

# Klasifikasi Kelayakan Siswa dalam Menentukan Kelas Unggulan Menggunakan Algoritma K-Nearest Neighbor

Diah Nurul Chasanah  
Universitas Buana Perjuangan  
Karawang, Indonesia  
if17.diahchasanah@mhs.ubpkarawang.ac.id

Amril Mutoi Siregar  
Universitas Buana Perjuangan  
Karawang, Indonesia  
amrilmutoi@ubpkarawang.ac.id

Rahmat  
Universitas Buana Perjuangan  
Karawang, Indonesia  
rahmat@ubpkarawang.ac.id

## Abstract—

Kegiatan belajar mengajar menjadi sebuah upaya timbal balik yang dilakukan antara tenaga pendidik dengan peserta didik. Dinamika kelas menjadi salah satu yang mempengaruhi kegiatan belajar mengajar. Banyak peserta yang memiliki kecerdasan rendah mendapat kelas unggulan, hal ini menyebabkan kesulitan dalam mengikuti proses pembelajaran dikarenakan pembentukan kelas yang dilakukan pihak sekolah belum objektif. Data yang digunakan dalam penelitian ini sebanyak 505 data nilai siswa. Penelitian ini bertujuan untuk mengklasifikasi kelayakan siswa dalam menentukan kelas unggulan, menggunakan metode algoritma *K-Nearest Neighbor*. Pembagian data yang digunakan sebesar 80% data training dan 20% data testing. Hasil penelitian dilakukan dengan beberapa cara perhitungan, pada perhitungan manual dengan *tools* Microsoft excel untuk nilai  $K=25$  hasilnya yaitu “Layak”. Sedangkan pada perhitungan RapidMiner dan bahasa python digunakan untuk menghitung nilai akurasi dengan *confusion matrix*. Hasil akurasi yang diperoleh dengan perhitungan RapidMiner pada  $K=25$  nilai akurasi sebesar 95,05%, dan hasil perhitungan bahasa python dengan *tools* google colaboratory pada  $K=25$  nilai akurasi yang diperoleh sebesar 97,00%.

**Kata kunci** — *data mining*, kelas unggulan, klasifikasi, *k-nearest neighbour*

## I. PENDAHULUAN

Pendidikan merupakan suatu proses pembelajaran yang aktif, serta sebagai wadah untuk mengembangkan potensi diri bagi peserta didik dalam menghadapi perkembangan zaman [1]. Dalam dunia Pendidikan, upaya timbal balik yang dilakukan antara tenaga pendidik dengan peserta didik dapat dilihat pada kegiatan belajar mengajar. Banyak hal yang mempengaruhi kegiatan belajar mengajar baik dalam faktor internal maupun eksternal, salah satunya adalah dinamika kelas. Dalam satu kelas biasanya tenaga pendidik melakukan kegiatan belajar mengajar dengan tingkat kecerdasan peserta didik yang berbeda-beda. bagi peserta didik dengan tingkat kecerdasan rendah mereka mengalami kesulitan mengikuti sistem pembelajaran yang diterapkan sekolah. Sebaliknya peserta didik dengan tingkat kecerdasan tinggi akan merasa bosan dengan pembelajaran yang disampaikan di kelas, karena sistem pembelajaran yang diterapkan tidak sesuai kebutuhan. Kelas unggulan merupakan kelas yang dibuat untuk meningkatkan sumber daya manusia yang lebih berkualitas [2]. Pembentukan kelas unggulan saat ini masih kurang objektif, terdapat beberapa peserta didik yang tidak layak ditempatkan pada kelas unggulan begitupun sebaliknya, hal ini dikarenakan seleksi yang kurang tepat. Maka dari itu sekolah perlu mengklasifikasikan kelayakan siswa dalam menentukan kelas unggulan dengan algoritma *K-Nearest Neighbor*.

Masalah klasifikasi kelayakan dengan menggunakan metode *K-Nearest Neighbor* telah banyak yang dilakukan oleh beberapa orang, diantaranya sebagai berikut. Seperti dalam penelitian yang dilakukan oleh Kartika, Santoso dan Sutrisno [3] tentang pengambilan keputusan untuk menentukan siswa berprestasi dengan sistem yang bersifat objektif dan cepat. Selain itu ada juga penelitian Satrian [4] mengenai klasifikasi usia siswa sekolah dalam bermain *game online* yang memberi dampak negatif terhadap keseharian dan pencapaian prestasi akademik. Hal ini dapat membantu para guru di sekolah membimbing siswa yang kecanduan *game*.

