

IJCIT

(Indonesian Journal on Computer and Information Technology)

Journal Homepage: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijcit>

Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa

Saifur Rohman Cholil¹, Titis Handayani², Rastri Prathivi³, Tria Ardianita⁴

^{1,2,4}Sistem Informasi, Universitas Semarang

Semarang, Indonesia

e-mail: cholil@usm.ac.id¹, titis@usm.ac.id², triaardianita@gmail.com⁴

³Teknik Informatika, Universitas Semarang

Semarang, Indonesia

e-mail: vivi@usm.ac.id

ABSTRAK

Pemberian beasiswa kepada siswa Sekolah Menengah Atas (SMA) sudah umum dilakukan. Hal ini terjadi sejak adanya dana pendidikan 20% dari Kementerian Pendidikan dan Kebudayaan (Kemendikbud). Selain untuk batuan kepada siswa yang kurang mampu, beasiswa juga diberikan kepada siswa yang mempunyai prestasi akademik maupun prestasi non akademik. Pemberian beasiswa yang terjadi selama ini baik di SMA ataupun yang lain masih menggunakan perhitungan dan pengolahan data secara manual. Proses perhitungan secara manual memungkinkan adanya penerima beasiswa yang tidak tepat sasaran. Pengolahan penerimaan beasiswa bisa menggunakan sebuah algoritma data mining untuk mengklasifikasikan calon penerima beasiswa berdasarkan data yang diambil dari data siswa penerima beasiswa sebelumnya (data training) dengan data yang diambil dari calon penerima beasiswa (data testing). Penelitian ini bertujuan membantu proses seleksi beasiswa di SMA menggunakan algoritma K-Nearest Neighbor (KNN) supaya penerima beasiswa tepat sasaran. Algoritma KNN bisa memberikan kebutuhan data yang akurat dan informasi yang diperlukan untuk menyeleksi calon penerima beasiswa. Hasil dari penelitian ini adalah terseleksinya 30 orang dari 89 data yang telah dilakukan klasifikasi. Pengujian sistem menggunakan pengujian akurasi metode *confusion matrix* dengan hasil pengujian sebesar 90.5%. Hal ini menunjukkan bahwa algoritma KNN bisa digunakan untuk mengklasifikasikan seleksi penerimaan beasiswa.

Katakunci: algoritma, beasiswa, data mining, KNN

ABSTRACTS

Providing scholarships to high school students (SMA) is common. This happened since there was a 20% education fund from the Ministry of Education and Culture (Kemendikbud). In addition to rocks to underprivileged students, scholarships are also given to students who have academic and non-academic achievements. Scholarships that have occurred so far both in high school and others still use manual calculation and data processing. The manual calculation process allows for scholarship recipients who are not on target. Processing scholarship receipts can use a data mining algorithm to classify prospective scholarship recipients based on data taken from previous scholarship recipient student data (training data) with data taken from prospective scholarship recipients (data testing). This study aims to help the scholarship selection process in high school using the K-Nearest Neighbor (KNN) algorithm so that scholarship recipients are on target. The KNN algorithm can provide accurate data and information needed to select prospective scholarship recipients. The result of this research is the selection of 30



Jurnal ini dapat diakses secara terbuka dan memiliki lisensi CC-BY-SA

(<https://creativecommons.org/licenses/by-sa/4.0/>) ©2021 by penulis dan IJCIT

people from 89 data that has been classified. System testing uses the accuracy of confusion matrix testing with 90.5% test results. This shows that the KNN algorithm can be used to classify scholarship acceptance selections.

