

APLIKASI PREDIKSI KELULUSAN MAHASISWA BERBASIS *K-NEAREST NEIGHBOR* (K-NN)

Lalu Abd Rahman Hakim¹, Ahmad Ashril Rizal², Dwi Ratnasari³

^{1,2,3}Universitas Bumigora

Email: ¹rahmanlalu11@gmail.com, ²ashril.rizal@gmail.com, ³tsabat.wyk@gmail.com

Abstrak

Mahasiswa adalah aset penting bagi sebuah institusi pendidikan dan untuk itu perlu diperhatikan tingkat kelulusan mahasiswa tepat pada waktunya. Presentasi naik turunnya kemampuan mahasiswa untuk menyelesaikan studi tepat waktu merupakan salah satu elemen penilaian akreditasi kampus. Berdasarkan data dari Bagian Prodi di 3 tahun terakhir presentasi kelulusan mahasiswa hanya 25% mahasiswa dari total yang keseluruhan mahasiswa yang dapat menyelesaikan studinya dengan tepat waktu. Pada penelitian ini menggunakan algoritma K-Nearest Neighbor yang bertujuan untuk dapat mengidentifikasi kelulusan mahasiswa pada kasus baru dengan cara mengadaptasi solusi dari kasus sebelumnya yang memiliki kedekatan dengan kasus baru. Algoritma ini berperan untuk mendapatkan nilai kedekatan kasus baru terhadap kasus lama, yang selanjutnya populasi terbanyak pada area K dengan nilai terdekat didapatkan mahasiswa tersebut di prediksi apakah lulus tepat waktu atau tidak tepat waktu. Penelitian ini menggunakan metode waterfall Roger S. Pressman yaitu Communication, Planning, Modelling, dan Construction. Berdasarkan pengujian yang dilakukan menggunakan K-Fold Cross Validation didapatkan akurasi tertinggi pada model yang ketiga sebesar 80% ketika k-fold ke-4 dan 61% ketika nilai K=1. Sedangkan pengujian menggunakan Confusion Matrix didapatkan akurasi tertinggi sebesar 98% pada K=1 untuk klasifikasi "Tepat Waktu", dan 98% pada K=2 untuk klasifikasi "Tidak Tepat Waktu". (1 baris kosong, 10pt)

Kata kunci: Mahasiswa, *K-Nearest Neighbor*, Kelulusan Mahasiswa, Data Mining

STUDENT GRADUATION PREDICTION USING *K-NEAREST NEIGHBOR* (K-NN)

Abstract

Students are important assets for an educational institution and for this reason, it is necessary to pay attention to the student's graduation rate on time. Presentation of the ups and downs of students' ability to complete their studies on time is one of the elements of campus accreditation assessment. Based on data from the Study Program Section in the last 3 years the student graduation presentation is only 25% of the total students who can complete their studies on time. In this study using the K-Nearest Neighbor algorithm which aims to be able to identify student graduation in new cases by adapting solutions from previous cases that have closeness to new cases. This algorithm has the role to get the value of the closeness of the new case to the old case, which in turn the most population in area K with the closest value obtained by the student is predicted whether to pass on time or not on time. This study uses Roger S. Pressman's waterfall method, namely Communication, Planning, Modeling, and Construction. Based on the tests carried out using K-Fold Cross Validation, the highest accuracy in the third model was 80% when folded 4th and 61% when the K value = 1. While testing using the Confusion Matrix obtained the highest accuracy of 98% at K = 1 for classification "Timely", and 98% at K = 2 for classification "Not Timely". (1 empty line, 10pt)

Keywords: Student, *K-Nearest Neighbor*, Student Graduation, Data Mining

1. PENDAHULUAN

STMIK Bumigora Mataram merupakan sekolah tinggi komputer yang memiliki mimpi di masa depan yang mampu berkompetisi dengan perguruan tinggi baik lulusan, sumber daya dan manajemennya. Selain itu di masa mendatang

