

# Analisis Metode K-Nearest Neighbors (K-NN) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa

## *(Analysis of K-Nearest Neighbors (K-NN) and Naive Bayes Methods in Predicting Student Graduation)*

Kartarina<sup>[1]</sup>, Ni Ketut Sriwinarti<sup>[2]\*</sup>, Ni luh Putu Juniarti<sup>[3]</sup>

<sup>[1],[3]</sup>Sistem Informasi, Universitas Bumigora Mataram

E-mail: [kartarina@universitasbumigora.ac.id](mailto:kartarina@universitasbumigora.ac.id)

<sup>[2]</sup>Akuntansi, Universitas Bumigora Mataram

E-mail: [sriwinarti@universitasbumigora.ac.id](mailto:sriwinarti@universitasbumigora.ac.id)

### KEYWORDS:

*Data Mining, K-NN Methods, Naive Bayes Methods, Prediction, Student Graduation*

### ABSTRACT

*In this research the author aims to apply the K-NN and Naive Bayes algorithms for predicting student graduation rates at Sekolah Tinggi Pariwisata (STP) Mataram. The comparison of these two methods was carried out because based on several previous studies it was found that K-NN and Naive Bayes are well-known classification methods with a good level of accuracy. But which one has a better accuracy rate than the two algorithms, that's what researchers are trying to do. The output of this application is in the form of information on the prediction of student graduation, whether to graduate on time or not on time. The selection of STP as the research location was carried out because of the imbalance between the entry and exit of students who had completed their studies. Students who enter have a large number, but students who graduate on time according to the provisions are far very small, resulting in accumulation of the high number of students in each period of graduation, so it takes the initial predictions to quickly overcome these problems. Based on the results of designing, implementing, testing, and testing the Student Graduation Prediction Application program using the K-NN and Naive Bayes Methods with the Cross Validation method, the result is an accuracy for the K-NN method of 96.18% and for the Naive Bayes method an accuracy of 91.94% with using the RapideMiner accuracy test. So based on the results of the two tests between the K-NN and Naive Bayes methods which produce the highest accuracy, namely the K-NN method with an accuracy of 96.18%. So it can be concluded that the K-NN method is more feasible to use to predict student graduation*

### KATA KUNCI:

*Data Mining, Metode K-NN, Metode Naive Bayes, Prediksi, Kelulusan Mahasiswa*

### ABSTRAK

*Penelitian ini memiliki tujuan untuk menerapkan algoritma K-NN dan Naive Bayes guna memprediksi tingkat kelulusan mahasiswa di Sekolah Tinggi Pariwisata (STP) Mataram, komparasi kedua metode ini dilakukan karena berdasarkan beberapa penelitian sebelumnya ditemukan bahwasanya K-NN dan Naive Bayes merupakan metode pengklasifikasi yang terkenal dengan tingkat keakuratan yang baik. Namun mana yang tingkat akurasi yang lebih baik dari kedua algoritma tersebut, itulah yang coba peneliti lakukan. Output dari aplikasi ini berupa informasi hasil prediksi kelulusan mahasiswa, apakah lulus tepat waktu atau tidak tepat waktu. Pemilihan STP sebagai lokasi penelitian dilakukan karena adanya ketidakseimbangan antara masuk dan keluarnya mahasiswa yang telah menyelesaikan studinya. Mahasiswa yang masuk memiliki jumlah besar, tetapi mahasiswa yang lulus tepat waktu sesuai dengan ketentuan ternyata jauh sangat kecil. Sehingga terjadi penumpukan mahasiswa dalam jumlah tinggi disetiap periode kelulusan, sehingga dibutuhkan prediksi awal untuk dapat segera mengatasi permasalahan tersebut. Berdasarkan hasil perancangan, implementasi, pengujian, dan uji coba program Aplikasi Prediksi Kelulusan Mahasiswa Menggunakan Metode K-NN dan Naive Bayes dengan metode Cross Validation menghasilkan akurasi untuk metode K-NN sebesar 96,18% dan untuk metode Naive Bayes akurasi sebesar 91,94% dengan menggunakan pengujian akurasi RapideMiner. Jadi berdasarkan hasil kedua pengujian antara metode K-NN dan Naive Bayes yang menghasilkan akurasi tertinggi yaitu metode K-NN dengan akurasi 96,18%. Jadi dapat disimpulkan bahwa metode K-NN lebih layak di gunakan untuk memprediksi kelulusan mahasiswa.*

