

LECTURERS ADMISSIONS SELECTIONS MODEL USING FUZZY K-NEAREST NEIGHBOR METHOD

Lely Meilina¹, Nyoman Putra Sastra², Dewa Made Wiharta³

^{1,2,3} Program Studi Magister Teknik Elektro, Fakultas Teknik, Universitas Udayana, Indonesia
Email: ¹lelymeilina452@gmail.com, ²a@unud.ac.id, ³wiharta@unud.ac.id

(Naskah masuk: 9 Desember 2022, Revisi: 15 Desember 2022, Diterbitkan: 23 Maret 2023)

Abstract

Higher Education, or tertiary education, is the final stage which is optional in formal education. It is usually organized in the form of a university, academy, seminary, high school, or institute. Every tertiary institution needs qualified and professional educators because they have an important role in the process of implementing the Tri Dharma of Higher Education. Recruitment for teaching staff usually has several stages and standardization of assessment in selection process. In order for the process of selecting educators to be carried out objectively, a support system is needed to carry out the assessment process. This study applies the Fuzzy K-Nearest Neighbor (FK-NN) method for the classification process in determining prospective educators who pass or not. Data classification is a new data or object grouping into classes or labels based on certain attributes. The application of the FK-NN method has several stages, namely weighting the criteria, then calculating the closeness of the test data and training data, finding the value of k-nearest neighbors between the training data and testing data and determining the membership of each data. Tests were carried out using the Confusion matrix method on several variations of the k value where the highest percentage was obtained from the value of $k = 5$. The test results for all k values obtained an average accuracy rate of 89.22%, 89.22% precision and 82.45% recall with 114 training data and 50 test data. Based on the average value of the test results, it can be concluded that the FK-NN method is feasible and good to use for the selection of educators with the classification of pass or not.

Keywords: Fuzzy K-Nearest Neighbor (FK-NN), Decision Support System (DSS), Data Classification

MODEL SELEKSI PENERIMAAN TENAGA PENDIDIK DENGAN METODE FUZZY K-NEAREST NEIGHBOR

Abstrak

Perguruan Tinggi adalah satuan penyelenggara pendidikan tinggi dengan tenaga pendidik yang biasa disebut dosen dan peserta didik yang biasa disebut mahasiswa. Setiap perguruan tinggi membutuhkan tenaga pendidik yang berkualitas dan profesional karena dalam proses pelaksanaan Tri Darma Perguruan Tinggi memiliki peran yang sangat penting. Seleksi penerimaan tenaga pendidik biasanya memiliki beberapa tahapan dan standar penilaian dalam proses seleksi. Agar proses seleksi penerimaan tenaga pendidik memiliki penilaian secara obyektif, maka dibutuhkan sebuah sistem pendukung keputusan untuk menjalankan proses penilaian. Penelitian ini membuat sebuah model dengan menggunakan algoritma Fuzzy K-Nearest Neighbor (FK-NN) sebagai proses klasifikasi untuk menentukan calon tenaga pendidik dengan hasil lolos atau tidak lolos. Klasifikasi data adalah pengelompokan data atau objek baru ke dalam kelas atau label berdasarkan atribut-atribut tertentu. Penerapan metode Fuzzy K-Nearest Neighbor (FK-NN) memiliki beberapa tahapan yaitu melakukan pembobotan terhadap kriteria, kemudian menghitung kedekatan data *testing* dan data *training*, mencari nilai k tetangga terdekat antara data latih dengan data uji dan menentukan keanggotan dari masing-masing data. Penelitian ini menggunakan pengujian dengan metode *Confusion matrix* terhadap beberapa variasi nilai k dengan $k = 5$ menghasilkan persentase tertinggi. Hasil pengujian terhadap seluruh nilai k memperoleh nilai rata-rata tingkat *accuracy* sebesar 89,22%, *precision* sebesar 89,22% dan *recall* 82,45% dengan 114 data *training* dan sebanyak 50 data uji. Berdasarkan nilai rata-rata dari hasil pengujian, dapat disimpulkan bahwa metode FK-NN layak dan baik digunakan untuk seleksi penerimaan tenaga pendidik dengan klasifikasi lolos atau tidak lolos.

Kata kunci: Fuzzy K-Nearest Neighbor (FK-NN), Sistem Pendukung Keputusan, Klasifikasi Data.

1. PENDAHULUAN

Peningkatan Sumber Daya Manusia (SDM) adalah sebuah hal yang memiliki peran penting untuk menunjang terlaksananya visi dan misi pada sebuah organisasi, perguruan tinggi maupun instansi lainnya. Dalam sebuah perguruan tinggi, seorang tenaga pendidik adalah sumber daya manusia yang sangat berperan dalam proses pelaksanaan Tri Darma Perguruan Tinggi yaitu pendidikan dan penelitian serta pengabdian masyarakat [1]. Setiap perguruan tinggi memiliki standar penilaian masing-masing untuk seleksi penerimaan seorang tenaga pendidik. Penilaian pada saat proses seleksi bertujuan agar setiap perguruan tinggi mendapatkan SDM yang berkualitas dan profesional [2]. Calon tenaga pendidik biasanya akan melewati tahapan-tahapan seleksi yang terdiri dari beberapa kriteria. Dalam hal ini biasanya pihak akademik akan memberikan beberapa kriteria penilaian, diantaranya kualifikasi pendidikan, IPK, tes kemampuan dasar, *micro-teaching* dan wawancara untuk dapat mengetahui kualitas dan kemampuan calon tenaga pendidik. Setelah calon tenaga pendidik melakukan semua tahapan seleksi, maka dibutuhkan sebuah sistem pendukung keputusan untuk menjalankan proses penilaian agar dapat mengurangi tindak kecurangan maupun subyektifitas penilaian sehingga proses akhir berjalan dengan lebih suportif dan efektif [3].