Penelitian mengenai klasifikasi juga dilakukan oleh Sudarsono dan Lestari [5] tentang klasifikasi diagnosa tingkat depresi pada mahasiswa akhir. Untuk menghindari buruknya reputasi kampus, pihak kampus dapat melakukan evaluasi untuk meningkatkan kepercayaan terhadap kampus. Hal ini dapat dilakukan dengan mendiagnosa tingkat depresi untuk mengurangi resiko yang terjadi terhadap mahasiswa tingkat akhir. Dengan metode yang sama penelitian Nikmatun dan Waspada [6] tentang klasifikasi masa studi mahasiswa untuk meningkatkan mutu program studi. Penilaian akreditasi sangat berpengaruh besar terhadap kampus, sehingga perlu evaluasi untuk optimasi ketepatan masa studi mahasiswa. Hasil akurasi dari penelitian tersebut yaitu 75.95% dengan memakai atribut mata kuliah pilihan. Penelitian juga dilakukan oleh Umaidah dan Purwanto [7] tentang pencarian optimal untuk siswa berprestasi. Hal ini dapat membantu pihak sekolah untuk mengetahui prestasi siswa. Hasil dari penelitian tersebut mampu memprediksi prestasi siswa dengan jumlah accuracy yaitu 93.63%.

Berdasarkan penelitian [3,4,5,6,7] bahwa algoritma *K-Nearest Neighbor* mampu menyelesaikan beberapa masalah, maka penelitian ini diberi judul klasifikasi kelayakan siswa dalam menentukan kelas unggulan dengan algoritma *K-Nearest Neighbor*. Hal ini agar memudahkan kegiatan belajar mengajar yang dilakukan oleh peserta didik.

II. DATA DAN METODE

A. Klasifikasi

Dalam data mining klasifikasi merupakan suatu teknik yang digunakan untuk menganalisis sekumpulan data yang di klasifikasikan menjadi sebuah data baru ke dalam kelompok yang sejenis. Klasifikasi digunakan untuk pengklasifikasian data uji dari sekumpulan data, hasil dari data tersebut sebelumnya dilakukan analisis terlebih dahulu. Klasifikasi termasuk ke dalam supervised learning, yang merupakan sebuah metode dimana pengetahuan didapat dari label yang terdapat pada set data pelatihan, kemudian data hasil tersebut digunakan untuk pengklasifikasian data uji [6]. Tujuan dari supervised learning yaitu menghubungkan pola-pola dari data yang sudah ada dengan data baru untuk menemukan pola baru dalam data.

Berdasarkan data proses klasifikasi terdiri dari pembelajaran dan klasifikasi. Proses pembelajaran dianalisis dari data training menggunakan algoritma. Selanjutnya, yang digunakan untuk memastikan tingkat akurasi dari proses klasifikasi yaitu data testing. K-Nearest Neighbour merupakan salah satu algoritma yang digunakan dalam pengklasifikasian data, namun ada banyak algoritma yang lainnya.

B. K-Nearest Neighbor

Salah satu metode yang digunakan dalam pengklasifikasian data yaitu algoritma *K-Nearest Neighbor* atau yang biasa disebut algoritma KNN. dibandingkan dengan metode pengklasifikasian lainnya cara kerja algoritma ini lebih sederhana [8]. Salah satu kekurangan dari algoritma K-Nearest Neighbor ini yaitu rendahnya hasil akhir dari nilai akurasi yang didapatkan. Hal ini dikarenakan penggunaan fitur perhitungan klasifikasi yang tidak perlu masuk dalam proses perhitungan. Namun kelebihanannya dari metode klasifikasi ini mudah untuk dipahami.