STMIK Bumigora akan menjadi perguruan yang bertaraf internasional, khususnya di bidang teknologi informasi. Di samping itu, STMIK Bumigora menjadi salah satu perguruan tinggi yang berperan aktif dalam menjamin hubungan dengan perguruan tinggi lain dalam rangka mengurangi disparitas keilmuan dan sumber daya. Salah satu

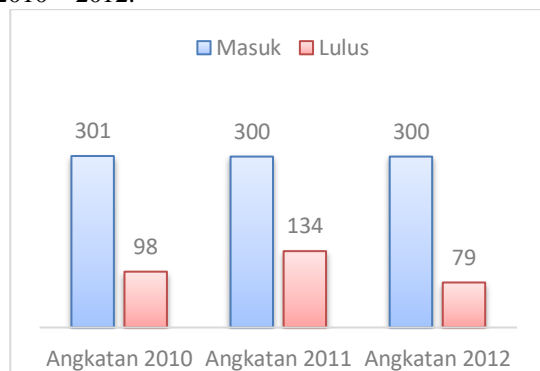
unsur terpenting dalam penyelenggaraan perguruan tinggi adalah mahasiswa [1].

Mahasiswa sering disebut sebagai kelompok masyarakat yang memiliki ciri intelektualitas yang lebih luas dibandingkan dengan kelompok seusia mereka yang bukan mahasiswa ataupun kelompok usia lain yang dibawah mereka. Dengan intelektualitasnya, mahasiswa akan mampu menghadapi dan mencari permasalahan secara sistematis yang nantinya diterapkan dalam kehidupan sehari-hari agar bisa bersaing dalam dunia kerja [2]

Perguruan tinggi dituntut untuk menyelenggarakan pendidikan yang berkualitas bagi mahasiswa sehingga menghasilkan sumber daya manusia yang berilmu, cakap, kreatif dan bersaing. Dalam sistem pendidikan, mahasiswa adalah aset penting bagi sebuah institusi pendidikan dan untuk itu perlu diperhatikan tingkat kelulusan mahasiswa tepat pada waktunya [3].

Presentasi naik turunnya kemampuan mahasiswa untuk menyelesaikan studi tepat waktu merupakan salah satu elemen penilaian akreditasi universitas[4]. Oleh karena itu, perlu adanya pemantuan dan evaluasi terhadap kecenderungan kelulusan mahasiswa lulus tepat waktu atau tidak.

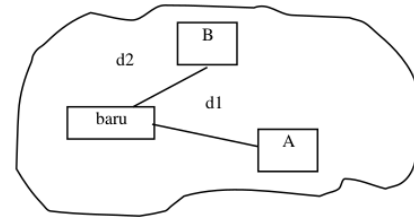
Sistem Kredit Semester (SKS) merupakan salah satu upaya pihak kampus untuk memudahkan mahasiswa dalam menyelesaikan masa studinya tepat waktu (Buku Pedoman Akademik STMIK Bumigora Mataram, 2013). Akan tetapi kenyataan yang terjadi adalah rendahnya tingkat kelulusan mahasiswa setiap tahun. Berikut kami sajikan grafik perbedaan jumlah mahasiswa yang masuk dan yang lulus program studi S1 Teknik Informatika angkatan 2010 – 2012.



Gambar 1. Grafik Perbedaan Jumlah Mahasiswa Masuk dan Lulus 2010 - 2012

Data mining adalah proses menemukan pola-pola didalam data, dimana proses penemuan tersebut dilakukan secara otomatis atau semi otomatis dan pola-pola yang ditemukan harus bermanfaat [5]. Algoritma K-Nearest Neighbor (KNN) adalah merupakan sebuah metode untuk melakukan klasifikasi terhadap obyek baru berdasarkan (K) tetangga terdekatnya. KNN termasuk algoritma supervised learning, dimana hasil dari query instance

yang baru, diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi [6].