Decision Support System (DSS) adalah suatu sistem yang digunakan untuk mendukung proses pengambilan keputusan. DSS atau yang biasa disebut Sistem Pendukung Keputusan (SPK) merupakan sebuah sistem terstruktur maupun tidak terstruktur yang digunakan untuk memberikan solusi dalam pemecahan permasalahan [4]. Suatu organisasi atau perusahaan biasanya menggunakan SPK sebagai bagian dari sistem informasi untuk membantu perusahaan dalam pengambilan keputusan [5]. Ada berbagai macam metode dari SPK untuk proses seleksi penerimaan tenaga pendidik. Salah satunya adalah algoritma *Fuzzy K-Nearests Neighbor* (FK-NN) yang merupakan sebuah algoritma untuk klasifikasi dari penggabungan antara teknik *Fuzzy* dengan *K-Nearest Neighbor Clasifier* [6]. Metode FK-NN digunakan untuk menghitung sebuah bobot dengan cara mencari nilai tetangga terdekat untuk mendapatkan hasil nilai terbaik.

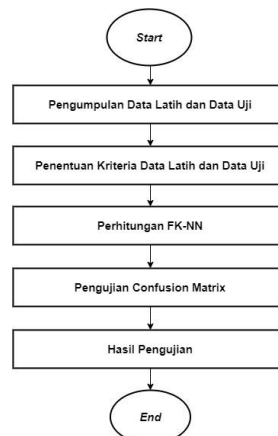
Penelitian yang dilakukan merujuk pada penelitian terdahulu yang menggunakan metode *Fuzzy* dan KNN namun memiliki objek penelitian yang berbeda. Penelitian Jiwandani, dkk [7] dengan menggunakan metode FK-NN, berhasil diimplementasi dan memiliki tingkat akurasi terbaik sebesar 86,785% untuk hasil dianogsis penyakit tanaman kentang. Penelitian berikutnya dilakukan oleh Nanda, dkk [8] yang menggunakan metode FK-NN untuk klasifikasi rating tempat wisata berdasarkan komentar yang didapat dari *social media* dengan hasil persentase sebesar 51,4% dari

nilai rata-rata hasil pengujian tingkat akurasi. Penelitian lain telah dilakukan oleh Nur Inzani, dkk dengan metode yang digunakan adalah *Fuzzy KNearest Neighbor* dan *Discrete Wavelet Transform* untuk mengidentifikasi mood mahasiswa berdasarkan ekspresi wajah. Penelitian ini melakukan tiga ujian pada ekspresi marah, sedih dan senang dengan nilai akurasi rata-rata sebesar 77,49% [9]. Metode *Fuzzy K-NN* pada penelitian terdahulu menghasilkan nilai rata-rata akurasi yang cukup tinggi dari hasil pengujian. Hal ini menunjukkan bahwa metode FK-NN ini layak digunakan sebagai metode pendukung keputusan.

Berdasarkan penelitian terdahulu dan latar belakang permasalahan maka tujuan dari penelitian ini adalah membuat model pendukung keputusan dengan menerapkan metode *FuzzysK-Nearest Neighbor* (FK-NN) pada seleksi penerimaan tenaga pendidik untuk mendukung keputusan pimpinan perguruan tinggi dalam penerimaan tenaga pendidik yang kompeten.

2. METODE PENELITIAN

Sumber data penelitian ini berasal dari artikel jurnal, publikasi konferensi, mesin pencarian seperti Google Scholar, Garuda Ristekbrin, dan IEEE. Pengumpulan data yang digunakan sebagai objek penelitian merupakan data *dummy* yang dibuat dan dirancang oleh penulis untuk dapat disesuaikan dengan keadaan di lapangan. Adapun tahapan proses klasifikasi pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Proses Klasifikasi

Tahap pertama pada penelitian ini adalah mengumpulkan data latih dan data uji. Kemudian menentukan kriteria setiap data yang akan diproses perhitungan. Data *training* dan data *testing* terdiri dari 5 kriteria dalam penilaian seleksi tenaga pendidik. Adapun kriteria penilaian tenaga pendidik yang digunakan dapat dilihat pada Tabel 1. Kemudian dari masing-masing kriteria akan diberikan bobot sebagai acuan proses klasifikasi menggunakan perhitungan FK-NN. Penelitian ini

menggunakan proses pengujian *Confussion Matrix* untuk mencari tingkat akurasi dari klasifikasi menggunakan metode FK-NN. Hasil *output* dari proses klasifikasi yang diharapkan adalah kandidat yang lolos atau tidak lolos berdasarkan nilai yang telah diproses menggunakan metode FK-NN.