## I. PENDAHULUAN

Metode *K-Nearest Neighbor* (K-NN) merupakan suatu metode yang menggunakan algoritma supervised dimana hasil dari *Query instance* dapat diklasifikasikan berdasarkan mayoritas dari label class pada K-NN[1]. Algoritma K-NN bekerja berdasarkan jarak terpendek dari *Query instance* ke training data untuk menentukan K-NNnya dan merupakan salah satu cara untuk menghitung jarak dekat atau jauhnya suatu data menggunakan metode *euclidian distance*[2]. Sedangkan metode *Naive Bayes* adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk *machine learning* dan *Data Mining*[3]. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan "naive" dimana diasumsikan kondisi antar atribut saling bebas. Algoritma *Naive Bayes* adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayesian *Classification* didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *Decision tree* dan *Neural network* [3]. Penelitian dengan menggunakan metode *Naive Bayes* telah dilakukan untuk Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu [4]. Di mana kedua metode ini merupakan bagian dari *Data Mining* yang biasanya digunakan untuk mencari faktor mana yang paling berperan atau paling mempunyai pengaruh yang cukup besar berkaitan dengan informasi yang ingin diperoleh [5]. *Data Mining* sendiri merupakan sebuah proses ekstraksi informasi untuk menemukan pola (*pattern recognition*) yang penting dalam tumpukan data sehingga menjadi pengetahuan (*knowledge discovery*). Fungsi-fungsinya *Data Mining* antara lain fungsi deskripsi, estimasi, prediksi, clustering klasifikasi dan asosiasi [6].

Terdapat beberapa penelitian yang membuat suatu prediksi mengenai kelulusan mahasiswa tepat waktu dengan memperhitungkan atribut – atribut dari database perguruan tinggi yang digunakan. Sedangkan untuk tingkat akurasi ketiga literatur menghasilkan akurasi di atas 90% walaupun dengan menggunakan jumlah atribut dan aplikasi *Data*

*Mining* yang berbeda. Atribut yang terdapat pada semua literatur yang dapat menentukan prediksi ketepatan waktu mahasiswa lulus adalah dengan menggunakan atribut IPK (Indeks Prestasi Kumulatif) [4]. Pada penelitian selanjutnya dilakukan oleh [7] pada tahun 2019 untuk Prediksi Kelulusan Mahasiswa, hasil penelitian tersebut menunjukkan bahwa metode *K-Nearest Neighbor* (K-NN) mampu memberikan hasil yang terbaik dari pengujian menggunakan *Confusion Matrix* bahwa didapatkan hasil prediksi kelulusan mahasiswa S1 Teknik Informatika angkatan 2014 bahwa dari 247 mahasiswa, hanya 51 mahasiswa diprediksi lulus tepat waktu dan sisanya yaitu 196 orang diprediksi tidak lulus tepat waktu [7].

Pada penelitian ini penulis menggunakan komparasi dua metode yaitu *Naive Bayes* dan *K-Nearest Neighbor* di karenakan dua metode ini yaitu KNN dan *Naive bayes* merupakan metode pengklasifikasi yang terkenal dengan tingkat keakuratan yang baik. Beberapa penelitian sebelumnya kebanyakan berbasiskan pada ilmu komputer atau informatika sehingga pada pembahasannya lebih ditekankan pada hasil analisis menggunakan aplikasi serta tema yang diambil berkaitan dengan hal-hal yang bersifat elektronik. Selain itu berbeda dengan metode pengklasifikasian dengan regresi logistik ordinal maupun nominal, pada metode *Naive Bayes* dan *K-Nearest Neighbor* pengklasifikasian tidak diperlukan adanya permodelan maupun uji statistik seperti uji signifikansi. Data set yang akan di gunakan dalam penelitian kali ini adalah Data set bersumber dari data Mahasiswa di Sekolah Tinggi Pariwisata (STP) Mataram. Sekolah Tinggi Pariwisata (STP) Mataram adalah sekolah Tinggi Pariwisata pertama di Nusa Tenggara Barat (NTB), STP merupakan yayasan pertama dan terbaik di NTB yang selalu menarik lumayan banyak mahasiswa setiap prodinya dalam tiap tahun. Dari data yang bersumber dari Administrasi Akademik dan Kemahasiswaan (BAAK) mahasiswa masuk pada tahun 2013 sekitar 214, 2014 sekitar 229, 2015 sekitar 203 dan 2016 sekitar 235 mahasiswa yang lulusan angkatan tahun 2013 sekitar 156 mahasiawa yang lulus angkatan 2014 sekitar 162 mahasiswa dan mahasiswa lulusan angkatan 2015 yang lulus sekitar 147 mahasiwa dan angkatan 2016 yang lulus sekitar 187. Dari data di atas bisa kita ketahui bahwa data tidak konsisten dalam artian jumlah peserta didik baru dengan