Gambar 2. Grafik Perbedaan Jumlah Mahasiswa Masuk dan Lulus 2010 - 2012

Penelitian sebelumnya yang terkait dengan topik ini adalah penelitian yang dilakukan oleh Agus Panoto, Yustina Retno Wahyu Utami dan Wawan Laksito WS dengan judul Penerapan Algoritma K-Nearest Neighbors Untuk Prediksi Kelulusan Mahasiswa Pada STMIK SINAR NUSANTARA SURAKARTA dan penelitian yang dilakukan oleh Abdul Rohman dengan judul Model Algoritma K-Nearest Neighbor (KNN) Untuk Prediksi Kelulusan Mahasiswa.

2. METODOLOGI PENELITIAN

Metode penelitian yang digunakan dalam penulisan ini adalah metode waterfall menurut Roger S. Pressman.

1. Communication

a. Tahap Pengumpulan Data

Wawancara dilakukan kepada bagian Program Studi (Prodi) yaitu Kepala Prodi S1 yaitu Ibu Ni Gusti Ayu Dasriani, M.Kom. Wawancara diperlukan untuk mengetahui permasalahan-permasalahan apa saja yang dihadapi oleh bagian jurusan dan solusi yang dibutuhkan.

Observasi dilakukan untuk mengumpulkan data dengan cara terjun langsung ke lokasi penelitian yaitu STMIK Bumigora Mataram yaitu bagian Program Studi (Prodi), dan Pusat Teknologi Informasi (Pustik). Data-data yang didapatkan dari hasil observasi yaitu data-data kelulusan mahasiswa dan data mahasiswa yang aktif.

Teknik pengumpulan data dengan cara mempelajari, mengkaji dan memahami data yang ada pada berbagai sumber internet, buku, jurnal maupun artikel ilmiah lainnya yang terkait dengan penelitian ini. Data-data yang didapatkan dari studi literatur yaitu teori-teori yang berkaitan dengan penelitian ini seperti teori tentang data mining, algoritma knn dan lain-lain.

b. Tahap Analisa

Setelah melakukan pengumpulan data, maka tahap yang dilakukan selanjutnya adalah menganalisa masalah, masalah yang terdapat dalam penelitian ini yaitu bagian Prodi STMIK Bumigora Mataram masih mendapat kesulitan dalam mengacak

mahasiswa yang berpotensi mengalami keterlambatan kelulusan tepat waktu karena banyaknya mahasiswa yang ada.

Setelah diketahui masalah yang terdapat dalam penelitian ini, maka tahap selanjutnya yaitu menganalisa penyebab dari masalah yang ada. Berdasarkan hasil analisa, yang menjadi penyebab masalah yaitu belum adanya sistem yang dapat memprediksi kelulusan mahasiswa apakah tepat waktu atau tidak.

c. Sistem Lama

Di STMIK Bumigora Mataram pelacakan tetap dilakukan untuk analisa kualitas dan kuantitas lulusan. Analisa yang dilakukan menggunakan metode konvensional dengan melihat jumlah lulusan di tiap tahunnya. Analisa yang lebih kompleks guna mendapatkan informasi yang lebih akurat belum dilakukan. Sehingga alternatif yang diajukan peneliti adalah dengan membuat rancangan desain alternatif untuk menggambarkan proses yang akan terjadi dalam sistem.

d. Desain Sistem Alternatif

Admin melakukan proses login ke aplikasi sesuai dengan username dan password. Admin melakukan penginputan data mahasiswa yang diberikan oleh bagian pustik. Data mahasiswa yang sudah masuk ke dalam sistem akan di normalisasikan. Normalisasi data dapat dilakukan dengan cara yaitu : pembersihan data, integrasi data, seleksi data, dan transformasi data. Proses perhitungan KNN ada beberapa langkah yaitu; menghitung jarak data baru ke data lama, mengurutkan data dari jaraknya yang terkecil ke terbesar, hitung nilai K, hitung perbandingan data yang lebih besar. Setelah diketahui data yang lebih besar, maka ia menjadi hasil prediksi kelulusan apakah tergolong tepat waktu atau tidak. Cetak laporan untuk diserahkan ke pimpinan.