Tabel 1. Kriteria Penilaian Tenaga Pendidik

No	Kode	Kriteria
1	K1	Kualifikasi Pendidikan
2	K2	Predikat Kelulusan
3	K3	Tes Kemampuan Dasar
4	K4	Wawancara
5	K5	<i>Micro Teaching</i>

2.1. Pengertian Tenaga Pendidik

Tenaga pendidik adalah tenaga yang bertugas untuk merencanakan, melaksanakan dan menilai proses pembelajaran serta dapat membimbing dan melatih peserta didik secara professional. Tenaga pendidik pada perguruan tinggi harus dapat melakukan penelitian dan pengabdian kepada masyarakat. Menurut UU RI nomor 20 Tahun 2003, tenaga pendidik adalah tenaga kependidikan yang berkualifikasi sebagai guru, dosen, konselor, pamong belajar, widyaiswara, tutor, instruktur, fasilitator dan sebutan lain yang sesuai dengan kekhususannya, serta berpartisipasi dalam penyelenggaraan pendidikan [10].

2.2. Klasifikasi

Klasifikasi adalah penyusunan dan penilaian data secara sistematis untuk dapat dimasukkan ke dalam jumlah kelas yang tersedia. Proses klasifikasi dilakukan dengan cara membuat sebuah model berdasarkan data training yang ada untuk digunakan dalam mengklasifikasikan data yang baru ditambahkan [11]. Klasifikasi harus memiliki target kategori yang dapat dijadikan sebagai variabel *output* [12]. Proses ini bertujuan untuk memprediksi kelas atau kategori dari objek yang digunakan sebagai dataset. Klasifikasi adalah metode data mining yang banyak diterapkan di berbagai bidang dan memiliki banyak algoritma yang dikembangkan [13]. Metode *Fuzzy K-Nearest Neighbor (FKNN)* merupakan salah satu algoritma klasifikasi pada data mining.

2.3. Fuzzy Logic

Fuzzy Logic atau Logika Fuzzy adalah sebuah algoritma yang digunakan untuk memetakan ruang *input* kedalam sebuah ruang *output* [14]. Cara kerja *Fuzzy Logic* adalah dengan menggunakan derajat keanggotaan pada sebuah nilai berdasarkan spesifikasi yang telah ditentukan guna mendapatkan hasil yang diinginkan. Logika fuzzy memiliki prinsip yang menghasilkan dua kondisi pernyataan yaitu benar (*true*) atau salah (*false*) seperti struktur logika binary boolean. Nilai yang dihasilkan oleh Logika fuzzy

bukan hanya *true* (1) atau *false* (0) akan tetapi dapat memungkinkan nilai keanggotaan diantara 0 dan 1.

2.4. Metode K-Nearest Neighbors (K-NN)

Algoritma data mining yang digunakan untuk dapat mengklasifikasikan sebuah obyek dengan data pembelajaran yang memiliki jarak terdekat dengan obyek tersebut adalah Metode *K-NearestNeighbor* (K-NN). Metode K-NN bergantung pada nilai *K* yang digunakan sebagai proses seleksi nilai dengan jumlah *record* data terdekat dengan obyek [15]. Langkah pertama yang dilakukan untuk menerapkan algoritma K-NN adalah menentukan data *testing* dan data *training* yang akan digunakan untuk menghitung jarak ke tetangga terdekat setelah menentukan nilai "K" [16]. Proses perhitungan jarak antara data *testing* dan data *training* disebut dengan *Euclidean Distance*. Adapun rumus dari proses tersebut dapat dilihat pada Persamaan (1).

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

dengan:

$d(x, y)$ adalah jarak data *testing* dengan data *training*

n adalah banyaknya data *training*

x_i adalah data *training*

y_i adalah data *testing*

Setelah menentukan nilai *K*, dilakukan proses *Euclidean Distance* pada data *testing* dengan seluruh data *training* menggunakan Persamaan (1). Hasil dari perhitungan jarak tersebut kemudian diurutkan berdasarkan nilai terdekat. Langkah terakhir adalah mengambil kelas data uji berdasarkan kelas terbanyak tetangga terdekat.

2.5. Metode Fuzzy K-Nearest Neighbor (FK-NN)

Fuzzy K-Nearest Neighbor (FK-NN) merupakan sebuah metode klasifikasi penggabungan teknik *Fuzzy* dengan metode K-NN. *Fuzzy K-Nearest Neighbors* merupakan sebuah metode klasifikasi dalam memprediksi data *testing* terhadap masing-masing kelas sesuai dengan nilai derajat keanggotaan [17]. Penelitian ini menggunakan metode FK-NN sebagai penentuan kelas yang memiliki jumlah data dan jarak pada tetangga terdekatnya. Setiap kelas memiliki perbedaan pada nilai derajat keanggotaan terhadap data *testing* dengan interval [0,1]. Langkah awal dari perhitungan FK-NN adalah melakukan proses inialisasi *fuzzy* untuk mengetahui nilai keanggotaan kelas dengan proses menggunakan rumus berikut [13].