Berikut langkah-langkah untuk menghitung metode algoritma *K-Nearest Neighbor* untuk mengklasifikasikan data baru yang belum diketahui:

1. Menentukan jumlah k tetangga, jumlah k yang diambil umumnya ganjil karena menghindari jumlah jarak yang sama dalam proses pengklasifikasian.
2. Dari data yang telah diberikan kemudian dilanjutkan dengan menghitung *euclidean distance* dari masing-masing objek.
3. Hasil dari perhitungan jarak kemudian objek-objek tersebut diurutkan dari jarak *euclid* terkecil.
4. Hasil klasifikasi tersebut dikumpulkan sesuai dengan kategori yang ada pada label.
5. Setelah dilakukan perhitungan maka akan muncul hasil prediksi dari kategori yang ditentukan dengan metode *K-Nearest Neighbor*.

Berikut merupakan persamaan untuk Algoritma K-Nearest Neighbor:

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \tag{1}$$

Keterangan :

- $d_{ij}$  : jarak antara objek i dengan j
- $x_{ik}$  : nilai objek i pada variabel ke-k
- $x_{jk}$  : nilai objek j pada variabel ke-k
- $p$  : banyaknya variabel

C. Kelas Unggulan

Peserta didik yang memiliki bakat dan kemampuan lebih dari peserta didik lainnya biasanya terdapat dalam kelas unggulan. Kelas unggulan merupakan kumpulan dari siswa yang cerdas dan memiliki nilai rata-rata yang tinggi seperti nilai rapor [9]. Proses pembelajaran pada peserta didik dapat berpengaruh terhadap pemilihan kelas unggulan yang berkualitas. Pada kelas unggulan peserta didik akan memperoleh pengajaran yang lebih spesifik sehingga potensi yang ada didalam dirinya bisa lebih berkembang. Secara akademik siswa kelas unggulan biasanya lebih dituntut sukses secara akademik dibanding siswa kelas reguler, hal ini bertujuan agar dapat mencapai ketuntasan belajar di atas kelas reguler. Dalam rangka meningkatkan mutu pendidikan dapat dikembangkan dari kelas unggulan.

D. Confusion matrix

Metode perhitungan akurasi dalam data mining dapat dilakukan dengan *confusion matrix*, metode ini dapat menghitung nilai akurasi, presisi dan *recall* dan *f-Measure* [10]. *Confusion matrix* juga merupakan alat yang kegunaannya memiliki fungsi untuk menganalisis apakah klasifikasi tersebut mengenali *tuple* dari kelas yang berbeda dengan baik [11].

Tabel 1 Confusion Matrix

<i>Correct Classification</i>	<i>Classification</i>	
	Positif	Negatif
Positif	TP	TN
Negatif	FP	FN

Keterangan:

TP = *True Positive* yaitu jumlah data dengan nilai sebenarnya positif dan nilai prediksi positif.

FP = *False Positive* yaitu jumlah data dengan nilai sebenarnya negatif dan nilai prediksi positif.

TN = *True Negative* yaitu jumlah data dengan nilai sebenarnya negatif dan nilai prediksi negatif.

FN = *False Negative* yaitu jumlah data dengan nilai sebenarnya positif dan nilai prediksi negatif.

Evaluasi pada confusion matrix dihitung menggunakan akurasi, presisi, dan *recall*, dan *f-Measure* seperti berikut ini:

#### 1. Akurasi

Akurasi dalam klasifikasi merupakan hasil pengujian yang berupa persentase ketepatan. Berikut merupakan perhitungan dari akurasi, yaitu:

$$\text{Akurasi} = \frac{\text{Jumlah prediksi benar}}{\text{Jumlah total prediksi}} \quad (2)$$

#### 2. Presisi

Presisi merupakan prediksi positif yang data sebenarnya juga bersifat positif. Berikut merupakan perhitungan dari presisi, yaitu:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

#### 3. Recall

*Recall* atau *sensitivity* merupakan hasil positif yang diprediksi dengan benar. Berikut merupakan perhitungan dari *recall*, yaitu:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

