

# [IMPLEMENTATION OF DATA MINING ON TELKOM CUSTOMER DATA USING THE K-NEAREST NEIGHBOR METHOD TO PREDICT SERVICE STATUS]

## [IMPLEMENTASI DATA MINING PADA DATA PELANGGAN TELKOM MENGGUNAKAN METODE K-NEAREST NEIGHBOR UNTUK MEMPREDIKSI STATUS PELAYANAN]

Ervina Agusetiana<sup>1)</sup>, Arif Senja Fitriani<sup>2)</sup>, Rohman Dijaya<sup>3)</sup>, Nuril Lutvi Azizah<sup>4)</sup>

<sup>1)</sup> Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

<sup>2)</sup> Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

<sup>3)</sup> Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

<sup>4)</sup> Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

\*Email Penulis Korespondensi: asfjim@umsida.ac.id

**Abstract.** *The use of the internet makes it easier for people to get information. The large number of wifi users resulted in an increase in Telkom Indonesia. This impact resulted in frequent delays in the service process. The purpose of this study is to classify service status as normal or abnormal. This study implements the k-nearest neighbor classification method. The data used is Telkom customer data of 15,113 (May to June 2018), dataset of 100 data. The input consists of 60% training data and 40% testing data. Weka processes calculations and displays classification results in the form of normal or abnormal. In Weka, the correctness prediction is 92.5%, the error prediction is 7.5%. In manual calculations, the correctness prediction is 98%, the error prediction is 3%. From the results of calculations using KNN, the highest results were obtained using manual with truth level of 98%, error rate of 3%.*

**Keywords** – data mining; knn; sidoarjo

**Abstrak.** *Penggunaan internet memudahkan masyarakat untuk mendapatkan informasi. Banyaknya jumlah pengguna wifi mengakibatkan peningkatan di Telkom Indonesia. Dampak tersebut mengakibatkan sering terjadinya keterlambatan dalam proses pelayanan. Tujuan dari penelitian ini adalah untuk mengklasifikasikan status pelayanan normal atau tidak normal. Penelitian ini mengimplementasikan metode klasifikasi k-nearest neighbor. Data yang digunakan adalah data pelanggan Telkom sebanyak 15.113 (Mei sampai dengan Juni 2018), dataset sebanyak 100 data. Input terdiri dari 60% data training dan 40% data testing. Weka memproses perhitungan dan menampilkan hasil klasifikasi berupa normal atau abnormal. Di Weka, prediksi kebenarannya adalah 92,5%, prediksi kesalahannya adalah 7,5%. Pada perhitungan manual, prediksi kebenarannya adalah 98%, prediksi kesalahannya adalah 3%. Dari hasil perhitungan menggunakan KNN diperoleh hasil tertinggi dengan menggunakan cara manual dengan tingkat kebenaran 98%, tingkat kesalahan 3%.*

**Kata Kunci** - data mining; knn; sidoarjo

## I. PENDAHULUAN

Berdasarkan hasil survei dari Asosiasi Penyelenggara Jasa Internet Indonesia telah didapatkan hasil bahwasanya jumlah pengguna internet di Indonesia sekitar 171,17 juta jiwa dari jumlah keseluruhan penduduk Indonesia yaitu sebesar 264,16 juta jiwa. Dilihat dari sisi profil pengguna internet di Indonesia pada zaman sekarang ini survei membuktikan, rentang usia 15-19 tahun merupakan kontributor utama dari sisi usia pengguna dengan 91 persen, rentang usia 20-24 tahun sebesar 88,5 persen, rentang usia 25-29 tahun sebesar 82,7 persen, rentang usia 30-34 tahun sebesar 76,5 persen, rentang usia 35-39 tahun sebesar 68,5 persen, rentang usia 40-44 tahun sebesar 51,4 persen, rentang usia 45-49 tahun sebesar 47,6 persen, dan rentang usia 50-54 tahun sebesar 40,9 persen. [1]

Dengan meningkatnya pengguna internet di Indonesia mengakibatkan meningkat pula jumlah pelanggan PT. Telkom Indonesia. Namun dari banyaknya *request order* pelanggan telkom, ada beberapa faktor yang mempengaruhi waktu pelayanan dari pihak telkom terhadap pelanggan sehingga status pelayanan ada yang normal dan adakalanya juga tidak normal. Pada permasalahan di atas, peneliti ingin mengimplementasikan data mining pada data pelanggan telkom menggunakan metode klasifikasi K-Nearest Neighbor untuk memprediksi status pelayanan dan dilaporkan kepada manajer guna mengetahui waktu dalam pelayanan pemasangan wifi sehingga bisa dijadikan evaluasi oleh manajer kepada teknisi agar lebih optimal dalam proses pelayanan.

### A. PT. Telkom

PT Telkom Indonesia ialah sebuah perusahaan yang bergerak pada bidang jasa berupa layanan komunikasi dan jaringan telekomunikasi di Indonesia. Indonesia adalah negara pemegang saham terbanyak yaitu sebesar 52,09% sedangkan 47,91% sisanya dimiliki oleh publik. Saham milik Telkom dengan kode “TLK” kepada New York Exchange (NYSE) dan dengan kode “TLKM” kepada Bursa Efek Indonesia (BEI). [2]

### B. Kualitas Pelayanan

Pelayanan ialah proses pemenuhan kebutuhan yang dilakukan oleh orang lain melalui suatu aktivitas secara langsung. Baik tidaknya pelayanan terletak pada kualitas pelayanan. Kualitas Pelayanan (dalam Tjiptono, 2008:8) merupakan “ukuran seberapa bagus tingkat layanan yang diberikan mampu sesuai dengan ekspektasi pelanggan”. Kualitas pelayanan sangat bergantung pada kemampuan suatu penyedia jasa dalam memenuhi keinginan konsumen yang dilakukan dengan konsisten. Kualitas pelayanan ialah suatu upaya pemenuhan kebutuhan dan keinginan konsumen serta ketepatan penyampaiannya dalam mengimbangi harapan konsumen. [3]

### C. Pelayanan Telkom

Ketersediaan sarana dan prasana yang dimiliki oleh PT. Telkom telah tersebar di seluruh Indonesia. Salah satu contoh akses internet yang dimiliki PT. Telkom ialah Speedy. Speedy yaitu layanan akses internet bagi perumahan yang berkualitas tinggi. PT. Telkom juga memberikan layanan jaringan wifi. Layanan yang diberikan ialah pasang wifi baru, penambahan maupun kecepatan akses wifi dan pergantian kabel. PT. Telkom mementingkan kepuasan pelanggan dalam pelayanannya. [4]