peserta yang lulus tepat waktu haruslah berbanding lurus. namun pada kenyataannya masih ada peserta didik yang lulus pada semester yang lebih dari waktu yang ditentukan di STP Mataram, sehingga membuat penumpukan mahasiswa/ dengan jumlah tinggi di setiap periode lulus, yang membuat STP Mataram kesulitan mengelola data mahasiswa/i dalam mengetahui pengelompokan mahasiswa lulus tepat waktu atau lulus tidak tepat waktu.

Tingkat lulus tepat waktu yang tinggi adalah keuntungan bagi dua pihak. Pertama, untuk pihak mahasiswa, mahasiswa akan mendapatkan pekerjaan dengan lebih mudah karena perusahaan cenderung mencari yang baru lulus. Kedua, untuk pihak universitas, karena seiring tepatnya waktu kelulusan mahasiswa, hal itu akan membantu memajukan kualitas universitas tersebut, seperti peningkatan akreditasi [8]. Oleh karena itu untuk meningkatkan kualitas dan akreditasi pada STP Mataram, maka harus ada penyaringan pada mahasiswa yang mendaftar untuk masuk dan perlu diprediksi mahasiswa yang lulus tepat waktu karena berpengaruh langsung terhadap kapasitas ruang kelas, rasio dosen terhadap mahasiswa/i, beserta kapasitas tempat parkir. Prediksi waktu kelulusan mahasiswa merupakan salah satu yang penting dan tepat untuk membentuk pola-pola yang mungkin memberikan indikasi yang bermamfaat pada data mahasiswa yang berjumlah besar. Untuk mencapai tujuan tersebut pada penelitian ini peneliti menggunakan algoritma K-NN dalam menentukan prediksi kelulusan berdasarkan atribut, Nim, jenis kelamin, Indeks Prestasi (IP) semester 1-5 dan IPK. Berdasarkan latar belakang yang sudah dipaparkan, maka dalam skripsi ini penulis bertujuan untuk menerapkan algoritma K-NN dan *Naive Bayes* untuk prediksi tingkat kelulusan mahasiswa di STP Mataram, dan membandingkan tingkat akurasi dari kedua algoritma sehingga pihak kampus dapat meminimalisir tingkat kelulusan tidak tepat waktu dan mengetahui algoritma mana yang memiliki tingkat akurasi yang lebih tinggi.

## II. METODOLOGI

### 2.1. Metode Pengumpulan Data

#### 2.1.1 Wawancara

Wawancara dilakukan dengan cara mengumpulkan data dengan cara bertanya atau melakukan tanya jawab secara langsung dengan pihak bagian

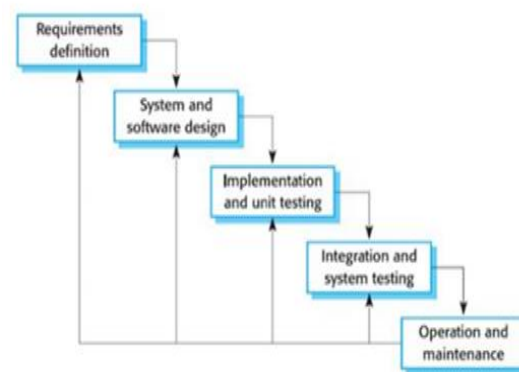
Akademik dan K-Prodi D3 Perhotelan yang bertujuan untuk Mengetahui data data Mahasiswa STP Mataram Jurusan D3 Perhotelan.

#### 2.1.2. Studi Pustaka

Metode ini dilakukan dengan cara mempelajari dan mencari referensi pada jurnal maupun literature lain yang berhubungan dengan penelitian yang dilakukan yaitu Prediksi Kelulusan Mahasiswa menggunakan Algoritma *K-Nearest Neighbor* (K-NN) dan *Naive Bayes*.

