasa pemrograman PHP dan MySQL sebagai *database*. Sistem ini terdiri dari dua jenis *user* yaitu Admin yang memerlukan akses *login* dan *user* umum yang tidak memerlukan akses login, tetapi *user* umum perlu memasukkan data diri terlebih dahulu jika ingin masuk ke sistem dan mengisi kuesioner. Setelah *user* umum berhasil memasukkan data diri, selanjutnya *user* diminta mengisi kuesioner. Tampilan halaman kuesioner dapat dilihat pada Gambar 2.



Gambar 2.
Tampilan *Interface* Kuesioner

Setelah *user* mengisi kuesioner, lalu setiap jawaban dari *user* kemudian disebut sebagai data latih dan akan dihitung menggunakan rumus *euclidean distance* terhadap data uji atau data master yang diperoleh dari Guru BK. Data uji ini yang akan dijadikan acuan untuk menentukan nilai jarak dari data latih. Jumlah pernyataan yang kemudian disebut sebagai jumlah variabel dalam penyebaran kuesioner ini yaitu 26 pernyataan dari 5 kriteria. Kriteria ekonomi memiliki 4 variabel, kriteria lingkungan sosial memiliki 9 variabel, kriteria cita-cita memiliki 5 variabel, kriteria kemampuan akademik memiliki 4 variabel, dan kriteria minat memiliki 4 variabel.

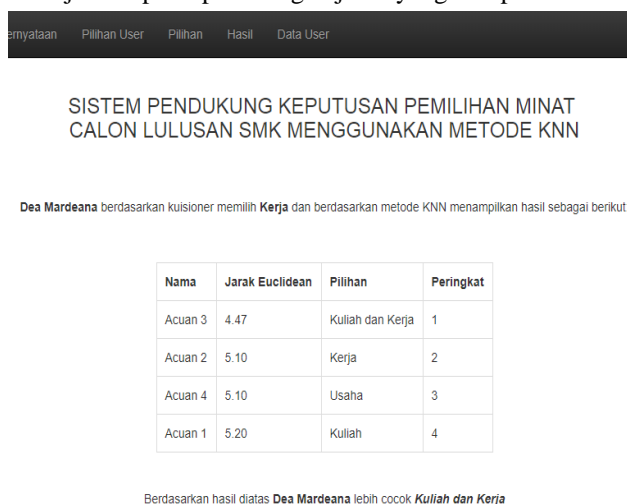
Berdasarkan *model base* perhitungan KNN pada Gambar 1, dijelaskan bagaimana metode KNN bekerja, dimulai dari

memasukkan data hingga didapat hasil pemilihan minat. Data yang dimasukkan merupakan data hasil kuesioner yang dilakukan kepada para pelajar SMK Pembangunan Bogor. Pada kuesioner akan terdapat beberapa pernyataan pada setiap kriteria yang harus dinilai oleh pelajar dengan nilai SS (Sangat Setuju), S (Setuju), TS (Tidak Setuju), STS (Sangat Tidak Setuju). Proses penilaian kuesioner sendiri menggunakan metode *Likert*. Penilaian metode *likert* dapat dilihat pada Tabel 1.

Tabel 1.
 Penilaian Metode *Likert* Berdasarkan Kriteria

No.	Kriteria	Nilai
1.	Ekonomi	Sangat Setuju (4), Setuju (3), Tidak Setuju (2), Sangat Tidak Setuju (1)
2.	Lingkungan Sosial	Sangat Setuju (4), Setuju (3), Tidak Setuju (2), Sangat Tidak Setuju (1)
3.	Cita-cita	Sangat Setuju (4), Setuju (3), Tidak Setuju (2), Sangat Tidak Setuju (1)
4.	Kemampuan Akademik	Sangat Setuju (4), Setuju (3), Tidak Setuju (2), Sangat Tidak Setuju (1)
5.	Minat	Sangat Setuju (4), Setuju (3), Tidak Setuju (2), Sangat Tidak Setuju (1)

Setelah setiap variabel mendapatkan nilai, selanjutnya dihitung menggunakan rumus KNN persamaan (2) untuk mendapatkan nilai jarak bagi data latih terhadap data uji. Tampilan perhitungan jarak yang didapat oleh *user* di *website* pada Gambar 3.