$$U_{ij}(x) = \begin{cases} 0.51 + \left(\frac{n_j}{k}\right) \times 0.49, & \text{jika } j = i \\ \left(\frac{n_j}{k}\right) \times 0.49, & \text{jika } j \neq i \end{cases} \quad (2)$$

dengan:

- u_{ij} adalah nilai keanggotaan kelas i pada vektor j
- n_j adalah jumlah anggota kelas j pada dataset K
- k adalah jumlah banyaknya tetangga terdekat
- j adalah kelas target (training/tidak training)

Tahap berikutnya adalah proses *Euclidean Distance* terhadap data uji dengan data *training* menggunakan Persamaan (1). Hasil diurutkan dimulai dari nilai terendah untuk kemudian menentukan nilai “K” pada *record* terdekat. Hitung derajat keanggotaan data baru pada setiap kelas dengan Persamaan (3).

$$u_i(x) = \frac{\sum_{j=1}^k u_{ij}(\|x-x_j\|^{-2/(m-1)})}{\sum_{j=1}^k (\|x-x_j\|^{-2/(m-1)})} \tag{3}$$

dengan:

- $u_i(x)$ adalah nilai keanggotaan pada data x ke kelas i
- k adalah jumlah tetangga terdekat yang digunakan
- $x - x_j$ adalah selisih jarak data x ke data x_j dalam k tetangga terdekat
- m adalah bobot pangkat (*weight exponent*) yang besarnya $m > 1$

2.6. Teknik Pengujian Confussion Matrix

Confussion matrix adalah sebuah metode pengujian untuk menilai tingkat kebenaran proses klasifikasi. Metode pengujian ini terdiri dari nilai *Accuracy*, *Precision* dan *Recall*. Pada penelitian ini menggunakan nilai *Accuracy* yang merupakan tingkat rasio antara jumlah prediksi yang terbukti benar dengan keseluruhan total prediksi. *Accuracy* adalah tingkat akurasi atau ketepatan antara nilai aktual dengan nilai yang diprediksi [18]. *Precision* adalah tingkat ketepatan yang menggambarkan sebuah akurasi dari hasil prediksi oleh model dengan data yang diinginkan. *Recall* adalah persentase keberhasilan dalam mengklasifikasikan data positif dari hasil prediksi. Adapun rumus *Accuracy* dan *Precision* serta *Recall* dari hasil prediksi dapat dilihat pada persamaan berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

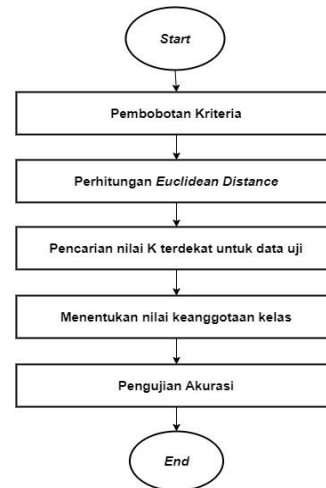
$$Recall = \frac{TP}{TP+FN} \tag{6}$$

dengan:

- TP adalah total data positif yang terdeteksi benar.
- FP adalah total data positif yang terdeteksi sebagai data negatif
- TN adalah total data negatif yang terdeteksi benar.
- FN adalah total data negatif yang terdeteksi sebagai data positif

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data *dummy* dalam melakukan proses klasifikasi. Adapun data yang digunakan adalah calon tenaga pendidik yang terdiri dari *data training* dengan jumlah sebanyak 114 data. Kemudian menggunakan sebanyak 50 data sebagai *data testing*. Setiap data yang digunakan sudah memiliki kriteria yang telah ditentukan pada masing-masing data yang kemudian akan diproses melalui perhitungan yang menggunakan algoritma *Fuzzy KNearest Neighbor* atau FK-NN. Perhitungan algoritma FK-NN memiliki 5 tahapan yang dimulai dari pembobotan kriteria pada masing-masing data yang akan digunakan. Tahap akhir bertujuan untuk mengetahui tingkat akurasi dari perhitungan metode ini. Adapun diagram alir tahapan dari FK-NN dapat dilihat pada Gambar 2.



Gambar 2. Diagram Alir Tahapan FK-NN

3.1. Pembobotan Kriteria

Berdasarkan observasi dan studi literature yang telah dilakukan, dokumen yang dapat dihasilkan berupa kriteria yang digunakan sebagai variabel penelitian untuk dijadikan acuan dalam proses seleksi penerimaan tenaga pendidik. Kemudian dilakukan pembobotan kriteria dengan memberikan nilai bobot secara acak agar dapat digunakan pada setiap sumber data. Nilai acak yang akan digunakan dalam pembobotan memiliki rentang nilai dari 0 hingga 1. Adapun tabel pembobotan kriteria adalah sebagai berikut.