2. Planning

Pada tahap ini dilakukan beberapa kegiatan seperti perencanaan kebutuhan user yang terdiri dari kebutuhan fungsional dan kebutuhan non fungsional. Selanjutnya adalah tahap pembuatan perangkat lunak yang berisi jadwal pembuatan aplikasi yang di rencanakan selesai dalam jangka waktu 109 hari.

3. Modelling

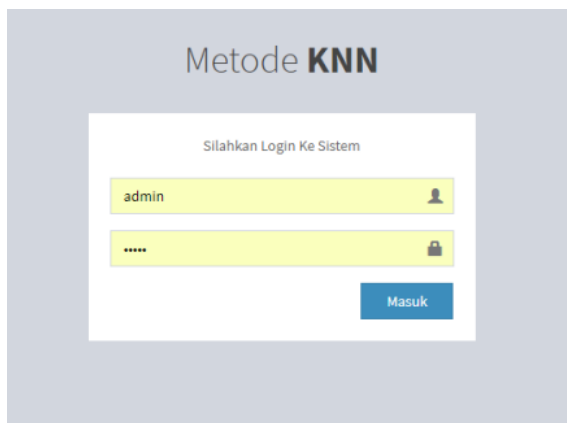
Pada tahap ini dilakukan perancangan basis data yang terdiri tabel data testing, tabel data training, tabel nilai, tabel nilai K, tabel temp, dan tabel data user. Selanjutnya dilakukan perancangan aplikasi dengan melakukan modelling dengan membuat *use case diagram*, *activity diagram*, *class diagram*, *sequence diagram*.

4. Konstruksi

Setelah dilakukan tahap konstruksi, selanjutnya dilakukan tahap konstruksi dengan melakukan beberapa kegiatan seperti pengkodean dengan menggunakan bahasa pemrograman PHP, CSS Bootstrap, dan menggunakan basis data MySQL. Kegiatan selanjutnya adalah pengujian hasil luaran aplikasi dengan menggunakan *confusion matrix* dan *cross validation*.

3. HASIL DAN PEMBAHASAN

3.1. Halaman Admin



Gambar 3. Halaman Login Admin

Halaman Admin merupakan halaman yang berfungsi untuk masuk kedalam sistem dengan menggunakan username dan password yang sudah terdaftar.

3.2. Halaman Dashboard



Gambar 4. Halaman Dashboard

Pada halaman dasborad terdapat beberapa fitur yang ada pada aplikasi ini, yaitu sebagai berikut:

- Dashboard; merupakan halaman untuk mengetahui jumlah keseluruhan data yang realtime dan *link* untuk menuju halaman yang dimaksud.
- Data Testing; merupakan halaman untuk manajemen data testing
- Data Training; merupakan halaman untuk manajemen data training
- Pengujian; merupakan halaman untuk melakukan pengujian data knn

- e. Hasil KNN; merupakan halaman untuk melihat hasil prediksi
- f. Data Pengguna; merupakan halaman untuk memanejemen data penggunaAdmin; merupakan halaman untuk memanejemen admin

3.3. Halaman Data Testing

ID	NIM	Nama Mahasiswa	JK	IP Band 1	IP Band 2	IP Band 3	IP Band 4	IP Band 5	IP Band 6	IP Band 7	IP Band 8	IP Band 9	IP Band 10	Kategori	Aksi
1	1410100001	L	2,04	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
2	1410100002	L	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
3	1410100003	P	2,47	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	Tidak Tepat Waktu	Tambah Hapus Ubah
4	1410100004	P	2,03	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	Tidak Tepat Waktu	Tambah Hapus Ubah
5	1410100012	P	2,22	2,28	2,43	2,4	2,4	2,47						Tidak Tepat Waktu	Tambah Hapus Ubah
6	1410100018	L	0,78	0,08	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
7	1410100019	L	2,17	0	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
8	1410100020	L	0	0	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
9	1410100021	P	2,03	2	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
10	1410100022	L	2,17	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	Tidak Tepat Waktu	Tambah Hapus Ubah

Gambar 5. Halaman Data Testing

Pada halaman ini untuk menginput data testing digunakan fitur upload, setelah tombol upload ditekan maka data akan muncul di table. Data yang ada di tabel dapat ditambah, dihapus dan diubah kemudian datanya tersimpan dalam database.