Keywords: algorithms, data mining, KNN, scholarship

1. PENDAHULUAN

Klasifikasi yang didasarkan pada sistem informasi merupakan teknik memetakan (mengklasifikasikan) data ke dalam satu atau beberapa kelas yang sudah didefinisikan sebelumnya. Ada banyak teknik klasifikasi yang dapat digunakan, diantaranya adalah *K-Nearest Neighbor* (KNN). Konsep penelitian dengan algoritma KNN telah banyak dilakukan oleh peneliti-peneliti sebelumnya diantaranya yaitu, algoritma KNN digunakan untuk pengenalan pola, pengenalan teks, pengolahan objek dan lain-lain. Algoritma KNN dianggap mempunyai kesederhanaan dalam pengolahan data training dan data testing dalam jumlah yang sangat besar (Tang, Jing, Li, & Atkinson, 2016). Algoritma KNN mampu melakukan training pada dataset penyakit diabetes untuk melihat dampak negatif hilangnya nilai imputasi dan solusi untuk penyembuhan. Akurasi algoritma KNN diatas rata-rata yang diterapkan untuk klasifikasi penerima beasiswa PPA dan BBM (Sumarlin, 2016).

Algoritma KNN merupakan metode yang digunakan untuk melakukan klasifikasi data berdasarkan jarak terpendek terhadap objek data. Penentuan nilai K yang terbaik untuk algoritma ini berdasarkan pada data yang ada. Nilai K yang tinggi dapat mengurangi efek noise pada klasifikasi, bisa juga membuat batasan antara setiap klasifikasi menjadi lebih kabur (Anshori, Regasari, & Putri, 2018). Algoritma KNN merupakan algoritma berbasis contoh atau non parametric dan dianggap metode paling sederhana di dalam proses data mining (Tharwat, Mahdi, Elhoseny, & Hassanien, 2018). Algoritma KNN salah satu metode klasifikasi data yang mudah diimplementasi pada jumlah data yang kecil, tetapi jika dataset yang diolah banyak dan kompleks maka algoritma KNN memiliki kelemahan dan waktu yang tidak efisien (Yahya & Hidayanti, 2020). Metode KNN menggunakan ukuran jarak yang sesuai untuk mengklasifikasikan data baru. Jarak tetangga terdekat K dihitung dan label kelas dari tetangga terdekat diprediksi sebagai label kelas dari instance baru. Akurasi KNN sangat terpengaruh

dengan memilih jumlah K tetangga terdekat. Jika terdapat nilai K yang kecil, maka akan sensitif terhadap *noise* dan jika terlalu besar, dapat menyebabkan bias model (Anshori et al., 2018). Algoritma KNN merupakan algoritma berbasis memori yang menggunakan iterasi pada data sampai atribut atau parameter data yang terdekat ditemukan. Jarak minimum data yang diproses pada data *testing* akan dibandingkan dengan data *training* dengan jarak yang terdekat (Suwirmayanti, 2017).

Algoritma KNN digunakan untuk memprediksi harga jual tanah dengan cara menentukan pilihan dalam proses pemilihan tanah yang strategis dengan harga yang sesuai (R Novita & Harsani, 2018). Algoritma KNN digunakan untuk prediksi penjualan furniture sehingga dapat mengurangi waktu pemrosesan dan algoritma tersebut bisa memberikan akurasi yang baik dalam sampel pengujian yang telah dilakukan (Hutami & Astuti, 2016). Algoritma KNN digunakan untuk mengklasifikasikan pemasaran dengan hasil dan akurasi menggunakan metode *cross validation* dengan hasil yang baik (Mustakim, 2016).

Pemberian beasiswa kepada siswa SMA sudah umum dilakukan. Beasiswa diberikan kepada siswa yang tidak mampu, siswa yang mempunyai prestasi akademik maupun non akademik. Pemberian beasiswa yang terjadi selama ini masih menggunakan perhitungan dan pengolahan data secara manual. Proses perhitungan secara manual memungkinkan adanya penerima beasiswa yang tidak tepat sasaran. Pengolahan penerimaan beasiswa bisa menggunakan sebuah algoritma data mining untuk mengklasifikasikan calon penerima beasiswa berdasarkan data training yang diambil dari data siswa penerima beasiswa sebelumnya (dataset) dengan data testing yaitu data yang diambil dari calon penerima beasiswa.