### D. Teori Data Mining

Data Mining merupakan proses pengumpulan data yang diolah dengan berbagai macam metode yang sesuai. data mining dikelompokkan menjadi deskripsi, estimasi, prediksi, klasifikasi, clustering dan asosiasi (Larose, 2005). Tujuan dari data mining yaitu untuk memanfaatkan data dan mengolahnya hingga mendapatkan sebuah informasi yang baru dan berguna. [5]

### E. Flowchart

Flowchart ialah sebuah teknik analitis yang digunakan untuk menjelaskan aspek-aspek sistem informasi dengan jelas, tepat dan logis (Krisnaji, 2005). Secara umum, pengertian flowchart merupakan gambaran dari langkah-langkah dan urutan prosedur sebuah program secara grafik. [9]

### F. WEKA

“WEKA” merupakan singkatan dari “Waikato Environment for Knowledge Analysis”, yang dibuat di Universitas Waikato, New Zealand untuk penelitian, pendidikan dan berbagai aplikasi. WEKA mengandung tools untuk preprocessing data, klasifikasi, regresi, clustering, aturan asosiasi, dan visualisasi. [10]

## II. METODE

### A. Metode KNN (*K-Nearest Neighbor*)

Metode *K-Nearest Neighbor* adalah salah satu metode yang digunakan dalam pengklasifikasian data. Prinsip kerja *K-Nearest Neighbor* adalah mencari jarak terdekat antara data yang akan di evaluasi dengan *K* tetangga terdekatnya dalam data pelatihan (Rismawan, dkk. 2008). [7]

Berikut rumus pencarian jarak menggunakan rumus Euclidean (Sutanto, 2009). [8]

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

Penjelasan dari formula umum algoritma k-nearest neighbor ialah sebagai berikut:

$X_j$  = Sample data/data training

$x_{ik}$  = Data uji/ testing

$I_j$  = Variabel Data

$d$  = Jarak

$p$  = Dimensi

Berikut ialah contoh soal menggunakan rumus KNN yaitu sebagai berikut: [6]

**Tabel 1.** Data Training

$X_1$ = Asam Durabilitas (detik)	$X_2$ = Kekuatan (Kg / meter persegi)	Klasifikasi
7	7	Buruk
7	4	Buruk
3	4	Baik
1	4	Baik

Saat ini pabrik kertas telah menghasilkan jaringan baru yang lulus uji laboratorium dengan  $X_1=3$  dan  $X_2=7$ . Untuk menebak klasifikasi jaringan baru ini maka dilakukan perhitungan dengan menggunakan algoritma KNN. Adapun langkah-langkah untuk menghitung  $K$  tetangga terdekat dengan algoritma KNN adalah sebagai berikut.

a.) Tentukan parameter  $K$  (jumlah tetangga terdekat). Misal  $K = 3$  b.) Hitung jarak antara permintaan (data testing) dan contoh-contoh latihan semua (data training). Data training yang akan dihitung kedekatannya mempunyai koordinat (3,7).

Tabel 2. Perhitungan Jarak.

$X_1 =$ Asam Durabilitas (detik)	$X_2 =$ Kekuatan (Kg / meter persegi)	Square Jarak ke contoh permintaan (3, 7)
7	7	$(7-3)^2 + (7-7)^2 = 16$
7	4	$(7-3)^2 + (4-7)^2 = 25$
3	4	$(3-3)^2 + (4-7)^2 = 9$
1	4	$(1-3)^2 + (4-7)^2 = 13$

c.) Urutkan jarak dan menentukan tetangga terdekat berdasarkan jarak terdekat ke -  $K$ .

Tabel 3. Urutan Peringkat Jarak Minimum

$X_1 =$ Asam Durabilitas (detik)	$X_2 =$ Kekuatan (Kg / meter persegi)	Square Jarak ke contoh permintaan (3, 7)	Peringkat jarak minimum	Apakah termasuk dalam tetangga 3?
7	7	$(7-3)^2 + (7-7)^2 = 16$	3	Ya
7	4	$(7-3)^2 + (4-7)^2 = 25$	4	Tidak
3	4	$(3-3)^2 + (4-7)^2 = 9$	1	Ya
1	4	$(1-3)^2 + (4-7)^2 = 13$	2	Ya

d.) Kumpulkan kategori  $Y$  dari baris tetangga terdekat. Pada baris kedua kategori tetangga terdekat ( $Y$ ) tidak dimasukkan karena peringkatnya lebih dari 3 tetangga terdekat.

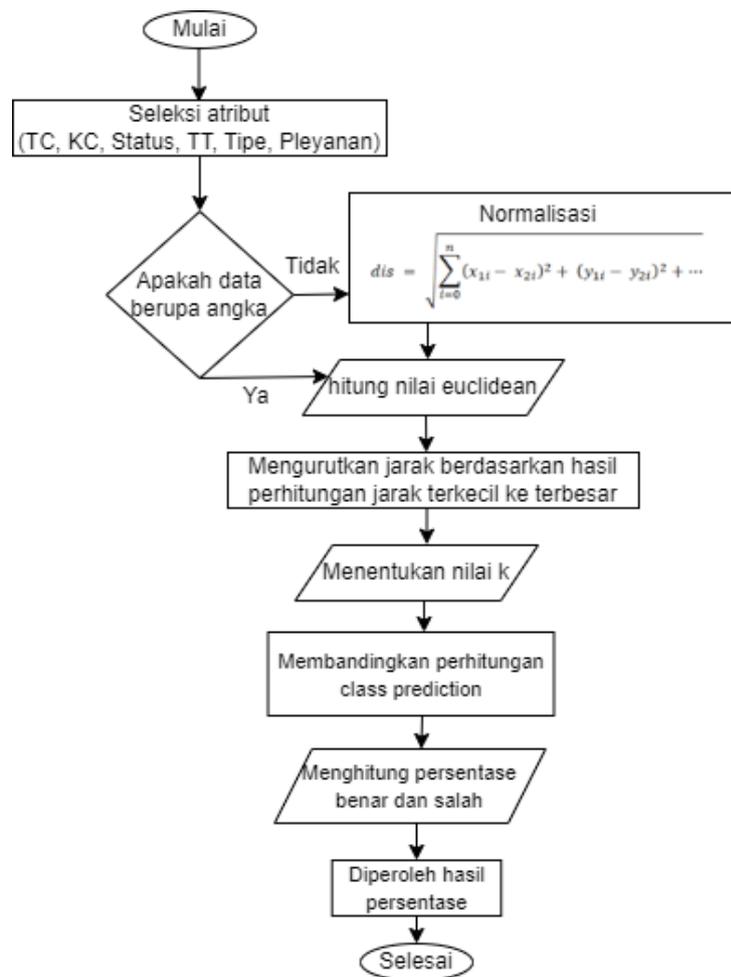
Tabel 4. Kumpulan Kategori  $Y$  Tetangga Terdekat.