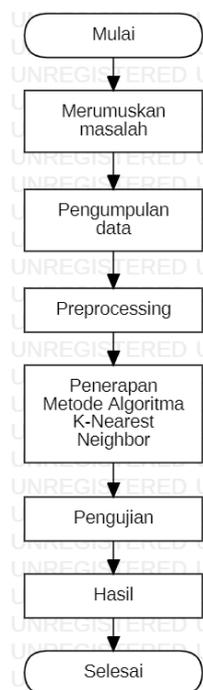
#### 4. F-Measure

Berikut merupakan perhitungan dari *F-Measure*, yaitu:

$$F\text{-Measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5)$$

### E. Gambaran Umum Penelitian

Berikut merupakan gambaran umum dari proses penelitian:



Gambar 1 Gambaran Umum Penelitian

Proses penelitian ini dimulai dari menentukan topik serta merumuskan masalah, dilanjutkan dengan proses pengumpulan data menggunakan studi literatur dan studi lapangan menggunakan teknik wawancara dengan pihak sekolah. Data yang sudah ada kemudian dilakukan pembersihan data meliputi penghapusan data yang tidak digunakan serta perbaikan format data yang kurang tepat, selanjutnya data diimplementasikan dengan algoritma *K-Nearest Neighbor*. Pengujian dilakukan terhadap data latih dan data uji yang sudah ditentukan untuk menentukan hasil perhitungan akurasi dari algoritma *K-Nearest Neighbor*.

F. Pengumpulan Data

Jenis data yang digunakan pada penelitian ini yaitu data kuantitatif. Data kuantitatif merupakan data-data analisis yang bersifat angka yang diolah dengan metode statistika. Data tersebut diperoleh dari nilai siswa SMK Lentera Bangsa Karawang. Data yang digunakan berupa nilai dari empat mata pelajaran utama yaitu nilai Bahasa Indonesia, Bahasa Inggris, Matematika dan Teori Kejuruan. Data nilai ini menggunakan data siswa kelas XI dengan jumlah 505 data. Metode penelitian yang digunakan adalah metode studi literatur dan studi lapangan. Studi literatur yaitu metode pengumpulan data yang diolah dengan cara membaca dan mencatat dari sumber yang berkaitan. Sedangkan, metode pengumpulan data yang dilakukan dengan cara pengamatan langsung pada objek yang diteliti disebut dengan studi lapangan. adapun teknik yang dilakukan adalah wawancara dengan kepala bidang kesiswaan SMK Lentera Bangsa Karawang.

G. Pembersihan Atribut Data

Dari data yang berjumlah 505 data dengan 8 atribut tidak semua atribut ini digunakan. Karena atribut yang tidak digunakan dalam klasifikasi dan tidak memiliki pengaruh terhadap penentuan klasifikasi kelayakan siswa dalam menentukan kelas unggulan. Maka dari atribut yang ada dilakukannya pembersihan. Dari proses pembersihan data atribut yang hanya digunakan yaitu, Nilai Bahasa Indonesia, Matematika, Bahasa Inggris, Kejuruan dan Keterangan yang akan digunakan dalam proses klasifikasi.

Tabel 2 Data yang akan di proses

No.	B.Indonesia	Matematika	B.Inggris	Kejuruan	Ket
1	78	78	70	79	L
2	80	78	70	79	L
3	80	70	81	80	L
4	78	73	76	75	L
5	70	77	77	79	L
·					
·					
·					
504	70	70	72	70	TL
505	70	81	74	70	?

III. HASIL DAN PEMBAHASAN

A. Hasil Implementasi Algoritma

Sebelum melakukan proses perhitungan, terlebih dahulu membagi data menjadi data latih dan data uji. Data latih diambil dari data lama sedangkan data baru diambil dari data ke 505. Setelah pembagian data kemudian menentukan nilai k, untuk nilai k yang akan digunakan dalam penelitian ini sebanyak 5 k yaitu dengan K=15, K=25, K=35, K=45, dan K=55. Berikut merupakan perhitungan algoritma yang merujuk pada persamaan (1) adalah :

$$(d_{ij}) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

$$(d_1) = \sqrt{(78 - 78)^2 + (78 - 70)^2 + (70 - 74)^2 + (79 - 75)^2}$$

$$= \sqrt{(0)^2 + (8)^2 + (-4)^2 + (4)^2}$$

$$= \sqrt{0 + 64 + 16 + 16}$$

$$= \sqrt{96}$$

$$= 9,79795897113271$$

$$(d_{ij}) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

$$(d_2) = \sqrt{(80 - 78)^2 + (78 - 70)^2 + (70 - 74)^2 + (79 - 75)^2}$$

$$= \sqrt{(2)^2 + (8)^2 + (-4)^2 + (4)^2}$$

$$= \sqrt{4 + 64 + 16 + 16}$$

$$= \sqrt{100}$$

$$= 10$$

Perhitungan ini dilakukan sampai dengan data terakhir atau data terakhir atau data ke-504.