### 2.2. Metode pengembangan sistem

Metode yang digunakan dalam penelitian ini mengadopsi dari metode waterfall memiliki yang tahapan – tahapan sebagai berikut[9] :



Gbr. 1 Metode Waterfall

#### a. Requirements analysis and definition

Layanan sistem, kendala, dan tujuan ditetapkan oleh hasil konsultasi dengan pengguna yang kemudian didefinisikan secara rinci dan berfungsi sebagai spesifikasi sistem.

#### b. System and software design

Tahapan perancangan sistem mengalokasikan kebutuhan-kebutuhan sistem baik perangkat keras maupun perangkat lunak dengan membentuk arsitektur sistem secara keseluruhan. Perancangan perangkat lunak melibatkan identifikasi dan penggambaran abstraksi sistem dasar perangkat lunak dan hubungannya.

#### c. Implementation and unit testing

Pada tahap ini, perancangan perangkat lunak direalisasikan sebagai serangkaian program atau unit program. Pengujian melibatkan verifikasi bahwa setiap unit memenuhi spesifikasinya.

#### d. Integration and system testing

Unit-unit individu program atau program digabung dan diuji sebagai sebuah sistem lengkap untuk memastikan apakah sesuai dengan kebutuhan perangkat lunak atau tidak. Setelah pengujian, perangkat lunak dapat dikirimkan ke customer.

#### e. Operation and maintenance

Biasanya (walaupun tidak selalu), tahapan ini merupakan tahapan yang paling panjang. Sistem dipasang dan digunakan secara nyata. Maintenance.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil Perhitungan Sistem Menggunakan K-NN

Hasil perhitungan sistem yang sudah dibuat dengan menggunakan metode K-NN dengan data testing yang digunakan sebagai berikut:

| NIM      | Nama         | Jenis Kelamin | IPS1 | IPS2 | IPS3 | IPS4 | IPS5 | IPK | Keterangan |
|----------|--------------|---------------|------|------|------|------|------|-----|------------|
| 15302026 | DWIN APRIADI | L             | 3.68 | 3.32 | 3.1  | 3    | 3.52 | 3.4 | ?          |

| Normalisasi Dataset |               |               |               |               |               |               |       |             |       |
|---------------------|---------------|---------------|---------------|---------------|---------------|---------------|-------|-------------|-------|
| 7 Nearest           |               |               |               |               |               |               |       |             |       |
| Nomor               | Jenis Kelamin | IP Semester 1 | IP Semester 2 | IP Semester 3 | IP Semester 4 | IP Semester 5 | IPK   | Keterangan  | Jarak |
| 82                  | 0.5           | 0.92          | 0.83          | 0.477         | 0.75          | 0.88          | 0.85  | TEPAT WAKTU | 0.045 |
| 225                 | 0.5           | 0.935         | 0.84          | 0.523         | 0.773         | 0.845         | 0.843 | TEPAT WAKTU | 0.07  |
| 213                 | 0.5           | 0.908         | 0.803         | 0.515         | 0.773         | 0.833         | 0.833 | TEPAT WAKTU | 0.093 |
| 484                 | 0.5           | 0.935         | 0.843         | 0.485         | 0.825         | 0.833         | 0.848 | TEPAT WAKTU | 0.095 |
| 254                 | 0.5           | 0.92          | 0.843         | 0.508         | 0.818         | 0.828         | 0.863 | TEPAT WAKTU | 0.1   |
| 228                 | 0.5           | 0.908         | 0.843         | 0.508         | 0.818         | 0.845         | 0.848 | TEPAT WAKTU | 0.101 |
| 129                 | 0.5           | 0.973         | 0.75          | 0.462         | 0.785         | 0.828         | 0.858 | TEPAT WAKTU | 0.101 |
| Total               |               |               |               |               |               |               |       |             |       |
| Keterangan          |               |               |               |               |               |               |       |             | Total |
| TEPAT WAKTU         |               |               |               |               |               |               |       |             | 7     |

Berdasarkan perhitungan, dengan Jenis Kelamin: 1, IP Semester 1: 3.86, IP Semester 2: 3.32, IP Semester 3: 3.1, IP Semester 4: 3, IP Semester 5: 3.52, IPK: 3.4, maka hasilnya: **TEPAT WAKTU**.