$X_1 =$ Asam Durabilitas (detik)	$X_2 =$ Kekuatan (Kg / meter persegi)	Square Jarak ke contoh permintaan (3, 7)	Peringkat jarak minimum	Apakah termasuk dalam tetangga 3- Terdekat?	$Y =$ Kategori tetangga Terdekat
7	7	$(7-3)^2 + (7-7)^2 = 16$	3	Ya	Buruk
7	4	$(7-3)^2 + (4-7)^2 = 25$	4	Tidak	-
3	4	$(3-3)^2 + (4-7)^2 = 9$	1	Ya	Baik
1	4	$(1-3)^2 + (4-7)^2 = 13$	2	Ya	Baik

#### A. Flowchart

Berikut ini merupakan gambaran flowchart metode klasifikasi algoritma K-NN pada data training dan data testing untuk memprediksi status pelayanan di PT Telkom Indonesia Kab Sidoarjo:





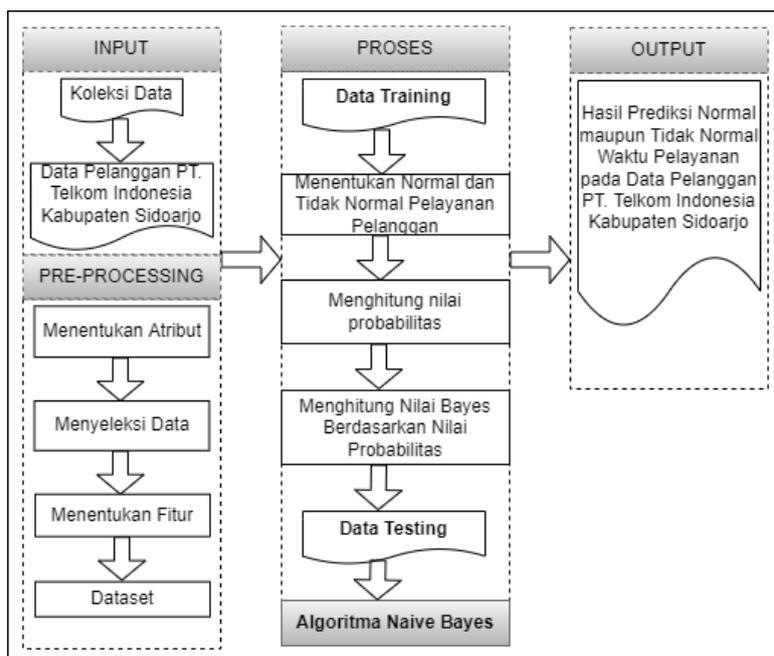
**Gambar 1.** Flowchart Proses K-NN

Dengan adanya gambaran flowchart diatas, Maka tahapan secara umum sesuai klasifikasi algoritma k-nn pada data training sebagai berikut

1. Mulai
2. Seleksi atribut dari NCLI, Nama, TC, KC, Status, TT, Tipe dan Pelayanan menjadi TC, KC, Status, TT, Tipe, Pelayanan
3. Menentukan nilai K yang digunakan
4. Menghitung jarak euclidean
5. Mengurutkan hasil perhitungan jarak dari nilai terkecil ke terbesar
6. Memilih objek yang termasuk dalam nilai K
7. Diperoleh hasil klasifikasi
8. Selesai

#### **B. Diagram Klasifikasi**

Perancangan diagram klasifikasi dengan menggunakan algoritma knn yaitu sebagai berikut:



Gambar 2. Diagram Klasifikasi K-NN

### III. HASIL DAN PEMBAHASAN

#### A. Data Pelanggan Telkom

Data yang digunakan sebagai penelitian ialah data request order dari pelanggan PT. Telkom di Kabupaten Sidoarjo. Data yang akan digunakan untuk memprediksi normal maupun tidaknya waktu pelayanan kepada pelanggan Telkom ialah sebanyak 15.113 data yang diperoleh dari bulan mei sampai dengan juni tahun 2018.

Tabel 5. Data Pelanggan Telkom

No	NCLI	NAMA	TC	KC	Status	TT	Tipe	Pelayanan
1.	37284313	AGUS SETYAWAN	ADA	PCT	Complete	AS	1P	N
2.	35835056	AHMAD KHOIRON	TIDAK	DLA	Complete	MO	1P	N
3.	37656944	ABDUL HUSIN	TIDAK	SPJ	UNSCC	AO	3P	TN
4.	37763048	ARIF ARIANTO	ADA	PDA	Process	AO	3P	TN
5.	37724973	NAWI	ADA	SDA	REVOKE	AO	2P	TN
6.	37624509	SITI AMINAH SITI MURTI	ADA	KRN	Complete	AO	3P	N
7.	32438404	NINGSIH	TIDAK	TUN	Process	MO	2P	TN
8.	37613772	SITI MUZAYANAH	ADA	SDN	REVOKE	AO	3P	TN
9.	30560673	SITI ROHMAH	ADA	BEJ	Complete	MO	1P	N
10.	34227683	SRI ANIFAH	TIDAK	DLA	Complete	MO	1P	N

#### B. Pre-processing

Pada tahap ini dilakukan pre-processing data berupa penentuan atribut dan fitur, seleksi atribut, penentuan dataset, penentuan data training dan penentuan data testing pada data pelanggan Telkom.