Penelitian ini bertujuan membantu proses seleksi beasiswa di SMA menggunakan algoritma *K-Nearest Neighbor* (KNN) supaya penerima beasiswa tepat sasaran. Keunikan yang terdapat dalam penelitian ini adalah mampu memetakan atau mengklasifikasikan data menjadi beberapa kelas yang didefinisikan sebelumnya.

2. METODE PENELITIAN

Beberapa cara digunakan untuk pengukuran jarak antara data baru (data testing) dengan data lama (data training), antara lain *manhattan distance* (*city block distance*) dan *euclidean distance*. Pengukuran jarak yang paling sering digunakan adalah *euclidean distance*. Jarak *euclidean distance* didefinisikan seperti pada persamaan 1 (Suwirmayanti, 2017).

$$d_i = \sqrt{\sum_{i=1}^p (x_1 - x_2)^2} \quad (1)$$

Keterangan :

- d = jarak
- i = variabel data
- p = dimensi data
- x_1 = sampel data
- x_2 = data uji

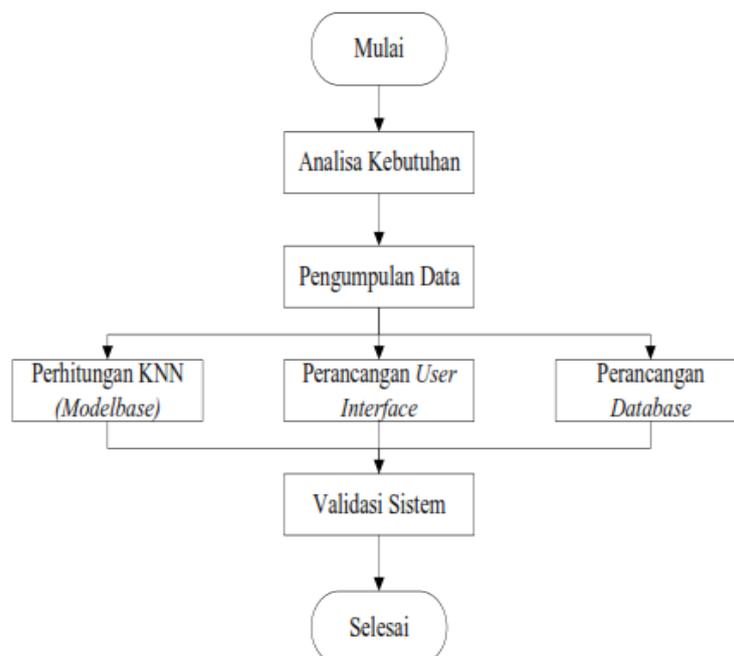
Proses perhitungan nilai akurasi dapat dilakukan dengan menggunakan persamaan 2.

$$akurasi = \frac{\text{jumlah klasifikasi benar}}{\text{jumlah data uji}} \times 100\% \quad (2)$$

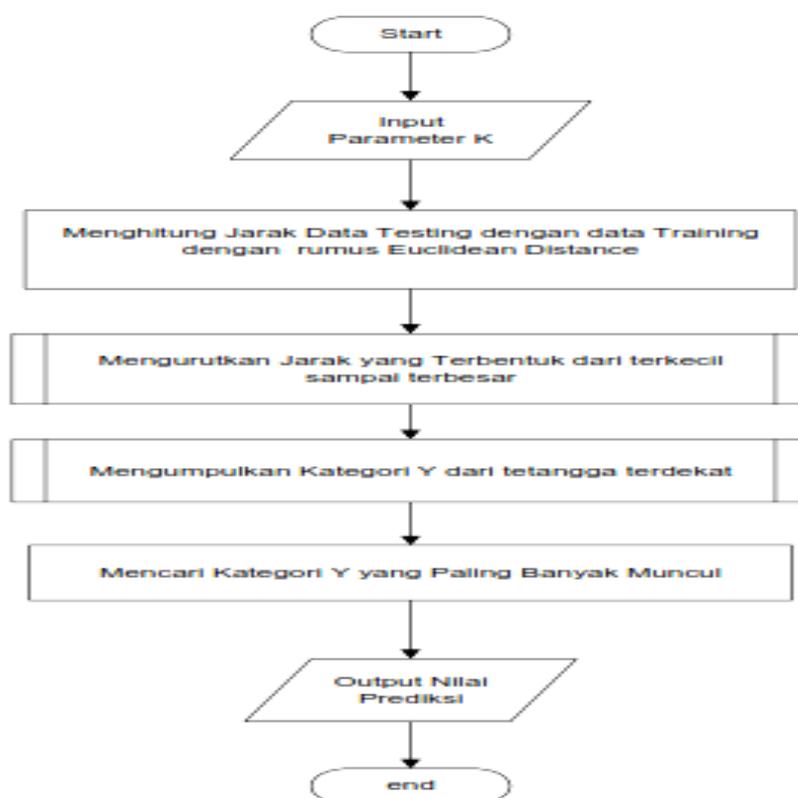
Algoritma KNN memiliki beberapa kelebihan diantaranya ketangguhan terhadap data training yang memiliki banyak noise dan data dalam jumlah yang besar. Kelemahan dari