**Gbr. 2** Hasil Perhitungan Sistem Menggunakan K-NN

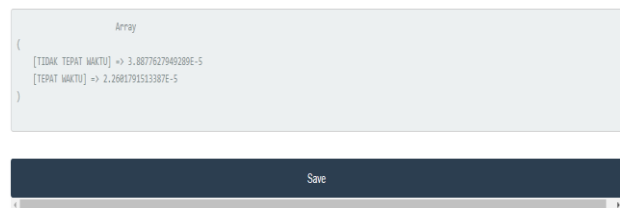
Dari hasil perhitungan K-NN diatas dengan nilai  $K = 7$ , didapatkan hasil keterangan **TEPAT WAKTU** adalah 7 dan **TIDAK TEPAT WAKTU** adalah 0. Maka dapat disimpulkan mahasiswa dengan nama Dwin Apriadi, nim 15302026, Jenis Kelamin: 1, IP Semester 1: 3.86, IP Semester 2: 3.32, IP Semester 3: 3.1, IP Semester 4: 3, IP Semester 5: 3.52, IPK: 3.4, hasilnya adalah: **TEPAT WAKTU**.

#### B. Hasil Perhitungan Sistem Menggunakan Naive Bayes

Hasil perhitungan sistem yang sudah dibuat dengan menggunakan metode *Naive Bayes* dengan data testing yang digunakan sebagai berikut:

| NIM      | Nama                 | Jenis Kelamin | IPS1 | IPS2 | IPS3 | IPS4 | IPS5 | IPK  | Keterangan |
|----------|----------------------|---------------|------|------|------|------|------|------|------------|
| 15302088 | AKHMAD ZAILAN FATONI | L             | 2.55 | 3.05 | 2.95 | 2.77 | 2.76 | 2.99 | ?          |

| NIM      | NAMA                 | Jenis Kelamin | IP SEMESTER 1 | IP SEMESTER 2 | IP SEMESTER 3 | IP SEMESTER 4 | IP SEMESTER 5 | IPK       | KETERANGAN LULUS  |
|----------|----------------------|---------------|---------------|---------------|---------------|---------------|---------------|-----------|-------------------|
| 15302088 | AKHMAD ZAILAN FATONI | L             | CUKUP         | MEMUASKAN     | CUKUP         | CUKUP         | CUKUP         | MEMUASKAN | TIDAK TEPAT WAKTU |



**Gbr. 3** Hasil Perhitungan Sistem Menggunakan Naive Bayes

Dari hasil perhitungan Naive Bayes diatas, didapatkan hasil keterangan **TIDAK TEPAT WAKTU**. Maka dapat disimpulkan mahasiswa dengan nama Akhmad Zailan Fatoni, nim 15302088, Jenis Kelamin: L, IP Semester 1: 2.55, IP Semester 2: 3.05, IP Semester 3: 2.95, IP Semester 4: 2.77, IP Semester 5: 2.76, IPK: 2.99, hasilnya adalah: **TIDAK TEPAT WAKTU**.

#### C. Tingkat Akurasi Menggunakan K-NN

##### 1). Data Awal

Data yang digunakan berjumlah 496 data dan diuji menggunakan *Cross validation*.