##### a) Variabel Dan Fitur

Variabel ialah suatu kumpulan yang terdiri dari bagian entitas. Variabel yang digunakan pada data pelanggan Telkom adalah sebanyak 6 atribut dan memiliki class prediksi yaitu pelayanan. Sedangkan fitur merupakan isi dari variabel. Data yang telah diperoleh dari PT. Telkom Indonesia Kabupaten Sidoarjo dapat diketahui variabel dan fitur yaitu sebagai berikut:

Tabel 6. Variabel Dan Fitur Data Pelanggan Telkom

No.	Atribut	Fitur	Keterangan	Bobot Fitur
1.	Telkom Cabang (TC)	Kategorikal {Ada, Tidak}	Ada atau tidaknya Telkom Cabang	Untuk menentukan ada maupun tidaknya Telkom cabang dan saling berelasi dengan kantor cabang
2.	Kantor Cabang (KC)	Kategorikal {BEJ, DLA, GDA, GEM, KRN, KRP, MIP, MJS, MOJ, MRT, NGI, PCT, PDA, PGE, PWS, SDA, SDN, SPJ, TUN}	Kantor Cabang Daerah : BEJ (Beji), DLA (Dlangu), GDA (Gedangan), GEM (Gempol), KRN (Krian), KRP (Karang Pilang), MIP (Mlirip), MJS (Mojosari), MOJ (Mojokerto), MRT (Kota Mojokerto), NGI (Ngoro Industri), PCT (Pacet), PDA (Pandaan), PGE (Prigen), PWS (Purwosari), SDA (Sidoarjo), SDN (Sukodono), SPJ (Sepanjang), TUN (Tulangan).	Menentukan letak kantor cabang pelanggan Telkom.
3.	Status	Kategorikal {Complete, Fallout, Process, Revoke, UNSCC}	Status Pelayanan : Complete (selesai), Fallout (Pencocokan data di lapangan dengan di kantor), Process (Proses Validasi), Revoke (), UNSCC (Pemasangan ditunda).	Untuk mengukur tingkat ketepatan dalam sebuah pelayanan
4.	Tipe Transaksi (TT)	Kategorikal {AO, AS, MO, MIGRATE}	Tipe Transaksi Pelanggan : AO (Pasang Baru), AS (Tambah Layanan), MO (Tambah atau Kurangi Kecepatan), MIGRATE (Ganti Kabel).	Tipe transaksi pada pelanggan berelasi dengan tipe layanan yang telah dipesan oleh pelanggan Telkom
5.	Tipe	Kategorikal {1P, 2P, 3P}	Tipe layanan yang terdiri dari : 1P : Telepon 2P : Telepon dan internet 3P : Telepon, Internet dan tv kabel	Menentukan tipe layanan pesanan pelanggan
6.	Pelayanan	Kategorikal {Normal, Tidak Normal}	Waktu Pelayanan	

#### b) Seleksi Atribut

Dari data yang telah diperoleh, telah dilakukan tahap seleksi atribut yang akan diolah melalui 6 atribut, diantaranya ialah variabel TC (Telkom Cabang), variabel KC (Kantor Cabang), variabel Status, variabel Tipe Transaksi (TT), variabel Tipe, variabel Pelayanan dan memiliki class prediksi yaitu Pelayanan. Berikut ini ialah hasil dari pre-processing data pada data request order.

Tabel 7. Dataset 100

NO	TC	KC	Status	TT	Tipe	Pelayanan
1.	TIDAK	SDN	Complete	AS	1P	N
2.	ADA	PCT	Complete	AS	1P	N
3.	TIDAK	SDN	Complete	MO	1P	N
4.	TIDAK	DLA	Complete	MO	1P	N
5.	ADA	SDN	Complete	MIGRATE	1P	N
6.	TIDAK	PWS	Complete	MIGRATE	1P	N
7.	TIDAK	SPJ	UNSCC	AO	3P	TN
8.	ADA	SDA	UNSCC	AO	3P	TN
9.	TIDAK	GEM	REVOKE	AO	3P	TN
10.	ADA	SDA	PROCESS	MO	3P	TN
..	..	..	..	..	..	..

Tabel 8. Data Training 60

NO	TC	KC	Status	TT	Tipe	Pelayanan
1.	TIDAK	TUN	Complete	MIGRATE	1P	N
2.	TIDAK	TUN	Complete	MIGRATE	1P	N
3.	ADA	SDA	Complete	MO	3P	N
4.	ADA	SDA	Complete	MO	3P	N
5.	ADA	SDA	Complete	MO	3P	N
6.	ADA	SDA	Complete	MO	3P	N
7.	ADA	SDA	Complete	MO	3P	N
8.	ADA	SDA	Complete	MO	3P	N
9.	ADA	GEM	Complete	AS	2P	N
10.	TIDAK	MIP	Complete	MO	3P	N
..	..	..	..	..	..	..

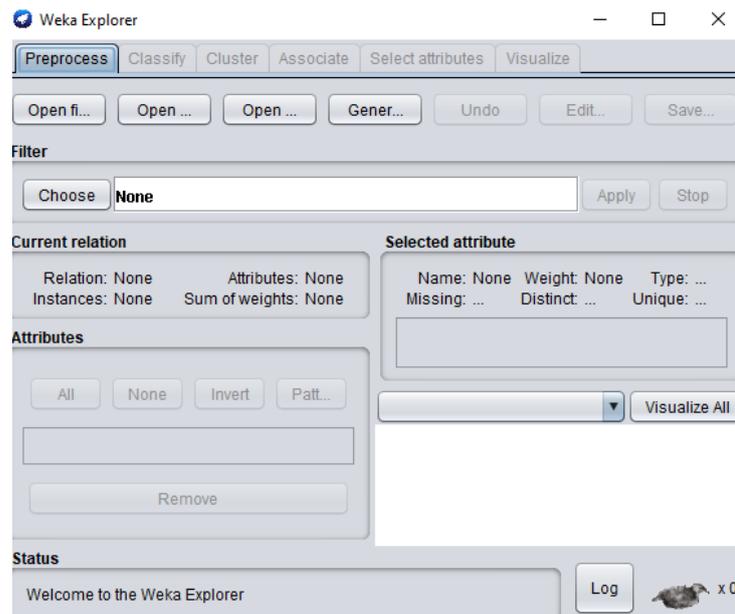
Tabel 9. Data Testing 40

NO	TC	KC	Status	TT	Tipe	Pelayanan
1.	ADA	KRN	Complete	MO	3P	?
2.	ADA	SDA	Complete	MO	1P	?
3.	TIDAK	KRN	Complete	MO	2P	?
4.	TIDAK	SPJ	Complete	MO	1P	?
5.	TIDAK	MJS	Complete	MO	3P	?
6.	ADA	PGE	Complete	MO	1P	?
7.	TIDAK	KRN	Complete	MO	2P	?
8.	TIDAK	SDA	REVOKE	AO	3P	?
9.	TIDAK	MRT	Complete	MO	1P	?
10.	TIDAK	PGE	Complete	MO	1P	?
..	..	..	..	..	..	..