3.4. Halaman Data Training

Pada halaman ini untuk menginput data training digunakan fitur upload, setelah tombol upload ditekan maka data akan muncul di tabel. Data yang ada di tabel dapat ditambah, dihapus dan diubah kemudian datanya tersimpan dalam database.

ID	NIM	Nama Mahasiswa	JK	IP Band 1	IP Band 2	IP Band 3	IP Band 4	IP Band 5	IP Band 6	IP Band 7	IP Band 8	IP Band 9	IP Band 10	Kategori	Aksi
1	920100001	Rizky Apriyanti DP	L	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	2,04	Tidak Tepat Waktu	Tambah Hapus Ubah
2	920100007	Lusy Adhita	P	2,78	2,85	2,84	2,85	2,85	2,85	2,85	2,85	2,85	2,85	Tidak Tepat Waktu	Tambah Hapus Ubah
3	920100004	Muhammad Aulian	L	2,86	2,86	2,86	2,86	2,86	2,86	2,86	2,86	2,86	2,86	Tidak Tepat Waktu	Tambah Hapus Ubah
4	920100002	L. Ray Ramadh Sugira	L	2,62	2,71	2,71	2,68	2,68						Tidak Tepat Waktu	Tambah Hapus Ubah
5	920100004	Kelga Reza Firmans Pratama	L	2,31	2,45	2,44	2,44	2,44	2,44	2,44	2,44	2,44	2,44	Tidak Tepat Waktu	Tambah Hapus Ubah
6	920100007	Eddy Setiawan	L	2,62	2,81	2,78	2,87	2,88						Tidak Tepat Waktu	Tambah Hapus Ubah

Gambar 5. Halaman Data Training

3.5. Halaman Pengujian

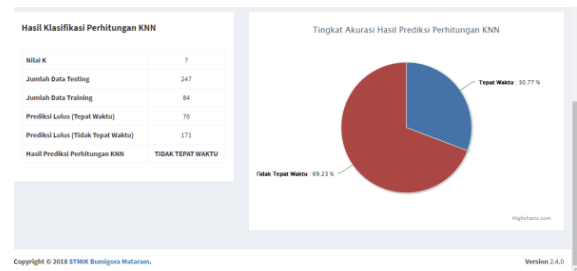
Gambar 6. Halaman Pengujian

Untuk melakukan pengujian menggunakan KNN, pengguna memasukkan nilai K yang dikehendaki kemudian sistem menampilkan hasil perhitungan jaraknya

3.6. Halaman Hasil KNN

ID	NIM	Nama	JK	IP Band 1	IP Band 2	IP Band 3	IP Band 4	IP Band 5	IP Band 6	IP Band 7	IP Band 8	IP Band 9	IP Band 10	Kategori	Aksi
1	1410100001	L	2,04	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
2	1410100002	L	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
3	1410100003	P	2,47	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	2,33	Tidak Tepat Waktu	Tambah Hapus Ubah
4	1410100004	P	2,03	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	2,4	Tidak Tepat Waktu	Tambah Hapus Ubah
5	1410100012	P	2,22	2,28	2,43	2,4	2,4	2,47						Tidak Tepat Waktu	Tambah Hapus Ubah
6	1410100018	L	0,78	0,08	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
7	1410100019	L	2,17	0	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
8	1410100020	L	0	0	0	0	0	0	0	0	0	0	0	Tidak Tepat Waktu	Tambah Hapus Ubah
9	1410100021	P	2,03	2	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	2,19	Tidak Tepat Waktu	Tambah Hapus Ubah
10	1410100022	L	2,17	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	2,07	Tidak Tepat Waktu	Tambah Hapus Ubah