Tabel 2. Pembobotan Kriteria

Kriteria	Nilai	Bobot
K1	S1	0.2
	S2	0.5
	S3	0.8
K2	Memuaskan	0.5
	Sangat Memuaskan	0.7
	Dengan Pujian	0.9
K3	10 – 50	0.3
	51 – 80	0.5
K4	81 – 90	0.8
	10 – 50	0.2

	51 – 80	0.5
	81 - 90	0.8
	10 – 50	0.4
K5	51 – 80	0.7
	81 – 90	0.9

Keterangan:

K1 adalah kriteria dengan kualifikasi pendidikan

K2 adalah kriteria dengan predikat kelulusan

K3 adalah kriteria tes kemampuan dasar

K4 adalah kriteria wawancara

K5 adalah kriteria *micro teaching*

Pembobotan kriteria pada tabel 2 digunakan sebagai acuan untuk memberikan nilai bobot pada data latih dan data uji. Penjelasan terkait kriteria K1 sampai dengan K5 yang digunakan sebagai proses pembobotan dapat dilihat pada Tabel 1. Pada kriteria K3 sampai dengan K5, pembobotan dilakukan berdasarkan rentang hasil skor yang diperoleh oleh calon kandidat. Data latih terdiri dari 5 kriteria calon tenaga pendidik yang telah diberikan nilai bobot dan kelas yang menyatakan lolos atau tidak lolos. Langkah berikutnya adalah memberikan pembobotan kriteria pada data uji yang dapat dilihat pada Tabel 3.

Tabel 3. Sampel Pembobotan Data Uji

Nama	K1	K2	K3	K4	K5
C1	0,5	0,5	0,3	0,2	0,4
C2	0,5	0,7	0,5	0,8	0,6
C3	0,8	0,9	0,8	0,8	0,9
C4	0,2	0,7	0,5	0,2	0,6
C5	0,8	0,5	0,8	0,5	0,9

Keterangan:

C1-C5 adalah inisial dari sampel data testing yang memiliki masing – masing kriteria.

3.2. Perhitungan *Euclidean Distance*

Perhitungan jarak *Euclidean* digunakan pada tahap ini untuk penentuan kedekatan dari data uji dengan data latih. Data kedekatan didapatkan dari penjumlahan selisih dari setiap kriteria dengan masing-masing data latih. Sebagai sampel perhitungan, akan digunakan *data testing* (C1) dan *data training* yang telah dilakukan pembobotan untuk menghitung nilai kedekatan dengan rumus pada Persamaan (1).

$$d(x,y) = \sqrt{(0,5 - 0,5)^2 + (0,9 - 0,5)^2 + (0,5 - 0,3)^2 + (0,5 - 0,2)^2 + (0,6 - 0,4)^2}$$

$$d(x,y) = \sqrt{0 + 0,16 + 0,04 + 0,09 + 0,04}$$

$$= 0,5744563$$

3.3. Pencarian Nilai K Tetangga Terdekat

Hasil dari perhitungan *Euclidean Distance* diurutkan berdasarkan nilai terendah hingga nilai tertinggi. Tabel 4 merupakan sampel perhitungan *euclidean distance* pada data uji C1.

Tabel 4. Sampel Pengurutan *Euclidean Distance*

Alternatif	Euclidean Distance	Klasifikasi
A20	0,2	Tidak Lolos
A25	0,2	Tidak Lolos
A24	0,3	Tidak Lolos
A9	0,458257569	Lolos
A10	0,458257569	Lolos
A11	0,458257569	Lolos
A17	0,469041576	Tidak Lolos
.....
.....
A4	1,053565	Lolos

Keterangan:

A1-A114 adalah inisial dari sampel data latih yang digunakan untuk perhitungan *euclidean distance* pada setiap data uji.

Langkah selanjutnya setelah proses pengurutan nilai *Euclidean Distance* adalah menentukan nilai kedekatan *k*. Dapat diketahui bahwa nilai *k* = 5 sebagai nilai tetangga terdekat. Tabel 4 menunjukkan bahwa data latih atau alternatif ke-20 dengan klasifikasi Tidak Lolos memiliki urutan pertama yang mendekati. Untuk perhitungan selanjutnya digunakan urutan jarak terpendek sesuai dengan nilai *k*, sehingga A10 digunakan sebagai *Euclidean Distance* terakhir.

3.4. Menentukan Nilai Keanggotaan Kelas

Algoritma dari Fuzzy K-NN digunakan untuk proses perhitungan nilai keanggotaan yang telah diujikan pada setiap kelas. Algoritma FK-NN mencari nilai tetangga terdekat sebanyak nilai *k* yang akan dijadikan acuan sebagai perhitungan nilai keanggotaan dari masing-masing kelas. Kemudian kelas akan diberikan bobot yang akan dikalikan dengan nilai *k*. Kelas hasil prediksi merupakan nilai keanggotaan terbesar dari masing-masing kelas. Sampel keanggotaan kelas yang digunakan pada data uji pertama adalah pada C1. Apabila data menampilkan hasil perhitungan yang sama, maka data tersebut akan bernilai 1. Namun apabila data menampilkan hasil yang berbeda akan bernilai 0. Adapun nilai keanggotaan yang dihasilkan dari setiap *k* terhadap kelas lolos dapat dilihat pada Tabel 5.