algoritma KNN adalah perlu menentukan jumlah tetangga terdekat dari target data, yang disimbolkan dengan nilai parameter K, data training yang didasarkan pada hasil perhitungan jarak kurang akurat karena harus memilih, mencoba dan menentukan jenis jarak yang digunakan dan atribut mana yang dipakai untuk mendapatkan hasil yang perhitungan jarak yang terbaik, serta biaya komputasi yang tinggi karena membutuhkan perhitungan jarak dari tiap *query instance* pada keseluruhan contoh data training.

Algoritma KNN merupakan salah satu metode *supervised learning* karena hasil *query instance* perlu diklasifikasikan berdasarkan mayoritas dari kategori KNN. Pengumpulan data menggunakan metode observasi. Data primer diperoleh dari bagian kepegawaian SMA Islam Sultan Agung I Semarang. Data primer yang dijadikan contoh dalam penelitian ini adalah data siswa calon penerima beasiswa tahun ajaran 2017/2018 dan tahun ajaran 2018/2019. Tahap perancangan penelitian yang dilakukan untuk memberikan gambaran secara jelas dapat ditunjukkan pada gambar 1. Dalam penelitian ini menggunakan data training yang berjumlah 101 data. Dan data testing berjumlah 89 data. Untuk metode pengolahan data menggunakan metode KNN untuk klasifikasi. Algoritma KNN dapat dilihat pada Gambar 2 berikut ini.



Gambar 1. Tahapan Penelitian



Gambar 2. Flowchart Algoritma KKN

Berikut penjelasan setiap tahapan dari algoritma KNN: (1). Menentukan jumlah tetangga terdekat yang disimbolkan dengan nilai parameter K. Nilai pada parameter K yang akurat untuk algoritma ini tergantung pada data training yang digunakan; (2). Menghitung dan menentukan nilai kuadrat jarak data testing objek terhadap data training yang diberikan dengan menggunakan persamaan 1; (3). Melakukan pengurutan dari hasil perhitungan no 2 secara ascending (urut dari nilai rendah ke nilai tinggi); (4). Mengumpulkan data pada kategori Y (Klasifikasi tetangga terdekat berdasarkan nilai K); (5). Kategori Y yang paling banyak muncul menjadi hasil akhir dari klasifikasi.

3. HASIL DAN PEMBAHASAN

3.1. Pengolahan Data Awal

Data calon penerima beasiswa yang terlihat seperti Tabel 1 adalah data mentah yang belum diolah dengan metode KNN. Data penerima beasiswa Tahun Akademik 2017/2018 dijadikan sebagai data *training* sedangkan data pendaftar beasiswa Tahun Akademik 2018/2019

seperti yang terlihat pada Tabel 2 dijadikan sebagai data *testing*.