### C. Pembahasan

Pada bagian pembahasan akan dijelaskan tahapan dalam melakukan perhitungan menggunakan metode klasifikasi K-NN untuk memprediksi normal maupun tidak normal status pelayanan pada data pelanggan Telkom. Data pelanggan Telkom tersebut akan diolah dan diuji menggunakan WEKA.

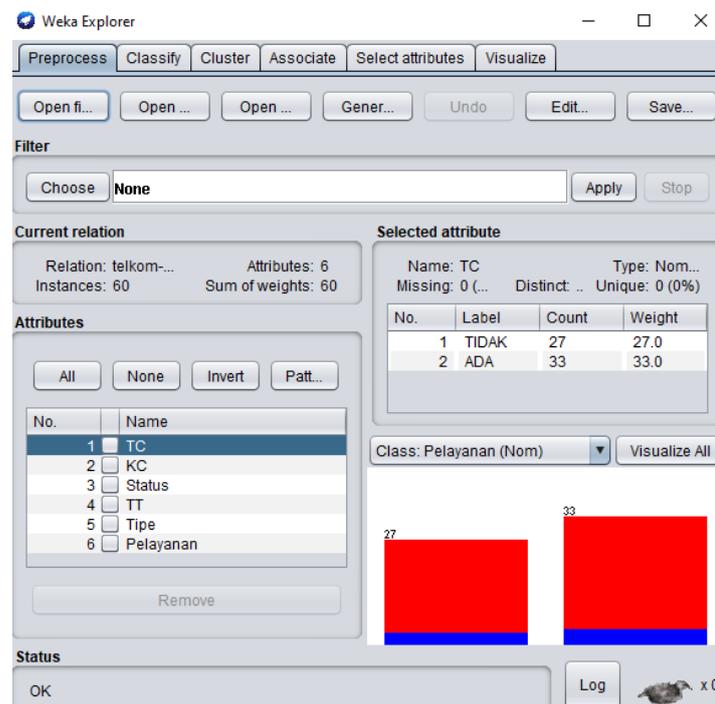
### D. Halaman Menu Explorer



**Gambar 3.** Halaman menu Explorer

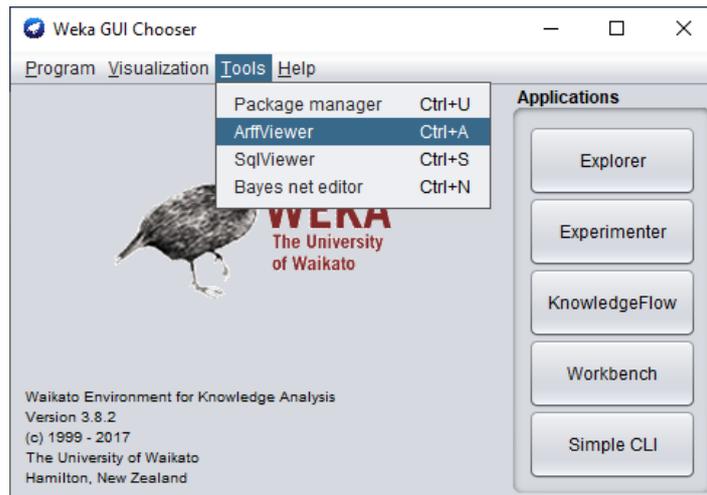
Pada halaman menu explorer terdapat 5 menu pilihan, diantaranya yaitu Preprocess, Classify, Cluster, Associate, Select attribute dan Visualize. Pada tahap pengujian ini akan dilakukan pada menu Preprocess dengan membuka file pada Open File, lalu masukkan data yang akan diuji yaitu data training.

### E. Halaman Atribut Data Training



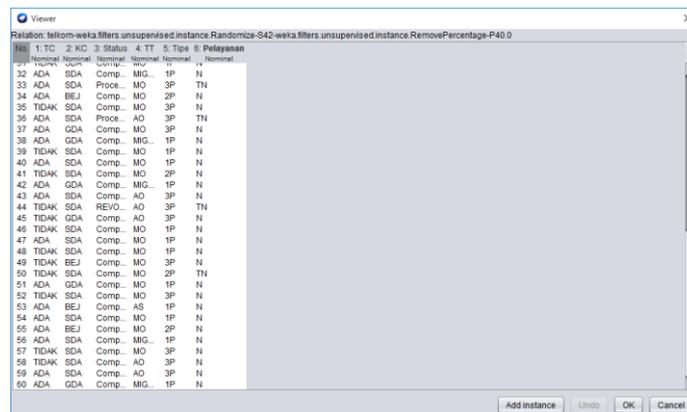
**Gambar 4.** Halaman atribut data training

Setelah memasukkan file, akan muncul atribut yang terdapat dalam data pelanggan. Pada halaman atribut terdapat 6 atribut yang digunakan, diantaranya ialah TC, KC, Status, TT, Tipe dan Pelayanan. Class yang digunakan untuk pengujian yaitu pelayanan dengan nilai ADA sebanyak 33 dan TIDAK sebanyak 27 dari 60 data pelanggan.



**Gambar 5.** Halaman Tools viewer atribut

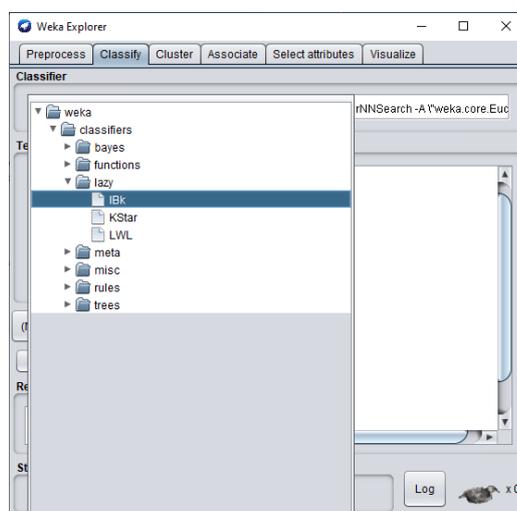
Untuk melihat atribut data training yang telah diubah pada format arff ialah masuk ke halaman awal Weka, masuk ke menu Tools lalu pilih ArffViewer. Pilih File untuk memasukkan data dengan cara klik Open lalu masukkan data yang digunakan.