Tabel 5. Sampel Nilai Keanggotaan Kelas Lolos

Data k ke-	Nilai	Nilai Keanggotaan
1	0	0
2	0	0
3	0	0
4	1	0,458257569
5	1	0,458257569
Total		0,916

Berdasarkan Tabel 5, total nilai keanggotaan pada setiap *k* terhadap sampel data testing C1 adalah 0,916. Langkah berikutnya adalah menghitung sampel nilai keanggotaan setiap *k* terhadap kelas tidak lolos yang dapat dilihat pada Tabel 6.

Tabel 6. Sampel Nilai Keanggotaan Kelas Tidak Lolos

Data k ke-	Nilai	Nilai Keanggotaan
1	1	0,2
2	1	0,2
3	1	0,3
4	0	0
5	0	0
Total		0,7

Berdasarkan Tabel 6, dapat diketahui bahwa total Nilai keanggotaan kelas tidak lolos terhadap data testing C1 adalah sebesar 0,7. Hasil perhitungan nilai keanggotaan menunjukkan bahwa total kelas lolos lebih dari total nilai keanggotaan kelas tidak lolos ($0,916 > 0,7$). Maka C1 dapat diklasifikasikan sebagai anggota dari kelas lolos. Proses klasifikasi nilai keanggotaan ini akan dilakukan pada seluruh data testing yang hasilnya dapat dilihat pada Tabel 7.

7. Sampel Hasil Klasifikasi Data Testing

Nama	Hasil Klasifikasi
C1	Lolos
C2	Lolos
C3	Lolos
C4	Tidak Lolos
C5	Lolos

3.5. Pengujian Akurasi

Penelitian ini menggunakan metode pengujian *Confusion Matrix*. Pengujian ini bertujuan untuk mendapatkan persentase tingkat *accuracy*, *precision* dan *recall*. Data *testing* yang digunakan memiliki jumlah sebanyak 50 data yang terdiri dari kelas Lolos dan Tidak Lolos. Kemudian pengujian dilakukan dengan memasukkan nilai *k* yang bervariasi dengan nilai 3, 5, 10 dan 15. Sampel perhitungan tingkat *accuracy*, *precision* dan *recall* pada masing-masing nilai *k* memiliki *actual* dan *predicted value* yang berbeda-beda. Sebagai sampel perhitungan pengujian akan digunakan nilai *k* = 3 dengan jumlah data positif benar (TP) sebanyak 20, jumlah data negatif benar (TN) sebanyak 25, jumlah data negatif yang terdeteksi positif (FN) sebanyak 2 dan data positif yang terdeteksi negatif (FP) sebanyak 3. Berdasarkan jumlah yang didapatkan dari *actual* dan *predicted value*, maka perhitungan pengujian menggunakan *Confusion Matrix* pada nilai *k* = 3 adalah sebagai berikut:

1. Perhitungan *accuracy* menggunakan rumus pada Persamaan (4).

$$\frac{20+24}{20+22+3+2} \times 100\% = 90\%$$

2. Perhitungan *precision* menggunakan rumus pada Persamaan (5).

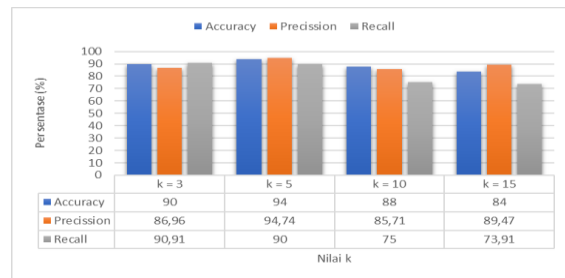
$$\frac{20}{20+3} \times 100\% = 86,95\%$$

3. Perhitungan *recall* menggunakan rumus pada Persamaan (6).

$$\frac{20}{20+2} \times 100\% = 90,90\%$$

Perhitungan pengujian menggunakan rumus diatas kemudian dilakukan terhadap seluruh nilai *k* untuk mengetahui persentase nilai *k* terbaik yang

digunakan dalam proses uji coba. Adapun perbandingan persentase dari seluruh variasi nilai *k* dapat dilihat pada Gambar 3.



Gambar 3. Grafik Perbandingan Setiap Nilai *k*

Berdasarkan grafik perbandingan setiap nilai *k* pada Gambar 3, dapat diketahui nilai *k* yang memiliki persentase tertinggi adalah nilai *k* = 5 dengan hasil tingkat *accuracy* yang dihasilkan adalah 94%, kemudian pengujian *precision* menghasilkan nilai sebesar 94,74% dan nilai pengujian *recall* sebesar 90%. Hasil presentase menunjukkan bahwa masing-masing nilai *k* memiliki pengaruh terhadap tingkat ketepatan yang dihasilkan. Grafik perbandingan pada Gambar 3 menunjukkan bahwa nilai *k* yang semakin tinggi cenderung memiliki tingkat akurasi semakin menurun. Nilai rata-rata pengujian terhadap nilai *k* pada tingkat *accuracy* adalah sebesar 89% sedangkan tingkat *precision* sebesar 89,22% dan *recall* 82,45%.