Pengisian kolom nilai raport didasarkan pada rata-rata nilai raport semester Gasal dan nilai raport semester Genap. Penilaian kolom Prestasi akademik diisi berdasarkan 5 parameter yaitu Kurang berprestasi, Prestasi tingkat sekolah, Prestasi tingkat kota, Prestasi tingkat propinsi dan Prestasi tingkat nasional. Pengisian nilai prestasi akademik ditunjukkan oleh Tabel 3. Prestasi akademik didasarkan pada prestasi siswa dalam mengikuti lomba-lomba akademik yang diselenggarakan baik tingkat sekolah, tingkat kota, tingkat propinsi maupun tingkat nasional. Diantaranya ada lomba olympiade sains, cerdas cermat, lomba mata pelajaran, lomba membaca puisi dan lain-lain.

Penilaian kolom prestasi non akademik diisi berdasarkan 4 parameter yaitu Kurang berprestasi, Prestasi tingkat kota, Prestasi tingkat propinsi dan Prestasi tingkat nasional. Pengisian nilai prestasi non akademik ditunjukkan oleh Tabel 4. Prestasi non akademik didasarkan pada prestasi siswa dalam mengikuti lomba-lomba non akademik yang diselenggarakan baik tingkat kota, tingkat propinsi maupun tingkat nasional. Diantaranya

ada lomba pramuka, lomba olahraga lomba kepemimpinan dan lain-lain.

Pengisian kolom penghasilan orangtua didasarkan pada pendapatan orangtua setiap bulannya. Penilaian tersebut diisi berdasarkan 5 parameter yaitu Penghasilan diatas 4 juta/bulan, Penghasilan 3-4 juta/bulan, Penghasilan 2-3 juta/bulan, Penghasilan 1-2 juta/bulan dan Penghasilan kurang dari 1 juta/bulan. Pengisian

nilai penghasilan orangtua ditunjukkan oleh Tabel 5.

Pengisian kolom kepribadian didasarkan pada sikap kepribadian selama di sekolah. Penilaian tersebut diisi berdasarkan lima parameter yaitu Sangat kurang, Kurang, Cukup, Baik dan Sangat baik. Pengisian nilai kepribadian ditunjukkan oleh Tabel 6.

Tabel 1. Data Training

No	NIS	Nama	Nilai Raport	Akademik	Non Akademik	Penghasilan Ortu	Kepribadian	Status Beasiswa
1	15291	Abdigusti R. M.	83,38	3	1	2	4	Tidak Beasiswa
2	15429	Achmad Alfarizi	83,06	2	1	3	4	Tidak Beasiswa
3	15361	Akmal F.	82,69	2	1	2	4	Tidak Beasiswa
4	15432	Alfian R.	82,50	2	2	4	4	Tidak Beasiswa
5	15435	Arif R.	82,44	2	3	4	4	Beasiswa
6	15364	Arya Aldi	82,38	2	1	1	4	Tidak Beasiswa
7	15438	Dharma N. P.	82,19	2	1	2	3	Tidak Beasiswa
8	15299	Dogan S.	82,19	2	1	3	3	Tidak Beasiswa
9	15367	Fadhil S.	82,00	2	2	3	5	Tidak Beasiswa
10	15303	Garuda B. N.	81,81	2	3	4	5	Beasiswa
11	15303	Hamzah A.	81,75	2	1	5	4	Tidak Beasiswa
12	15309	Humam M.	81,69	2	1	1	3	Tidak Beasiswa
13	15370	Husen T. A.	81,63	2	1	3	4	Tidak Beasiswa
14	15371	Irfan R.Putra	81,31	2	1	5	3	Tidak Beasiswa
15	15312	Jovian Chandara	74,88	1	1	5	3	Tidak Beasiswa
16	15642	M. Fariz Sukma	74,81	1	2	4	5	Tidak Beasiswa
17	15375	Maskharis	74,75	1	1	2	4	Tidak Beasiswa
18	15378	Mohamad Adrian	74,75	1	1	2	4	Tidak Beasiswa
19	15441	Mohamad Akmal	74,69	1	1	2	3	Tidak Beasiswa
20	15448	M. Faiz S.	74,54	1	1	1	4	Tidak Beasiswa