Gambar 7. Halaman Hasil KNN



Gambar 8. Halaman Prediksi KNN

Pada halaman ini berisi informasi mengenai beberapa hal yaitu sebagai berikut:

- a. Hasil Prediksi Kelulusan; Hasil prediksi kelulusan setiap mahasiswa apakah tepat waktu atau tidak.
- b. Hasil Klasifikasi Perhitungan KNN; Data klasifikasi hasil perhitungan KNN Tingkat Presentasi Kelulusan ;
- c. Jumlah presentasi kelulusan mahasiswa tepat waktu dan tidak tepat waktu.

3.7. Halaman Data Pengguna

ID	Nama Pengguna	Email	Username	Password	Aksi
1	administrator	administrator@gmail.com	admin		Tambah Hapus Ubah
2	user	user@gmail.com	user		Tambah Hapus Ubah

Gambar 9. Halaman Data Pengguna

Pada halaman ini digunakan untuk menginput data pengguna, kemudian hasilnya ditampilkan di tabel. Data yang ada di tabel dapat ditambah, dihapus dan diubah kemudian datanya tersimpan dalam database.

3.8. Pengujian Algoritma KNN

3.8.1. K-Fold Cross Validation

Dalam penelitian ini digunakan 4 fold-cross validation dimana 122 data set dibagi dengan 3 cara dalam 4 bagian yaitu sebagai berikut

Data Random

K-Fold Ke-	Nilai K							Rata-Rata
	1	2	3	4	5	6	7	
1	53%	78%	45%	67%	39%	47%	36%	52%
2	50%	36%	50%	42%	44%	33%	53%	44%
3	78%	75%	81%	78%	78%	72%	69%	76%
4	29%	7%	21%	14%	14%	7%	21%	16%
Rata-Rata	53%	49%	49%	44%	44%	40%	45%	

Gambar 10. Hasil Pengujian Akurasi Data Random

Berdasarkan tabel diatas, sistem memiliki tingkat rata-rata akurasi tertinggi ketika k-fold ke 3 sebesar 76 % dan ketika nilai K=1 sebesar 53.

Data Teratur dari Keterangan Lulus Tidak Tepat Waktu ke Tepat Waktu

K-Fold Ke-	Nilai K							Rata-Rata
	1	2	3	4	5	6	7	
1	53%	78%	45%	58%	39%	53%	39%	52%
2	53%	55%	53%	64%	61%	61%	55%	57%
3	36%	14%	33%	22%	31%	8%	22%	24%
4	93%	50%	86%	71%	71%	64%	79%	74%
Rata-Rata	59%	50%	54%	54%	51%	47%	49%	

Gambar 11. Hasil Pengujian Akurasi Data Teratur

Berdasarkan tabel diatas, sistem memiliki tingkat rata-rata akurasi tertinggi ketika k-fold ke 4 sebesar 74 % dan ketika nilai K=1 sebesar 59%.

Data Teratur dari Keterangan Lulus Tepat Waktu ke Tidak Tepat Waktu

K-Fold Ke-	Nilai K							Rata-Rata
	1	2	3	4	5	6	7	
1	47%	25%	39%	22%	22%	11%	19%	26%
2	67%	56%	78%	69%	78%	69%	72%	70%
3	50%	61%	53%	67%	56%	58%	53%	57%
4	79%	86%	71%	79%	71%	86%	86%	80%
Rata-Rata	61%	57%	60%	59%	57%	56%	58%	

Gambar 12. Hasil Pengujian Akurasi Data Teratur

Berdasarkan tabel diatas, sistem memiliki tingkat rata-rata akurasi tertinggi ketika k-fold ke 4 sebesar 80 % dan ketika nilai K=1 sebesar 61%.