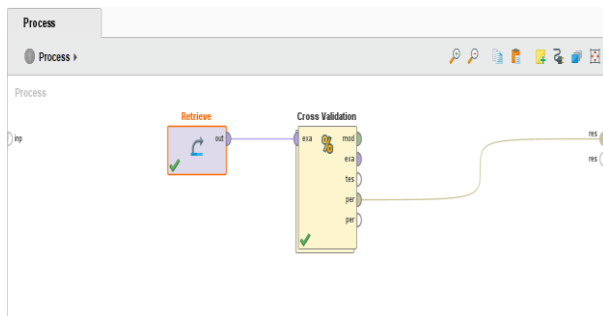
| Row No. | Keterangan                                                                  | JK | IPS1  | IPS2  | IPS3  | IPS4  | IPS5  | IPK   |
|---------|-----------------------------------------------------------------------------|----|-------|-------|-------|-------|-------|-------|
| 1       | [The position of the example in the (filtered) view on the example table.]0 |    |       |       | 3     | 3.410 | 3.430 | 3.450 |
| 2       | TEPAT WAKTU                                                                 | 1  | 3.630 | 4     | 3.600 | 3.910 | 4     | 3.850 |
| 3       | TEPAT WAKTU                                                                 | 2  | 3.420 | 3.840 | 3.100 | 3.770 | 4     | 3.680 |
| 4       | TEPAT WAKTU                                                                 | 2  | 3.630 | 3.630 | 3.350 | 3.590 | 3.810 | 3.650 |
| 5       | TEPAT WAKTU                                                                 | 1  | 3.890 | 4     | 3.450 | 3.910 | 3.620 | 3.800 |
| 6       | TEPAT WAKTU                                                                 | 1  | 3.630 | 4     | 3.550 | 4     | 3.810 | 3.820 |
| 7       | TEPAT WAKTU                                                                 | 2  | 3.630 | 4     | 3.550 | 3.910 | 4     | 3.840 |
| 8       | TEPAT WAKTU                                                                 | 1  | 3.740 | 3.890 | 3.850 | 3.770 | 4     | 3.870 |
| 9       | TEPAT WAKTU                                                                 | 2  | 3.050 | 4     | 3.700 | 3.820 | 3.620 | 3.690 |
| 10      | TEPAT WAKTU                                                                 | 2  | 3.840 | 3.680 | 3.600 | 3.910 | 3.900 | 3.820 |
| 11      | TEPAT WAKTU                                                                 | 2  | 3.890 | 4     | 3.700 | 3.910 | 3.710 | 3.860 |
| 12      | TEPAT WAKTU                                                                 | 2  | 3.890 | 4     | 3.700 | 4     | 4     | 3.930 |
| 13      | TEPAT WAKTU                                                                 | 1  | 3.680 | 4     | 3.700 | 4     | 4     | 3.900 |
| 14      | TEPAT WAKTU                                                                 | 2  | 3.680 | 4     | 3.700 | 4     | 4     | 3.900 |
| 494     | TEPAT WAKTU                                                                 | 1  | 2.890 | 2.740 | 2.850 | 3.350 | 2.330 | 2.930 |
| 495     | TEPAT WAKTU                                                                 | 1  | 3.630 | 3.630 | 3.250 | 3.950 | 3.810 | 3.620 |
| 496     | TEPAT WAKTU                                                                 | 2  | 3     | 3.680 | 3.300 | 3.700 | 3.710 | 3.570 |

**Gbr. 4** Data Awal Perhitungan Akurasi K-NN

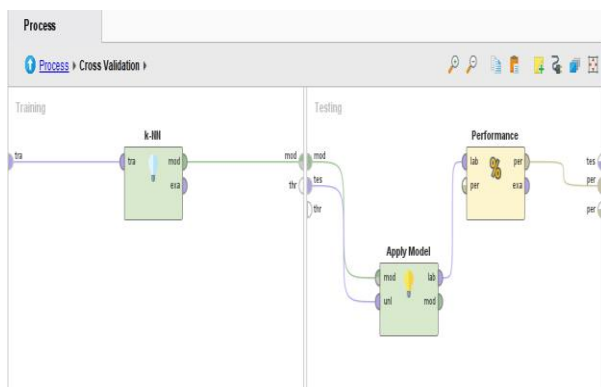


Gbr. 5 Grafik Akurasi K-NN

## 2) Design Sistem Perhitungan Akurasi



Gbr. 6 Design Sistem Perhitungan Akurasi K-NN



Gbr. 7 Design Proses Cross Validation K-NN

## 3) Hasil Akurasi

accuracy: 96.18% +/- 1.12% (micro average: 96.17%)

|                         | true TEPAT WAKTU | true TIDAK TEPAT WAKTU | class precision |
|-------------------------|------------------|------------------------|-----------------|
| pred. TEPAT WAKTU       | 477              | 19                     | 96.17%          |
| pred. TIDAK TEPAT WAKTU | 0                | 0                      | 0.00%           |
| class recall            | 100.00%          | 0.00%                  |                 |

Gbr. 7 Hasil Akurasi K-NN RapidMiner

Dari hasil pengujian yang sudah dilakukan menggunakan 496 data dengan metode pengujian

Cross Validation didapatkan hasil Akurasi metode K-NN sebesar 96,18%.

### - Sensitivitas (Precision)

$$TP = 477$$

$$TN = 0$$

$$FP = 0$$

$$FN = 19$$

$$Sensitivitas(Precision) = \frac{TP}{TP + FN} * 100\% = \frac{477}{477 + 19} * 100\% = 96,17\%$$