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

$$d_{504} = \sqrt{(70 - 78)^2 + (70 - 70)^2 + (72 - 74)^2 + (70 - 75)^2}$$

$$= \sqrt{(-8)^2 + (0)^2 + (-2)^2 + (5)^2}$$

$$= \sqrt{64 + 0 + 4 + 25}$$

$$= \sqrt{93}$$

$$= 9,64365076099295$$

Tabel 3 Perhitungan jarak

No.	B.Indonesia	Matematika	B.Inggris	Kejuruan	Ket	Euclidean Distance
1	78	78	70	79	L	9,79795897113271
2	80	78	70	79	L	10
3	80	70	81	80	L	8,83176086632785
4	78	73	76	75	L	3,60555127546399
⋮						
503	70	70	74	73	TL	8,24621125123532
504	70	70	72	70	TL	9,64365076099295

Setelah jarak dihitung dengan *euclidean distance*, maka selanjutnya data akan diurutkan berdasarkan nilai k. pada Tabel 3 perhitungan diambil dari k=15, maka data akan diurutkan dari 1 sampai dengan 15.

Tabel 4 Hasil perhitungan ranking

No	B.Indonesia	Matematika	B.Inggris	Kejuruan	Ket	Euclidean Distance	Rank
401	78	70	72	73	TL	2,82842712474619	1
296	77	70	74	78	L	3,08474066981327	2
⋮							
37	80	73	76	75	L	4,1821047332653	12
34	78	70	70	73	TL	4,23792402008342	13
307	77	70	73,5	70	TL	4,27383024931969	14
332	78	70	73,5	70	TL	4,27931069215592	15

Perhitungan nilai K=15 hasil yang didapatkan bahwa 8 data memiliki kategori Layak, sedangkan 7 data lainnya memiliki kategori Tidak Layak. Maka kesimpulannya data baru dengan nilai K=15 termasuk dalam kategori Layak. Untuk mengetahui hasil K=25, K=35, K=45, dan K=55 lakukan hal yang sama. Hasil perhitungan klasifikasi terhadap data baru dari nilai k tersebut dapat dilihat pada tabel 4.

Tabel 5 Hasil perhitungan manual

K	15	25	35	45	55
Data Baru	Layak	Layak	Layak	Tidak Layak	Layak

B. Hasil Akurasi dengan RapidMiner dan Python

Data yang digunakan pada perhitungan RapidMiner dan Python ini sebanyak 505 data. Data tersebut dibagi menggunakan *split data* yang dihubungkan dengan operator algoritma *K-Nearest Neighbor* pada *tools* RapidMiner. Data tersebut akan dibagi menjadi 2 bagian, yaitu sebesar 80% sebagai data latih dan 20% data uji. Nilai k yang akan digunakan sama halnya dengan perhitungan manual yaitu K=15, K=25, K=35, K=45, dan K=45. Berikut hasil klasifikasi pada perhitungan nilai k=25 :

Row No.	KET	prediction(K...	confidence(L)	confidence...	NAMA	B.INDONESIA
1	L	L	0.680	0.320	NENENG HO...	70
2	L	L	0.950	0.040	RIRIN AGUS...	80
3	TL	TL	0	1.000	WIDA APRIDA	70
4	L	TL	0.400	0.600	AULIA WULA...	80
5	TL	TL	0.120	0.880	FAUZAN M A...	70
6	TL	TL	0	1.000	AHYAR KASO...	70
7	TL	TL	0	1.000	ANTON FIRMI...	70
8	TL	TL	0	1.000	ANWAR	70
9	L	L	0.950	0.040	ARDIANSYAH...	84
10	L	L	0.950	0.040	ARIF RAHMAT	80
11	TL	TL	0	1.000	DESTI LESTA...	70
12	TL	TL	0	1.000	SARMILA	70
13	TL	TL	0	1.000	SARYA SOBA...	70