3.8.2. Confusion Matrix

Dalam pengujian ini menggunakan data testing dan data training sejumlah 122 data dengan rentang nilai K 1-7.

Nilai K=1

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	60	1
Tidak Tepat Waktu	1	60

Gambar 13. Hasil Pengujian Nilai K = 1

$$\begin{aligned}\text{Tingkat Akurasi} &= (60+60) / (60+1+1+60) \\ &= 120/122 = 0.98 \\ &= 0.98 \times 100 = 98\%\end{aligned}$$

Nilai K=2

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	39	22
Tidak Tepat Waktu	0	61

Gambar 14. Hasil Pengujian Nilai K = 2

$$\begin{aligned}\text{Tingkat Akurasi} &= (39+61) / (39+22+0+61) \\ &= 100/122 = 0.82 \\ &= 0.82 \times 100 = 82\%\end{aligned}$$

Nilai K-3

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	48	13
Tidak Tepat Waktu	12	49

Gambar 15. Hasil Pengujian Nilai K = 3

$$\begin{aligned}\text{Tingkat Akurasi} &= (48+49)/(48+13+12+49) \\ &= 97/122 = 0.80 \\ &= 0.80 \times 100 = 80\%\end{aligned}$$

Nilai K-4

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	40	21
Tidak Tepat Waktu	7	54

Gambar 15. Hasil Pengujian Nilai K = 4

$$\begin{aligned}\text{Tingkat Akurasi} &= (40+54)/(40+21+7+54) \\ &= 94/122 = 0.77 \\ &= 0.77 \times 100 = 77\%\end{aligned}$$

Nilai K-5

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	48	13
Tidak Tepat Waktu	13	48

Gambar 16. Hasil Pengujian Nilai K = 5

$$\begin{aligned}
 \text{Tingkat Akurasi} &= (48+48)/(48+13+13+48) \\
 &= 96/122 = 0.79 \\
 &= 0.79 \times 100 = 79 \%
 \end{aligned}$$

Nilai K-6

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	38	23
Tidak Tepat Waktu	4	57

Gambar 17. Hasil Pengujian Nilai K = 6

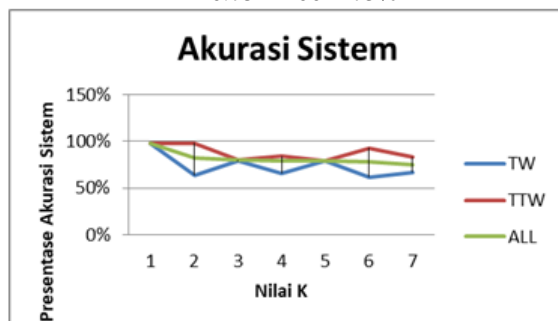
$$\begin{aligned}
 \text{Tingkat Akurasi} &= (38+57)/(38+23+4+57) \\
 &= 95/122 = 0.78 \\
 &= 0.78 \times 100 = 78\%
 \end{aligned}$$

Nilai K-7

Klasifikasi	Prediksi	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	41	20
Tidak Tepat Waktu	10	51

Gambar 18. Hasil Pengujian Nilai K = 7

$$\begin{aligned}
 \text{Tingkat Akurasi} &= (41+51)/(41+20+10+51) \\
 &= 92/122 = 0.75 \\
 &= 0.75 \times 100 = 75\%
 \end{aligned}$$



Gambar 19. Grafik Presentasi Akurasi Sistem

Berdasarkan grafik diatas diperoleh nilai presentase sebagai berikut:

1. Akurasi terendah berada pada K=6 dengan presentase 62% dan tertinggi berada pada K=1 dengan presentase 98% untuk klasifikasi “Tepat Waktu”.
2. Akurasi terendah berada pada K=5 dengan presentase 79% dan tertinggi berada pada K=2

dengan presentase 98% untuk klasifikasi “Tidak Tepat Waktu”.