Tabel 2. Data Testing

No	NIS	Nama	Nilai Raport	Akademik	Non Akademik	Penghasilan Ortu	Kepribadian
1	15861	Adi Haryanto	75,23	3	1	2	4
2	15862	Afiandani N.	80,00	2	1	3	4
3	15863	Ahmad Kristiyan	74,23	2	1	2	4
4	15864	Ahmad Thariq Z.	77,56	2	2	4	4
5	15865	Athaya Adeer	76,00	2	3	4	4
6	15866	BAyu Adi L.	88,25	2	1	1	4
7	15867	Bima Adi S.	75,45	2	1	2	3
8	15868	Chandra Yudha	74,00	2	1	3	3
9	15869	Daffa Tsani	72,35	2	2	3	5
10	15870	Davin Yusuf	72,66	2	3	4	5
11	15871	Ginangjar	76,40	2	1	5	4
12	15872	Irfan A.	81,00	2	1	1	3
13	15873	Krishan Yudha	75,20	2	1	3	4
14	15874	Lufi Darmawa	70,45	2	1	5	3
15	15875	M. Deva Prabwo	85,00	1	1	5	3
16	15876	M. Fadil	74,80	1	2	4	5
17	15877	Maskharis	72,50	1	3	3	5
18	15878	Mohamad Rafi	73,15	1	1	2	4
19	15879	Mohamad Reza	84,23	1	1	2	3
20	15880	Rizqi Fadillah	81,45	1	1	4	3

Tabel 3. Prestasi akademik

Parameter	Nilai
Kurang Berprestasi	1
Prestasi Tingkat Sekolah	2
Prestasi Tingkat Kota	3
Prestasi Tingkat Propinsi	4
Prestasi Tingkat Nasional	5

Tabel 4. Prestasi Non Akademik

Parameter	Nilai
Kurang Berprestasi	1
Prestasi Tingkat Kota	2
Prestasi Tingkat Propinsi	3
Prestasi Tingkat Nasional	4

Tabel 5. Parameter Penghasilan Orangtua

Parameter	Nilai
Penghasilan > 4 juta/bulan	1
Penghasilan 3-4 juta/bulan	2
Penghasilan 2-3 juta/bulan	3
Penghasilan 1-2 juta/bulan	4
Penghasilan < 4 juta/bulan	5

Tabel 6. Parameter Nilai Kepribadian

Parameter	Nilai
Sangat kurang	1
Kurang	2
Cukup	3
Baik	4
Sangat baik	5

3.2. Perhitungan *Euclidean Distance*

Data penerima beasiswa Tahun Akademik 2017/2018 dijadikan sebagai data training sedangkan data pendaftar beasiswa Tahun Akademik 2018/2019 dijadikan sebagai data testing. Proses perhitungan euclidean distance dilakukan dengan menggunakan rumus persamaan 1 seperti contoh perhitungan di bawah ini dan hasilnya ditunjukkan oleh Tabel 7 dengan nilai $K=30$. Data testing yang digunakan sebagai contoh perhitungan adalah data dengan nomor 1, NIS: 15858 atas nama Adi Hariyanto.

$$d_i = \sqrt{(83,38 - 75,23)^2 + (3 - 3)^2 + (1 - 1)^2 + (2 - 2)^2 + (4 - 4)^2} = 8,15$$

Tabel 7. Perhitungan *Euclidean Distance*

No	NIS	Nama	Nilai E D	Sorting	KNN	Klasifikasi
1	15291	Abdigusti Rachman Maulana	8,150	64	Tidak	-
2	15429	Achmad Alfarizi	7,957	63	Tidak	-
3	15361	Akmal Fikri Nugroho	7,527	59	Tidak	-
4	15432	Alfian Reza Fahluzi Yusuf	7,672	61	Tidak	-
5	15435	Arif Roshyan Nuruddin	7,809	62	Tidak	-
..
..
..
10	15394	Zhilal Algeiba	5,264	35	Tidak	-