Gambar 3.
 User Interface Hasil Perhitungan User

Berikut adalah perhitungan jarak *Euclidean* dari seluruh data *user* yang mengisi kuesioner pada sistem ini terhadap 4 data uji yang akan disajikan pada Tabel 2.

Tabel 2.
 Hasil Perhitungan 30 Data Latih

Nama	Nilai <i>Euclidean</i>				Nilai Teratas	
	Kuliah	Kerja	Kuliah & Kerja	Usaha	Peringkat 1*	Peringkat 2*
Dea M	5,20	5,10	4,47	5,10	3	2
Octa A	4,90	5,00	4,58	5,00	3	1
Abdullah	5,48	4,12	4,36	4,80	2	3
Ahmad I	7,00	6,32	6,63	4,90	3	2
Saeful B	5,39	5,83	3,74	6,16	3	1
Aris R	6,08	5,66	5,66	4,69	3	2
Maya Y	5,48	4,80	4,36	4,58	3	4
Alan P	4,90	4,36	3,00	5,20	3	2
Rizki	5,48	5,57	4,58	5,57	3	1
Marwanti	5,29	4,80	5,39	4,58	4	2
Julfiansyah	5,48	6,40	5,39	6,24	3	1
Mia W	6,56	6,48	5,83	6,48	3	2
Eka E	5,20	4,69	4,23	5,29	3	2
Sylviawan	4,58	5,66	3,46	5,66	3	1
Aswaludin	5,83	6,71	3,61	6,40	3	1
Rahma R	5,48	5,00	4,58	5,57	3	2
Adinda L	6,00	5,74	4,12	5,92	3	2
Rifa A	6,24	5,83	4,47	5,29	3	4
Dhistya N	4,36	6,48	5,10	6,63	1	3
Nur O	5,66	5,92	4,80	7,00	3	1
Laudia V	4,69	4,12	3,61	5,00	3	2
M Rifki	5,39	4,69	5,48	4,47	4	2

Nama	Nilai <i>Euclidean</i>			Nilai Teratas		
	Kuliah	Kerja	Kuliah & Kerja	Usaha	Peringkat 1*	Peringkat 2*
Neysha F	4,90	4,12	4,58	5,39	2	3
Ria L	3,87	5,83	4,90	5,29	1	3
Edo P	5,48	4,12	4,36	4,80	2	3
Amira S	4,36	6,32	4,69	6,93	1	3
Nur S	5,39	6,00	4,24	5,83	3	1
Galuh G	4,12	4,69	4,90	4,47	1	4
Arya A	3,87	4,24	3,74	5,10	3	1
Ario F	4,47	5,00	4,36	5,74	3	1

* Keterangan :

Pilihan 1 = Kuliah

Pilihan 2 = Kerja

Pilihan 3 = Kuliah dan Kerja

Pilihan 4 = Usaha

Berdasarkan data perhitungan dan setelah diurutkan maka tahap selanjutnya menghitung nilai K untuk menentukan nilai teratas yang akan diambil sebagai hasil keputusan. Nilai K diperoleh dari rumus $K = \sqrt{n}$ dimana nilai n = banyaknya data uji yang dihitung terhadap data latih. Sehingga nilai K dalam kasus ini adalah $K = \sqrt{4} = 2$. Pada Tabel 2 di atas dapat dilihat hasil perhitungan dan nilai *Euclidean* tertinggi yang didapat dari setiap *user*. Karena nilai K pada kasus ini adalah 2, maka diambil 2 nilai teratas dari perhitungan yang dilakukan oleh sistem. Akan tetapi, pada halaman *user*, sistem hanya akan menampilkan 1 pilihan yang nilainya paling tinggi diantara pilihan lain.

Mengapa hasil yang ditampilkan sistem kepada *user* hanya 1 sedangkan nilai K yang didapat dari perhitungan ini adalah 2? Hal ini disebabkan keinginan Guru BK SMK Pembangunan yang ingin menonjolkan hasil paling teratas yang dihasilkan oleh sistem untuk ditunjukkan ke *user*. Dan jika *user* memang ingin mengetahui hasil alternatif dari pengisian kuesioner pada sistem ini, *user* yang merupakan siswa SMK Pembangunan Bogor dapat menghubungi admin dan admin akan memperlihatkan hasil perhitungan sistem keseluruhan.