ExampleSet (101 examples, 4 special attributes, 5 regular attributes)

Gambar 2 Hasil Klasifikasi pada RapidMiner

Terlihat pada Gambar 4.1 perbandingan antara data awal dengan data hasil klasifikasi yang telah di proses pada RapidMiner yang ditandai dengan kolom warna hijau. Setelah hasil klasifikasi maka hasil akurasi juga dapat diperoleh, seperti pada gambar berikut :

accuracy: 95.05%

	true L	true TL
pred. L	25	1
pred. TL	4	71
class recall	86.21%	98.61%

Gambar 3 Confusion Matrix dengan RapidMiner

Untuk mengetahui nilai akurasi, recall, presisi dan F-Measure dapat dilihat pada penjelasan berikut:

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} \\
 &= \frac{TP+TN}{TP+FP+FN+TN} \\
 &= \frac{25+71}{25+1+4+71} = \frac{96}{101} \\
 &= 0,9504 \times 100\% = 95,05\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Presisi} &= \frac{TP}{TP+FP} \\
 &= \frac{25}{25+1} = \frac{25}{26} = 0,9615 \times 100\% \\
 &= 96,15\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP+FN} \\
 &= \frac{25}{25+4} = \frac{25}{29} = 0,8620 \times 100\% \\
 &= 86,20\%
 \end{aligned}$$

$$\begin{aligned}
 \text{F-Measure} &= 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \\
 &= 2 \times \frac{0,9615 \times 86,20}{0,9615 + 86,20} = \frac{1,6576}{1,8235} \\
 &= 0,9090 \times 100\% = 90,90\%
 \end{aligned}$$

Hasil perhitungan akurasi RapidMiner dengan nilai K=25 diperoleh akurasinya sebesar 95.05%. Tabel 6 merupakan hasil akurasi dari semua nilai k yang digunakan, sebagai berikut :

Tabel 6 Hasil akurasi dari setiap nilai k

K	15	25	35	45	55
Accuracy	95.05%	95.05%	94.06%	49.06%	92.08%

C. Hasil Akurasi dengan Python

Implementasi dengan menggunakan python dilakukan dengan google colaboratory, dengan pembagian data latih dan data uji sebesar 80% dan 20%. Nilai k yang digunakan yaitu K=15, K=25, K=35, K=45, dan K=55. Berikut merupakan hasil klasifikasi pada perhitungan python dengan nilai k=25 :

	Data Asli	Data Prediksi
0	0	0
1	1	1
2	1	1
3	1	1
4	0	0
...	...	...
96	1	1
97	0	0
98	0	0
99	0	1
100	1	1

Gambar 4 Hasil klasifikasi pada python

Gambar 4.3 merupakan hasil dari perhitungan dengan nilai k=25 berupa hasil perbandingan antara data asli dengan data hasil klasifikasi python, label yang bernilai 0 sama dengan “Layak”, sedangkan label dengan nilai 1 sama dengan “Tidak Layak”. Setelah mengetahui hasil klasifikasinya berikut ini adalah hasil dari akurasi perhitungan dengan python pada k=25 :

	precision	recall	f1-score	support
0	1.00	0.89	0.94	28
1	0.96	1.00	0.98	73
accuracy			0.97	101
macro avg	0.98	0.95	0.96	101
weighted avg	0.97	0.97	0.97	101

Gambar 5 Hasil akurasi dengan python pada k=25

Tabel 6 merupakan hasil akurasi dari semua nilai k yang digunakan pada perhitungan python, terlihat hanya k=15 yang memiliki hasil akurasi yang berbeda, sedangkan untuk hasil klasifikasi lainnya memiliki jumlah yang sama yaitu 97.00%.