Perhitungan dari nilai *euclidean distance* selanjutnya diurutkan berdasarkan tetangga terdekat dari nilai K . Berikutnya dilakukan klasifikasi apakah termasuk mendapatkan beasiswa atau tidak berdasarkan nilai K . Dengan menggunakan data testing nomor 1 dengan NIS: 15858 atas nama Adi Hariyanto yang dikategorikan tidak mendapatkan beasiswa maka didapatkan hasil perhitungan metode KNN dengan $K=30$, siswa yang mendapatkan beasiswa tidak ada, sedangkan siswa yang tidak mendapatkan beasiswa sebanyak 30 orang. Jika data testing diubah berdasarkan data nomor 56 dengan NIS: 15913 atas nama Orisza Sativa maka didapatkan hasil siswa yang mendapatkan

beasiswa sebanyak 22 orang dan yang tidak mendapatkan beasiswa sebanyak 8 orang.

3.3. Implementasi Sistem

Implementasi sistem adalah proses penerapan ke dalam bentuk program pada perangkat lunak. Berikut implementasi yang dilakukan pada perangkat lunak bentuk antarmuka aplikasi dan server programnya. Program menggunakan bahasa python dan tampilan web menggunakan Django. Implementasi program yang pertama kali dilakukan adalah menyalakan server pada command windows seperti pada Gambar 2.

Selanjutnya buka browser untuk memanggil file seperti pada Gambar 3.

3.4. Evaluasi Metode KNN

Pada penelitian ini telah dilakukan evaluasi atau pengujian metode KNN dengan menghitung nilai akurasi metode KNN dengan metode *confusion matrix*. Hasil dari evaluasi ini dapat terlihat pada gambar 4 dan gambar 5.

Pada gambar 5 dan gambar 6 menunjukkan hasil dari perhitungan KNN dengan pengujian menggunakan metode *confusion matrix* pada data training. Sedangkan gambar 7 dan gambar 8 merupakan hasil pengujian KNN dengan *confusion matrix* pada data testing. Hasil evaluasi perbandingan pengujian data *training* dan data testing ditunjukkan pada tabel 7.

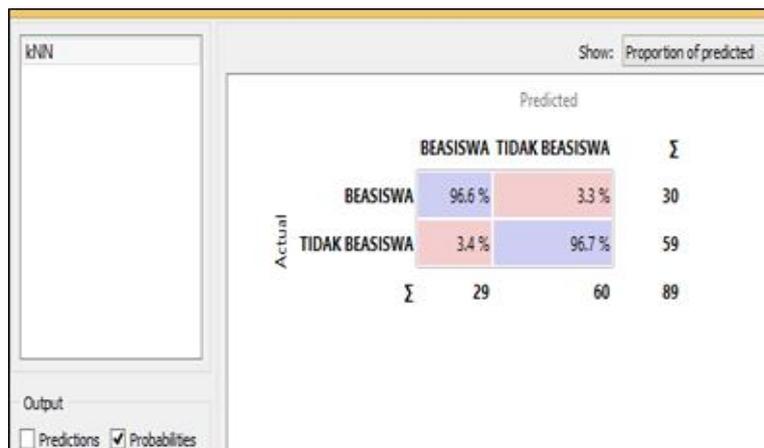
```

C:\WINDOWS\system32\cmd.exe - manage.py runserver
File "c:\python34\Lib\socketserver.py", line 673, in __init__
  self.handle()
File "c:\python34\Lib\wsgiref\handlers.py", line 144, in run
  self.close()
File "D:\knn\env\lib\site-packages\django\core\servers\basehttp.py", line 154, in handle
  handler.run(self.server.get_app())
File "c:\python34\Lib\wsgiref\handlers.py", line 144, in run
  self.close()
File "c:\python34\Lib\wsgiref\simple_server.py", line 35, in close
  self.status.split(' ',1)[0], self.bytes_sent
AttributeError: 'NoneType' object has no attribute 'split'
-----
File "c:\python34\Lib\wsgiref\simple_server.py", line 35, in close
  self.status.split(' ',1)[0], self.bytes_sent
File "c:\python34\Lib\wsgiref\simple_server.py", line 35, in close
  self.status.split(' ',1)[0], self.bytes_sent
AttributeError: 'NoneType' object has no attribute 'split'
-----
AttributeError: 'NoneType' object has no attribute 'split'
-----
[19/Aug/2019 16:10:47] "POST /pengabdian/result HTTP/1.1" 200 84732
  