IV. PENGUJIAN SISTEM

Dari hasil perhitungan jarak *Euclidean* pada masing-masing data latih (data *user*) terhadap masing-masing data uji (data master) yang disajikan pada sub bab pembahasan, kemudian didapatkan hasil perbandingan antara pilihan data latih dengan pilihan yang dihasilkan oleh sistem menggunakan perhitungan KNN, yang dapat dilihat pada Tabel 3.

Tabel 3.
 Pengujian Sistem

No.	Nama	Pilihan	KNN	Hasil
1	Dea M	Kerja	Kuliah & Kerja	Tidak Sesuai
2	Octa A	Kuliah & Kerja	Kuliah & Kerja	Sesuai
3	Abdullah	Kerja	Kerja	Sesuai
4	Ahmad I	Kuliah & Kerja	Kuliah & Kerja	Sesuai
5	Saeful B	Kuliah & Kerja	Kuliah & Kerja	Sesuai
6	Aris R	Kuliah & Kerja	Kuliah & Kerja	Sesuai
7	Maya Y	Kerja	Kuliah & Kerja	Tidak Sesuai
8	Alan P	Kuliah & Kerja	Kuliah & Kerja	Sesuai
9	Rizki	Kuliah & Kerja	Kuliah & Kerja	Sesuai
10	Marwanti	Kerja	Usaha	Tidak Sesuai
11	Julfiansyah	Kuliah & Kerja	Kuliah & Kerja	Sesuai
12	Mia W	Kerja	Kuliah & Kerja	Tidak Sesuai
13	Eka E	Kuliah & Kerja	Kuliah & Kerja	Sesuai
14	Sylviawan	Kuliah & Kerja	Kuliah & Kerja	Sesuai
15	Aswaludin	Kuliah & Kerja	Kuliah & Kerja	Sesuai
16	Rahma R	Kerja	Kuliah & Kerja	Tidak Sesuai
17	Adinda L	Kuliah & Kerja	Kuliah & Kerja	Sesuai
18	Rifa A	Kuliah & Kerja	Kuliah & Kerja	Sesuai
19	Dhistrya N	Kuliah	Kuliah	Sesuai
20	Nur O	Kuliah & Kerja	Kuliah & Kerja	Sesuai
21	Laudia V	Kuliah & Kerja	Kuliah & Kerja	Sesuai
22	M Rifki	Usaha	Usaha	Sesuai
23	Neysha F	Kerja	Kerja	Sesuai
24	Ria L	Kuliah	Kuliah	Sesuai
25	Edo P	Kerja	Kerja	Sesuai
26	Amira S	Kuliah	Kuliah	Sesuai
27	Nur S	Kuliah	Kuliah & Kerja	Tidak Sesuai
28	Galuh G	Kuliah & Kerja	Kuliah	Tidak Sesuai
29	Arya A	Kuliah & Kerja	Kuliah & Kerja	Sesuai
30	Ario F	Kuliah & Kerja	Kuliah & Kerja	Sesuai

Dari hasil pengujian sistem pada Tabel 3 dapat dihitung validitas keakurasian sistem yang dibuat terhadap pilihan yang diambil oleh *user*. Dan dari tabel tersebut didapatkan hasil yaitu dari 30 data yang ada, terdapat 23 data yang sesuai dan 7 data yang tidak sesuai Antara pilihan *user* dengan perhitungan KNN. Maka keakurasian sistem dapat diketahui dengan perhitungan berikut :

$$\text{Nilai Akurasi} = \frac{23}{30} \times 100 = 76,67\%$$

Hasil pengujian sistem dengan menggunakan metode KNN dan perhitungan keakurasian menunjukkan hasil bahwa kinerja sistem mencapai 76,67%, sehingga sistem ini diharapkan dapat dipergunakan oleh Guru BK SMK Pembangunan Bogor untuk membantu siswa dalam mengambil keputusan. Selain pengujian antara nilai pada kuesioner terhadap nilai pada sistem, dilakukan juga pengujian terhadap nilai kuesioner yang dilakukan oleh alumni. Pengujian ini dilakukan untuk menguji apakah sistem dapat memberikan pilihan yang tepat sesuai dengan kenyataan di lapangan atau tidak. Adapun hasil kuesioner terhadap alumni disajikan pada Tabel 4.