### - Spesifitas (Recall)

$$TP = 477$$

$$TN = 0$$

$$FP = 0$$

$$FN = 19$$

$$Spesifitas(Recall) = \frac{TN}{TN + FP} * 100\% = \frac{0}{0} * 100\% = 0\%$$

## D. Tingkat Akurasi Menggunakan Naive Bayes

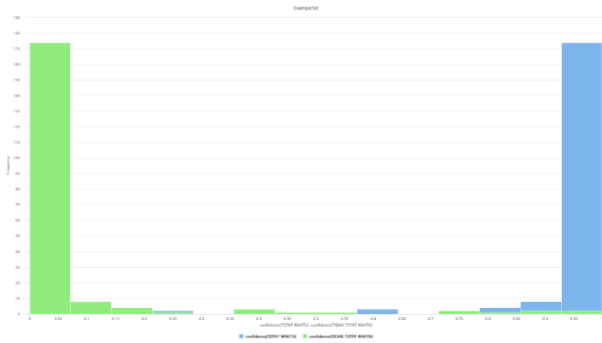
### 1) Data Awal

Data yang digunakan berjumlah 496 data, dan dibagi menjadi 60% data training dan 40% data testing.

| Row No. | Keterangan  | JK | IPS1         | IPS2         | IPS3         | IPS4         | IPS5         | IPK          |
|---------|-------------|----|--------------|--------------|--------------|--------------|--------------|--------------|
| 1       | TEPAT WAKTU | L  | MEMUASKAN    | DENGAN PU... | MEMUASKAN    | MEMUASKAN    | MEMUASKAN    | MEMUASKAN    |
| 2       | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 3       | TEPAT WAKTU | P  | MEMUASKAN    | DENGAN PU... | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 4       | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 5       | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 6       | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 7       | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 8       | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 9       | TEPAT WAKTU | P  | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 10      | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 11      | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 12      | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 13      | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 14      | TEPAT WAKTU | P  | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 494     | TEPAT WAKTU | L  | CUKUP        | CUKUP        | CUKUP        | MEMUASKAN    | KURANG       | CUKUP        |
| 495     | TEPAT WAKTU | L  | DENGAN PU... | DENGAN PU... | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... |
| 496     | TEPAT WAKTU | P  | MEMUASKAN    | DENGAN PU... | MEMUASKAN    | DENGAN PU... | DENGAN PU... | DENGAN PU... |

Gbr. 8 Data Awal Perhitungan Akurasi Naive Bayes

Sedangkan grafik akurasi Naive Bayes tampak seperti pada gambar dibawah ini :



Gbr. 9 Grafik Akurasi Naive Bayes

Dari hasil pengujian yang sudah dilakukan menggunakan 496 data dengan metode pengujian *Cross Validation* didapatkan hasil Akurasi metode *Naive Bayes* sebesar 91,94%.

- *Sensitivitas (Precision)*

TP = 447

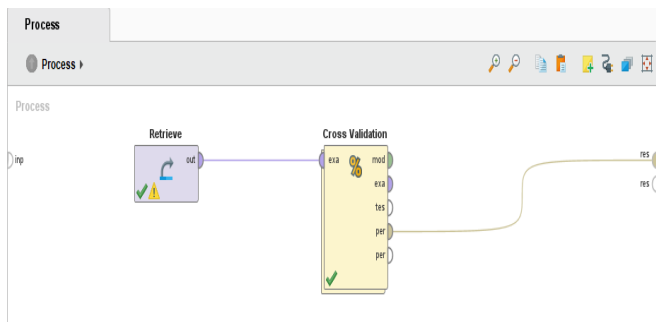
TN = 9

FP = 30

FN = 10

$$\text{Sensitivitas(Precision)} = \frac{TP}{TP + FN} * 100\% = \frac{447}{447 + 10} * 100\% = 97,81\%$$

## 2) Design Sistem Perhitungan Akurasi



Gbr. 10 Design Sistem Perhitungan Akurasi Naive Bayes

- *Spesifitas (Recall)*

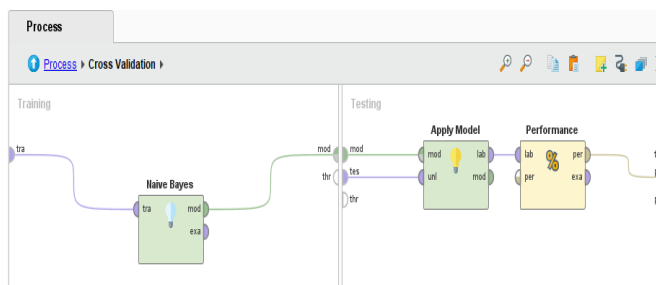
TP = 447

TN = 9

FP = 30

FN = 10

$$\text{Spesifitas(Recall)} = \frac{TN}{TN + FP} * 100\% = \frac{9}{9 + 30} * 100\% = 23,08\%$$