4. DISKUSI

Beberapa sumber referensi dan data pendukung dibutuhkan dalam melakukan penelitian ini, untuk dijadikan sebagai bahan pengembangan dan perbandingan antara penelitian terdahulu. Adapun referensi penelitian terdahulu yang berkaitan adalah sebagai berikut.

Penelitian terdahulu dilakukan oleh Lestari, dkk yang bertujuan untuk menerapkan metode Fuzzy K-NN dalam memprediksi klasifikasi royalti batubara pada CV. Fazar Utama. Pengujian *sampling* data batubara dilakukan sebanyak 56 kali berdasarkan nilai tetangga terdekat dengan nilai *k* optimal adalah 3. Pengujian akurasi prediksi klasifikasi menghasilkan persentase sebesar 100% dengan 45 data *testing* [19].

Penelitian terdahulu selanjutnya dilakukan oleh Ahmad, dkk bertujuan untuk melakukan prediksi terhadap tingkat kelulusan mahasiswa. Penelitian ini menggunakan 50 sample data latih dan 1 sample data uji. Jumlah *k* atau tetangga terdekat yang digunakan adalah 10. Pemodelan untuk prediksi kelulusan mahasiswa menggunakan metode *Fuzzy KNearest Neighbor* dengan akurasi lulus tepat waktu sebesar 77,35% dan akurasi prediksi lulus terlambat sebesar 63,42% [20].

Penelitian yang dilakukan oleh Chintya, dkk menerapkan metode FK-NN yang memiliki tujuan untuk mendeteksi penyakit hipertensi lebih dini. FK-NN digunakan untuk klasifikasi resiko hipertensi dengan nilai $k = 5$ mendapatkan hasil pengujian tingkat akurasi sebesar 84.0002%. Kemudian menerapkan seleksi fitur *Information Gain* sebanyak 6 fitur yaitu konsumsi kafein, konsumsi makanan berlemak, umur, olahraga, berat badan dan tekanan darah [21].

Berdasarkan uraian beberapa penelitian terdahulu, algoritma FK-NN layak digunakan dalam melakukan sebuah prediksi atau klasifikasi karena memiliki tingkat ketepatan atau akurasi yang cukup tinggi. Sehingga metode FK-NN digunakan pada penelitian ini untuk klasifikasi penerimaan tenaga pendidik. Penelitian ini memiliki beberapa tahapan yaitu melakukan pembobotan terhadap kriteria, kemudian menghitung kedekatan data *testing* dan data *training*, mencari nilai k tetangga terdekat antara data latih dengan data uji dan menentukan keanggotaan dari masing-masing data. Hasil dari penelitian ini diharapkan dapat memberikan solusi kepada pihak perguruan tinggi untuk mengambil keputusan dalam seleksi penerimaan tenaga pendidik yang berkualitas dan profesional.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa metode FK-NN mampu digunakan sebagai acuan untuk menentukan calon tenaga pendidik dengan klasifikasi lolos atau tidak lolos. Data *training* yang digunakan pada penelitian ini berjumlah sebanyak 114 dan data *testing* sebanyak 50. Metode *Confusion Matrix* digunakan sebagai pengujian pada penelitian ini untuk mendapatkan tingkat *accuracy*, *precision* dan *recall*. Pengujian dilakukan terhadap beberapa variasi nilai k dengan $k = 5$ menghasilkan persentase tertinggi dan nilai $k = 15$ menghasilkan persentase terendah. Hal ini menunjukkan bahwa nilai k sangat memiliki pengaruh terhadap tingkat akurasi yang dihasilkan. Pengujian terhadap seluruh nilai k memperoleh nilai rata-rata pengujian yang didapatkan pada tingkat *accuracy* adalah sebesar 89%, tingkat *precision* sebesar 89,22% dan *recall* sebesar 82,45%. Berdasarkan nilai rata-rata yang didapat dari hasil pengujian, metode FK-NN baik digunakan untuk seleksi penerimaan tenaga pendidik.