Tabel 7 Hasil akurasi python

K	15	25	35	45	55
Accuracy	96.00%	97.00%	97.00%	97.00%	97.00%

#### IV. KESIMPULAN

Hasil akurasi tertinggi dengan algoritma *K-Nearest Neighbor* untuk mengklasifikasi kelayakan siswa dalam menentukan kelas unggulan terdapat pada nilai K=25. Hasil perhitungan manual terdapat pada kategori Layak, sedangkan hasil perhitungan dengan RapidMiner pada nilai K=25 diperoleh akurasi sebesar 95.05%, dan hasil dari google colabory pada nilai K=25 dengan nilai akurasi sebesar 97,00%. Hasil akurasi yang tinggi tersebut diharapkan mampu membantu pihak sekolah dalam menentukan kelas unggulan.

#### PENYAKSIAN

Naskah ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Diah Nurul Chasanah dengan judul Klasifikasi Kelayakan Siswa dalam Menentukan Kelas Unggulan Menggunakan Algoritma *K-Nearest Neighbor*, yang dibimbing oleh Amril Mutoi Siregar dan Rahmat.

#### DAFTAR PUSTAKA

[1] N. Sari and S. A. Mulasari, “Pengetahuan, Sikap Dan Pendidikan Dengan Perilaku Pengelolaan Sampah Di Kelurahan Bener Kecamatan Tegalgrejo Yogyakarta,” *J. Med. Respati*, vol. 12, no. 2, pp. 74–84, 2017.

[2] A. R. Z. Wati and S. Trihanto, “Strategi Pengelolaan Kelas Unggulan dalam Meningkatkan Prestasi Belajar Siswa,” *J. Din. Manaj. Pendidik.*, vol. 5, no. 1, pp. 46–57, 2020.

[3] J. I. Kartika, E. Santoso, and Sutrisno, “Penentuan Siswa Berprestasi Menggunakan Metode K-Nearest Neighbor dan Weighted Product (Studi Kasus: SMP Negeri 3 Mejayana),” *J. Pengemb. Teknol. Inf. dan Ilmu Komputer.*, vol. 1, no. 5, pp. 352–360, 2017.

[4] B. Satrian, “Penerapan Algoritma K-Nn untuk Klasifikasi Gamers Usia Sekolah,” *J. Mhs. Apl. Teknol. Komput. dan Inf.*, vol. 2, no. 1, pp. 19–23,

2020.

- [5] B. G. Sudarsono and S. P. Lestari, "Diagnosa Tingkat Depresi Mahasiswa Akhir Terhadap Penelitian Ilmiah Menggunakan Algoritma K-Nearest Neighbor," *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 1094–1099, 2020, doi: 10.30865/mib.v4i4.2448.
- [6] I. A. Nikmatun and I. Waspada, "Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *J. SIMETRIS*, vol. 10, no. 2, pp. 421–432, 2019.
- [7] Y. Umaidah and Purwanto, "Penerapan Algoritma K-Nearest Neighbor (KNN) dengan Pencarian Optimal untuk Prediksi Siswa Berprestasi," *JISICOM (Journal Inf. Syst. Informatics Comput. )*, vol. 3, no. 2, pp. 1–8, 2019.
- [8] G. Harsemadi, M. Sudarma, and N. Pramaita, "Implementasi Algoritma K-Nearest Neighbor pada Perangkat Lunak Pengelompokan Musik untuk Menentukan Suasana Hati," *Maj. Ilm. Teknol. Elektro*, vol. 16, no. 1, pp. 14–20, 2017, doi: 10.24843/mite.1601.03.
- [9] A. S. Kusuma and K. S. Aryati, "Sistem Informasi Akademik Serta Penentuan Kelas Unggulan Dengan Metode Clustering Dengan Algoritma K-Means Di Smp Negeri 3 Ubud," *J. Sist. Inf. dan Komput. Terap. Indones.*, vol. 1, no. 3, pp. 143–152, 2019, doi: 10.33173/jsikti.29.
- [10] V. S. Ginting, K. Kusri, and E. Taufiq, "Implementasi Algoritma C4.5 untuk Memprediksi Keterlambatan Pembayaran Sumbangan Pembangunan Pendidikan Sekolah Menggunakan Python," *Jurnal Teknol. Inf. dan Komun.*, vol. 10, no. 1, pp. 36–44, 2020, doi: 10.35585/inspir.v10i1.2535.
- [11] M. F. Fibrianda and A. Bhawiyuga, "Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3112–3123, 2018.