**Gambar 6.** Halaman Viewer atribut data training

Pada gambar 3.8 menampilkan halaman Viewer yang digunakan untuk melihat atribut data training yang telah diubah pada format arff. Pada halaman tersebut terdapat atribut data training yang digunakan yaitu sebanyak 60 data training.

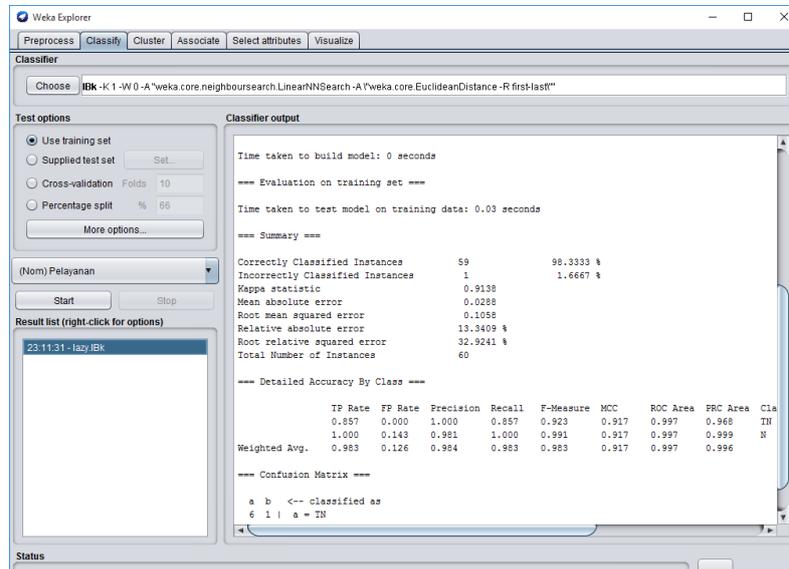
## F. Halaman Menu Klasifikasi



**Gambar 7.** Halaman menu klasifikasi

Pada halaman menu klasifikasi, terdapat bermacam-macam metode klasifikasi data mining. Pengujian klasifikasi data pelanggan ini akan diuji menggunakan metode KNN. Tahap berikutnya ialah memasukkan metode KNN yang akan digunakan yaitu pada bagian Filter Choose pilih folder lazy lalu pilih IBk.

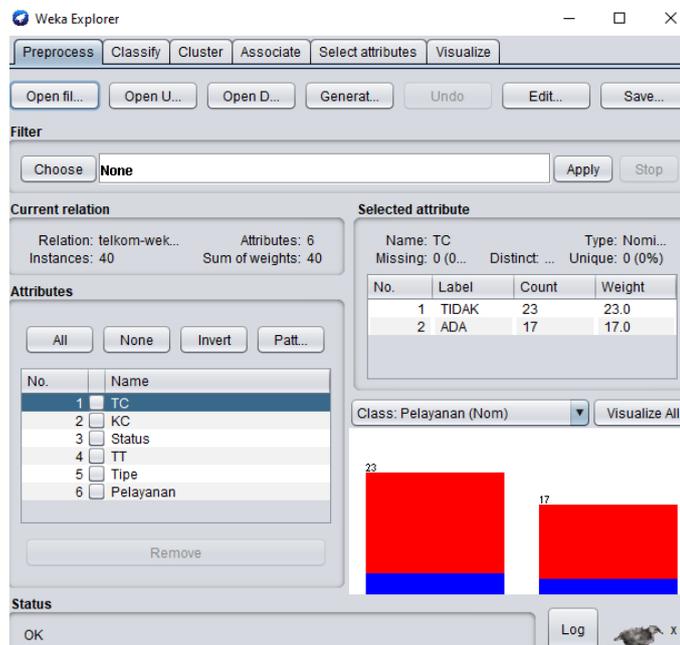
### G. Halaman Klasifikasi Data Training



**Gambar 8.** Hasil Klasifikasi Data Training

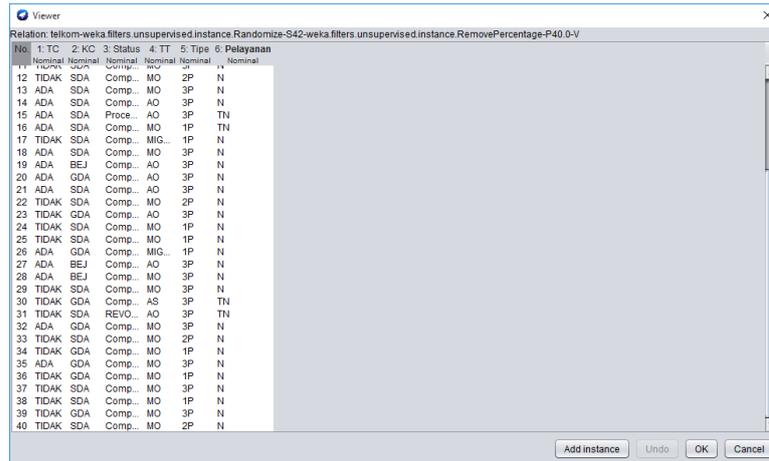
Pada tahap berikut ini pilih Start untuk melihat hasil klasifikasi dari data 60 data training. Dari hasil pengujian diperoleh tingkat kebenaran sebesar 98,34% pada *Correctly Classified Instances* dan tingkat kesalahan sebesar 1,67% pada *Incorrect Classified Instances*.