3. Akurasi terendah berada pada K=7 dengan presentase 75 % dan tertinggi berada pada K=1 dengan presentasi 98% untuk klasifikasi penggabungan antara “Tepat Waktu” dan “Tidak Tepat Waktu”.

3.8.3. Penerapan Model Terbaik dengan Data Uji

Data Uji yang digunakan adalah data mahasiswa angkatan 2014 dengan jumlah data 247 yang belum diketahui label kelasnya. Adapun data training dan nilai K yang digunakan adalah: Data Training 108 Data dan Nilai K=1

Hasil yang diperoleh adalah jumlah mahasiswa yang memiliki kecenderungan lulus tepat waktu sebesar 143 mahasiswa dan tidak lulus tepat waktu sebesar 104 mahasiswa.

Data Training 122 data dan Nilai K=2

Hasil yang diperoleh adalah jumlah mahasiswa yang memiliki kecenderungan lulus tepat waktu sebesar 51 mahasiswa dan tidak lulus tepat waktu sebesar 196 mahasiswa

4. KESIMPULAN

Berdasarkan hasil analisa dan pembahasan serta hasil uji coba program pada skripsi yang berjudul “Sistem Prediksi Kelulusan Mahasiswa Menggunakan Metode KNN”, maka dapat diambil kesimpulan sebagai berikut:

- a. Pengujian dengan K-Fold Cross Validation didapatkan akurasi tertinggi pada model yang ketiga sebesar 80% ketika k-fold ke-4 dan 61% ketika nilai K=1
- b. Pengujian dengan Confusion Matrix didapatkan akurasi tertinggi sebesar 98% pada K=1 untuk klasifikasi “Tepat Waktu”.
- c. Pengujian dengan Confusion Matrix didapatkan akurasi tertinggi sebesar 98% pada K=2 untuk klasifikasi “Tidak Tepat Waktu”.
- d. Dengan menggunakan model KNN yang terbaik dari pengujian menggunakan *K-Fold Cross Validation* bahwa didapatkan hasil prediksi kelulusan mahasiswa S1 Teknik Informatika angkatan 2014 sebesar 143 orang diprediksi lulus tepat waktu dan 104 orang diprediksi tidak lulus tepat waktu .
- e. Dengan menggunakan model KNN yang terbaik dari pengujian menggunakan *Confusion Matrix* bahwa didapatkan hasil prediksi kelulusan mahasiswa S1 Teknik Informatika angkatan 2014 sebesar 51 mahasiswa diprediksi lulus tepat waktu dan 196 orang diprediksi tidak lulus tepat waktu.
- f. Semakin data tidak teratur maka tingkat akurasi menurun, sebaliknya jika data teratur maka tingkat akurasi semakin tinggi. Hal ini menunjukkan bahwa algoritma KNN yang

diimplementasi pada kasus prediksi kelulusan mahasiswa sangat dipengaruhi oleh sebaran data.

5. UCAPAN TERIMA KASIH

Kami ucapkan terima kasih kepada Universitas Bumigora atas dukungannya sehingga artikel ini bisa diterbitkan

REFERENSI

- [1] Buku Pedoman Akademik STMIK Bumigora Mataram, 2013.
- [2] Azwar.(2004). Penyusunan Skala Psikologi. Yogyakarta: Pustaka Pelajar.
- [3] Buku VI Matriks Penilaian Instrumen Akreditasi Program Studi Badan Akreditasi Nasional Perguruan Tinggi, 2008
- [4] Rohman, Abdul. "Model Algoritma K-Nearest Neighbor (KNN) Untuk Prediksi Kelulusan Mahasiswa."Jurnal Unpad 1.1 (2015).Web.
- [5] Kamber, H. &. (2006). Data Mining Concept and Tehniques. San Fransisco: Morgan Kauffman
- [6] Gorunescu, F., 2011, Data Mining: Concepts, Models and Techniques, Springer, Berlin Heidelberg.