DAFTAR PUSTAKA

- [1] N. Lestari, J. Karman, and B. Santoso, "Komparasi Metode Weighted Aggregated Sum Product Assessment (WASPAS) dan Multi-Objective Optimization on The Basis of Ratio (MOORA) Dalam Penerimaan Dosen," 2021.
- [2] M. Sadikin, R. Rosnelly, and T. Surya Gunawan, "Perbandingan Tingkat Akurasi Klasifikasi Penerimaan Dosen Tetap Menggunakan Metode Naive Bayes Classifier dan C4.5," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, pp. 1100–1109, 2020, doi: 10.30865/mib.v4i4.2434.
- [3] Kharisma Lalu Puji Indra, "Sistem Pendukung Keputusan untuk Seleksi Penerimaan Dosen menggunakan Metode AHP dan SAW," *JTIM: Jurnal Teknologi Informasi dan Multimedia*, vol. 1, no. 2, pp. 160–165, 2019.
- [4] A. Pradana and E. Bu'ulolo, "Sistem Pendukung Keputusan Dalam Menilai Kinerja Tenaga Pendidikan Terbaik Menggunakan Metode WASPAS," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 5, no. 1, 2021, doi: 10.30865/komik.v5i1.3666.
- [5] O. S. K. Bancin, "Sistem Pendukung Keputusan Pemilihan Kinerja Karyawan Terbaik Menggunakan Metode Simple Additive Weight," *Jurnal Teknik, Komputer, Agroteknologi Dan Sains*, vol. 1, no. 1, pp. 1–9, May 2022, doi: 10.56248/marostek.v1i1.7.
- [6] W. G. Akbari, N. Hidayat, and N. Santoso, "Diagnosis Penyakit Cabai Menggunakan Metode Fuzzy K-Nearest Neighbor (FKNN)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 1, pp. 1070–1074, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [7] J. Andromeda, N. Hidayat, and R. K. Dewi, "Implementasi Metode FKNN (Fuzzy K-Nearest Neighbor) Untuk Diagnosis Penyakit Tanaman Kentang," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 12, pp. 10900–10905, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [8] N. A. Kartini, F. A. Bachtiar, and Indriati, "Klasifikasi Rating Berdasarkan Komentar Tempat Wisata Di Media Sosial Dengan Menggunakan Metode Fuzzy K-Nearest Neighbor," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 5, pp. 4164–4169, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] N. I. R. Amalia and J. Yusmah Sari, "Mengidentifikasi Mood Mahasiswa Berdasarkan Ekspresi Wajah dengan Menggunakan Discrete Wavelet Transform dan Fuzzy K-Nearest Neighbor," *ULTIMA Computing*, vol. XI, no. 1, 2019.
- [10] A. Riadin, M. Jailani, and U. Qudsiyah, "OPTIMALISASI KOMPETENSI DAN KINERJA GURU EKONOMI DALAM PEMBELAJARAN DARING PADA MASA PANDEMIC BERBASIS ICT," *JMM (Jurnal*

- Masyarakat Mandiri*), vol. 4, no. 6, pp. 1219–1261, 2020, doi: 10.31764/jmm.v4i6.3119.
- [11] D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung, *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 2, p. 437, Apr. 2020, doi: 10.30865/mib.v4i2.2080.
- [12] A. P. S. Iskandar, K. Arionadi Shobirin, and K. Oka Saputra, “Analysis of Shopping Cart At Drugs Store By Using An Apriori Algorithm,” *International Journal of Engineering and Emerging Technology*, vol. 2, no. 1, 2017.
- [13] D. Fadma Ristianti, “Komparasi Algoritma Klasifikasi pada Data Mining,” *PROCEEDINGS OF THE 1 st STEEEM*, vol. 1, no. 1, pp. 148–156, 2019.
- [14] S. Kusumadewi, *Artificial Intelligence (Teknik & Aplikasinya)*. Yogyakarta: Graha Ilmu, 2003.
- [15] D. N. Aini, B. Oktavianti, M. J. Husain, D. A. Sabillah, and T. Rizaldi, “Seleksi Fitur untuk Prediksi Hasil Produksi Agrikultur pada Algoritma K-Nearest Neighbor (KNN),” *Jurnal Sistem Komputer dan Informatika (JSON) Hal: 140–*, vol. 145, no. 1, 2022, doi: 10.30865/json.v4i1.4813.
- [16] M. Diki Hendriyanto and N. Sari, “Penerapan Algoritma K-Nearest Neighbor Penerapan Algoritma K-Nearest Neighbor dalam Klasifikasi Judul Berita Hoax,” *Jurnal Ilmiah Informatika (JIF)*, vol. 10, no. 2, 2022.
- [17] A. T. Putri, M. Tanzil Furqon, and R. C. Wihandika, “Klasifikasi Standar Produk Baja PT. Krakatau Steel (Persero) Tbk. Berdasarkan Komposisi Kimia dan Sifat Mekanis Baja Menggunakan Fuzzy K-Nearest Neighbor (Fuzzy K-NN),” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 1, pp. 184–189, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [18] A. N. Yuliarina, “COMPARISON OF PREDICTION ANALYSIS OF GOFOOD SERVICE USERS USING THE KNN & NAIVE BAYES ALGORITHM WITH RAPIDMINER SOFTWARE,” *Jurnal Teknik Informatika (JUTIF)*, vol. 3, pp. 847–856, 2022, doi: 10.20884/1.jutif.2022.3.4.294.
- [19] H. P. Lestari, S. Wahyuningsih, F. Deny, and T. Amijaya, “Prediksi Klasifikasi Royalti Batubara Menggunakan Algoritma Fuzzy K-Nearest Neighbor (Studi Kasus: CV.Fazar Utama),” *Jurnal EKSPONENSIAL*, vol. 10, no. 1, 2019.
- [20] I. Ahmad, H. Sulistiyani, H. Saputra, S. Informasi, and F. T. Dan, “Using Fuzzy K-Nearest Neighbor for Predicting University Students Graduation In Teknokrat,” *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, vol. 1, no. 1, pp. 47–52, 2018.
- [21] D. Chintya and B. Darma Setiawan, “Klasifikasi Risiko Hipertensi menggunakan Fuzzy K-Nearest Neighbor,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 11, pp. 10813–10819, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>