Tabel 4.
 Hasil Pengujian Sistem terhadap Data Alumni

No.	Nama	Pilihan	KNN	Hasil
1	Dahlia	Kuliah	Kuliah	Sesuai
2	Ersya	Kuliah	Kuliah	Sesuai
3	M Riskam	Kuliah	Kerja	Tidak Sesuai
4	Widi Dwi	Kerja	Kerja	Sesuai
5	Githa	Kuliah	Kuliah & Kerja	Tidak Sesuai
6	Putri	Kuliah	Kuliah	Sesuai
7	Abdul	Kerja	Kuliah & Kerja	Tidak Sesuai
8	Grace	Kerja	Kerja	Sesuai
9	Surya	Kerja	Kerja	Sesuai
10	Ashilla	Kuliah	Kuliah	Sesuai
11	Guntur	Kerja	Kerja	Sesuai
12	Sella	Kerja	Kuliah	Tidak Sesuai

Dari hasil pengujian sistem terhadap nilai alumni sesuai data *real* alumni pada Tabel 4, maka dapat dihitung nilai keakurasiannya sebagai berikut :

$$\text{Nilai Akurasi} = \frac{8}{12} \times 100 = 66,67\%$$

V. KESIMPULAN

Berdasarkan hasil penerapan metode *K-Nearest Neighbor* pada pembuatan Sistem Pendukung Keputusan Pemilihan Minat Calon Lulusan SMK yang berbasis *website* ini, dapat diambil kesimpulan sebagai berikut :

1. Metode KNN atau *K-Nearest Neighbor* adalah salah satu metode klasifikasi yang dapat digunakan untuk mendapatkan keputusan dalam memilih minat siswa atau calon lulusan SMK Pembangunan Bogor berdasarkan data-data yang diperoleh dari hasil wawancara dengan Guru BK dan pengisian kuesioner oleh siswa-siswi SMK Pembangunan Bogor.
2. Hasil perhitungan keakurasian sistem mencapai 76,67% yang berarti pembuatan sistem menggunakan metode KNN cukup berhasil dan sistem dapat digunakan untuk membantu mengambil keputusan bagi siswa atau *user*.
3. Sistem juga diuji coba dengan data *real* dari alumni, dan setelah dihitung menggunakan metode KNN pada sistem, didapatkan hasil keakurasian mencapai 66,67% antara data alumni dengan hasil keputusan di sistem.

Hasil pilihan yang dihasilkan sistem adalah 2 pilihan karena nilai K dari kasus ini adalah 2, maka hal ini juga dapat membantu memberikan pilihan keputusan alternatif bagi siswa atau user dan bagi Guru BK bila keputusan alternatif pertama tidak dapat diambil.

DAFTAR PUSTAKA

- [1] C. K. Marini and S. Hamidah, "Pengaruh self-efficacy, lingkungan keluarga, dan lingkungan sekolah terhadap minat berwirausaha siswa SMK jasa boga," *J. Pendidik. Vokasi*, vol. 4, no. 2, pp. 195–207, 2014.
- [2] F. Aryani and M. Rais, "Bimbingan Karir Masa Depan untuk Meraih Sukses ke Perguruan Tinggi." Badan Penerbit Universitas Negeri Makassar, 2018.
- [3] M. R. Fahdia, D. Riana, F. Amsury, I. Saputra, and N. Ruhjana, "Komparasi Algoritma Klasifikasi untuk Orientasi Minat Mahasiswa dalam Penuntasan Studi," *JIRA J. Inov. dan Ris. Akad.*, vol. 2, no. 7, pp. 970–1007, 2021.
- [4] E. Prasetyo, *Data mining mengolah data menjadi informasi menggunakan matlab*. Yogyakarta: Andi, 2019.
- [5] M. Bramer, *Principles of Data Mining*. London: Springer, 2007.
- [6] Y. Agusta, "K-means-penerapan, permasalahan dan metode terkait," *J. Sist. dan Inform.*, vol. 3, no. 1, pp. 47–60, 2007.
- [7] N. Hermaduanty and S. Kusumadewi, "Sistem Pendukung Keputusan Berbasis SMS untuk Menentukan Status Gizi dengan Metode K-Nearest Neighbor," 2008.