Gbr. 11 Design Proses Cross Validation

## IV. KESIMPULAN

Berdasarkan hasil perancangan, implementasi, pengujian, dan uji coba program Aplikasi Prediksi Kelulusan Mahasiswa Menggunakan Metode K-NN dan *Naive Bayes* dengan perbandingan data training 60% dan data testing 40% menghasilkan kesimpulan sebagai berikut:

1. Berdasarkan hasil pengujian dapat dikatakan bahwa prediksi kelulusan mahasiswa dengan metode *Naive Bayes* dan metode *K-Nearest Neighbor* sudah baik atau sudah akurat
2. Berdasarkan perhitungan akurasi menggunakan RapidMiner Metode K-NN yaitu sebesar 96,18% dan metode *Naive Bayes* sebesar 91,94%. Dari kedua nilai tersebut dapat dikatakan bahwa baik metode *Naive Bayes* ataupun metode *K-Nearest Neighbor* memiliki peluang yang kecil untuk melakukan kesalahan dalam proses prediksi.
3. Dilihat dari nilai pengujian dan akurasi dapat disimpulkan bahwa pengklasifikasian menggunakan metode *K-Nearest Neighbor* lebih baik dibandingkan dengan metode *Naive Bayes* dalam proses prediksi kelulusan mahasiswa

## 3) Hasil Akurasi

Table View Plot View

accuracy: 91.94% +/- 3.15% (micro average: 91.94%)

|                         | true TEPAT WAKTU | true TIDAK TEPAT WAKTU | class precision |
|-------------------------|------------------|------------------------|-----------------|
| pred. TEPAT WAKTU       | 447              | 10                     | 97.81%          |
| pred. TIDAK TEPAT WAKTU | 30               | 9                      | 23.08%          |
| class recall            | 93.71%           | 47.37%                 |                 |

Gbr. 12 Akurasi Naive Bayes RapidMiner

## UCAPAN TERIMA KASIH

Terima kasi kami ucapkan pada Sekolah Tinggi Pariwisata (STP) Mataram, atas bantuan dan kesediaannya dalam memberikan data sehingga penelitian ini dapat terselesaikan dengan baik. Terima kasih juga kepada sekawan team atas bantuan dan masukan atas tulisan dalam penelitian ini.

## REFERENSI

- [1] H. Romadhona, A., Suprapedi, S. dan Himawan, "Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, Dan Indeks Prestasi Menggunakan Algoritma Decision Tree," *J. Teknol. Inf.*, vol. 13, pp. 69–83, 2017.
- [2] I. Budiman, D. T. Nugrahadi, and R. A. Nugroho, "Implementasi Algoritma K-Nearest Neighbour untuk Prediksi Waktu Kelulusan Mahasiswa," *Pros. SNRT (Seminar Nas. Ris. Ter.)*, vol. 5662, pp. 9–10, 2016.
- [3] A. Azahari, Y. Yulindawati, D. Rosita, and S. Mallala, "Komparasi Data Mining Naive Bayes dan Neural Network memprediksi Masa Studi Mahasiswa S1," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, p. 443, 2020, doi: 10.25126/jtiik.2020732093.
- [4] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review," *Fakt. Exacta*, vol. 13, no. 1, pp. 38–47, 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [5] A. A. Murtopo, "Prediksi Kelulusan Tepat Waktu Mahasiswa STMIK YMI Tegal Menggunakan Algoritma Naïve Bayes," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 7, no. 3, p. 145, 2016, doi: 10.22303/csrid.7.3.2015.145-154.
- [6] D. Kartini, "Penerapan Data Mining dengan Algoritma Neural Network (Backpropagation) Untuk Prediksi Lama Studi Mahasiswa," *Pros. Semin. Nas. Sisfotek*, vol. 3584, pp. 235–241, 2017.
- [7] L. A. R. Hakim, A. A. Rizal, and D. Ratnasari, "Aplikasi Prediksi Kelulusan Mahasiswa Berbasis K-Nearest Neighbor (K-NN)," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 1, no. 1, pp. 30–36, 2019, doi: 10.35746/jtim.v1i1.11.
- [8] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, "Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa indonesia," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, p. 427, 2018, doi: 10.25126/jtiik.201854773.
- [9] G. W. Sasmito, "Penerapan Metode Waterfall Pada Desain Sistem Informasi Geografis Industri Kabupaten Tegal," *J. Inform. Pengemb. IT*, vol. 2, no. 1, pp. 6–12, 2017.