### H. Halaman Atribut Data Testing



**Gambar 9.** Halaman atribut data testing

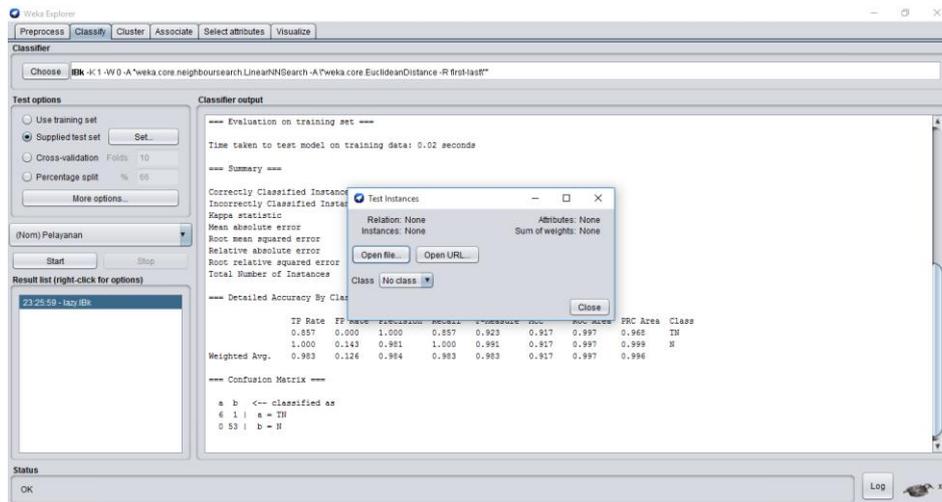
Setelah memasukkan file, akan muncul atribut yang terdapat dalam data pelanggan. Pada halaman atribut terdapat 6 atribut yang digunakan, diantaranya ialah TC, KC, Status, TT, Tipe dan Pelayanan. Class yang digunakan untuk pengujian yaitu pelayanan dengan nilai ADA sebanyak 17 dan TIDAK sebanyak 23 dari 40 data pelanggan.



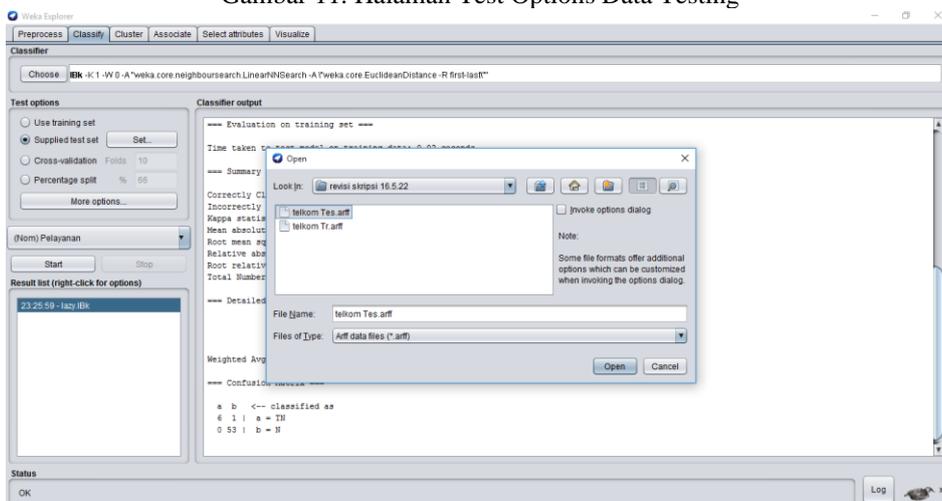
**Gambar 10.** Halaman Viewer atribut data testing

Pada gambar 4.9 menampilkan halaman Viewer yang digunakan untuk melihat atribut data training yang telah diubah pada format arff. Pada halaman tersebut terdapat atribut data training yang digunakan yaitu sebanyak 60 data training.

### I. Pengujian Data Training dengan Data Testing

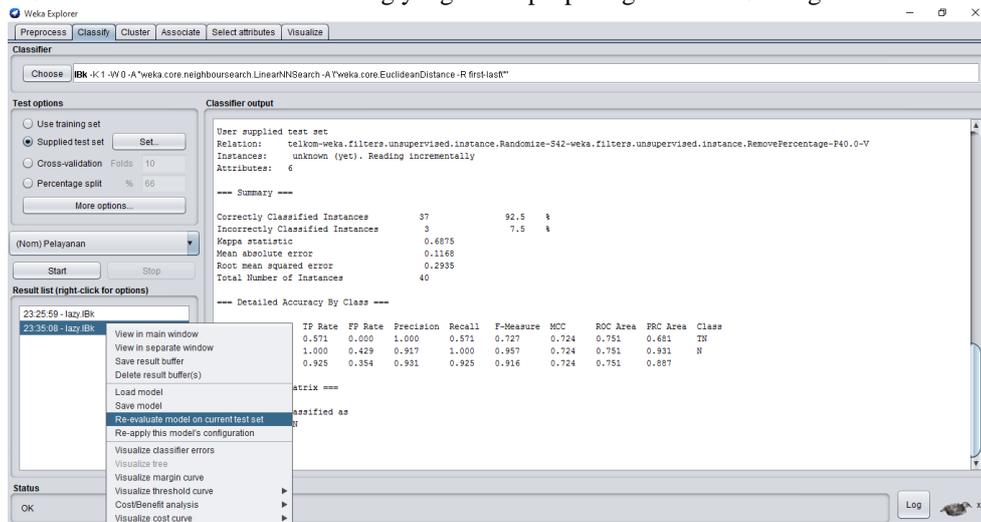


**Gambar 11.** Halaman Test Options Data Testing



**Gambar 12.** Halaman Supplied Test Set Data Testing

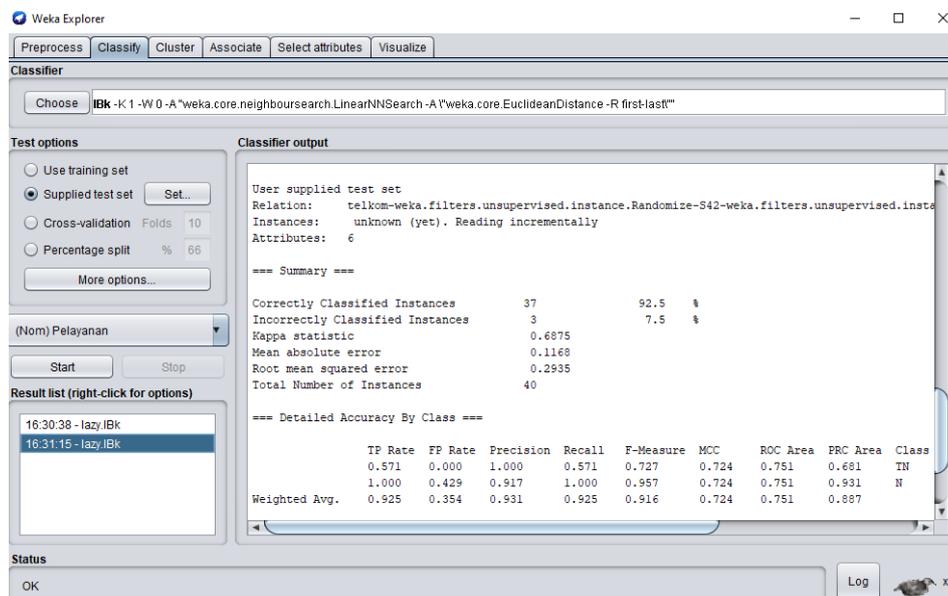
Pada tahap ini pengujian data training akan diuji dengan menggunakan data testing. Pada Test Options pilih Supplied Test Set dan masukkan file data testing yang terlampir pada gambar 4.10 dan gambar 4.11



**Gambar 13.** Proses Pengujian menggunakan data testing

Setelah memasukkan data testing, lalu klik kanan dan pilih Re-evaluate mode on current test set kemudian klik Start untuk melakukan proses pengujian.

**J. Halaman Klasifikasi Data Testing**



**Gambar 14.** Hasil Klasifikasi Data Testing

Pada tahap berikut ini pilih Start untuk melihat hasil klasifikasi dari data 40 data testing. Dari hasil pengujian diperoleh tingkat kebenaran sebesar 92,5% pada *Correctly Classified Instances* dan tingkat kesalahan sebesar 7,5% pada *Incorrect Classified Instances*.

**K. Hasil Persentase**

Dari hasil perbandingan yang diperoleh dari data training sebanyak 60% dan data testing sebanyak 40% berdasarkan dataset sebanyak 100 data yang terdiri dari 6 atribut yaitu Telkom Cabang (TC), Kantor Cabang (KC), Status, TT, Tipe dan Pelayanan dengan menggunakan perhitungan manual dan weka ialah sebagai berikut:

**Tabel 10.** Hasil Persentase Pengujian Dataset

	Dataset	Training	Testing	Benar	% Benar	Salah	% Salah
Website	100	60	40	37	92,5%	3	7,5%
Manual	100	60	40	39	98%	1	3%

Pada dataset 100 menggunakan weka diperoleh prediksi kebenaran sebesar 92,5% dan prediksi kesalahan sebesar 7,5%. Pada dataset 100 menggunakan perhitungan manual diperoleh prediksi kebenaran sebesar 98% dan prediksi kesalahan sebesar 3%. Dari hasil klasifikasi menggunakan metode KNN untuk memprediksi status pelayanan pada 100 dataset, diperoleh hasil tertinggi menggunakan perhitungan manual dengan perolehan persentase sebesar 98%.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Prediksi klasifikasi status pelayanan pada data pelanggan Telkom dilakukan menggunakan metode klasifikasi KNN serta menggunakan 6 variabel yang telah ditentukan.
2. Hasil prediksi pada data pelanggan Telkom dibagi menjadi 100 dataset. Pada setiap dataset sebanyak 60% dari tiap dataset dijadikan sebagai data training dan sebanyak 40% dari tiap dataset dijadikan sebagai data testing.
3. Dari hasil klasifikasi pada metode KNN untuk memprediksi status pelayanan dari 100 dataset, diperoleh hasil tertinggi menggunakan weka dengan perolehan tingkat kebenaran sebesar 92,5% dan tingkat kesalahan sebesar 7,5%.

#### Ucapan Terima Kasih

Terima kasih kepada pihak PT. Telkom Sidoarjo yang telah menyediakan fasilitas sebagai tempat penelitian pada kegiatan ini.

Dr. Hidayatullah, M.Si selaku Rektor Universitas Muhammadiyah Sidoarjo yang telah menyediakan fasilitas pada kegiatan perkuliahan.

Dr. Hindarto, S.Kom., M.T., selaku Dekan Fakultas Sains dan Teknologi yang telah menyediakan fasilitas pada kegiatan perkuliahan.

Ir Sumarno, MM, selaku Kaprodi Informatika yang telah menyediakan fasilitas dalam bagi perkuliahan.

#### REFERENSI

- [1] A. Penyelenggara and J. Internet, "RESPONDEN Survei Nasional PENETRASI Pengguna Internet 2018," 2018.
- [2] D. A. Kesawasidhi, "Pengaruh Kinerja Kualitas Pelayanan Terhadap Kepuasan Dan Loyalitas Pelanggan Pt. Telekomunikasi Indonesia Tbk. Cabang Yogyakarta," *Fak. Ekon. Manaj. Univ. Islam Indones.*, pp. 1–21, 2017.
- [3] A. Wibisono and , S., "Pengaruh Kualitas Jasa Pelayanan Terhadap Kepuasan Pelanggan," *Perform. " J. Bisnis Akuntansi*, vol. 6, no. 2, p. 32, 2016, doi: 10.24929/feb.v6i2.268.
- [4] I. Oktavianis, "Upaya PT. Telkom Dalam Meningkatkan Kualitas Pelayanan Internet Speedy Dalam Memuaskan Pelanggan (Studi Pada PT. Telkom Malang)," *J. Adm. Publik Mhs. Univ. Brawijaya*, vol. 1, no. 1, p. 72569, 2013.
- [5] H. Leidiyana, "Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor," *J. Penelit. Ilmu Komputer, Syst. Embed. Log.*, vol. 1, no. 1, pp. 65–76, 2013.
- [6] S. Sumarlin, "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM," *J. Sist. Inf. Bisnis*, vol. 5, no. 1, pp. 52–62, 2015, doi: 10.21456/vol5iss1pp52-62.
- [7] W. Yustanti, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," *J. Mat. Stat. dan komputasi*, vol. 9, no. 1, pp. 57–68, 2012.
- [8] S. A. Hasan, S. Suhada, and L. Hadjaratie, "Klasifikasi Status Gizi Menggunakan K-Nearest Neighbor," vol. 3, no. 1, pp. 6–11, 2015.
- [9] gunadharna, "Definisi dan Simbol Flowchart," *Defin. Dan Simbol Flowchart*, 2016.
- [10] D. Purnamasari, J. Henharta, Y. P. Sasmita, F. Ihsani, and I. W. S. Wicaksana, "Machine Learning 'Get Easy Using WEKA,'" *Dapur Buku*, pp. 1–40, 2013

#### **Conflict of Interest Statement:**

*The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.*