```

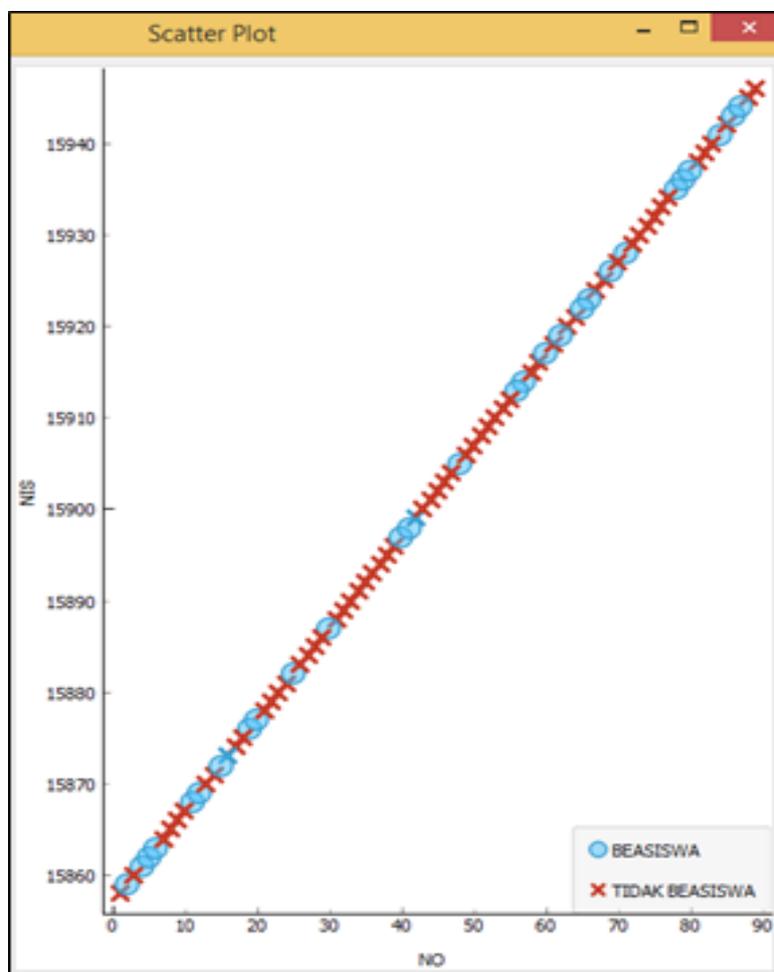
Gambar 2. Server Aplikasi

NIS	Nama	Kelas	Nilai Raport	Prestasi Akademik	Prestasi Non Akademik	Penghasilan Orang Tua	Kepribadian	Hasil
15858	ADI HARIWANTO	X-MIPA.1	75.23	3.0	1.0	2.0	4.0	TIDAK BEASISWA
15839	AFIANDANNI NAUFAL R	X-MIPA.1	80.0	2.0	1.0	3.0	4.0	TIDAK BEASISWA
15860	AHMAD KRISTANTO	X-MIPA.1	74.23	2.0	1.0	2.0	4.0	TIDAK BEASISWA
15861	AHMAD THARIQ ZHAFFRAN	X-MIPA.1	77.56	2.0	2.0	4.0	4.0	TIDAK BEASISWA
15862	ATHAYA ADEEV RABBANI	X-MIPA.2	76.0	2.0	3.0	4.0	4.0	TIDAK BEASISWA
15863	BAYU ADI LAKSONO	X-MIPA.2	88.25	2.0	1.0	1.0	4.0	BEASISWA

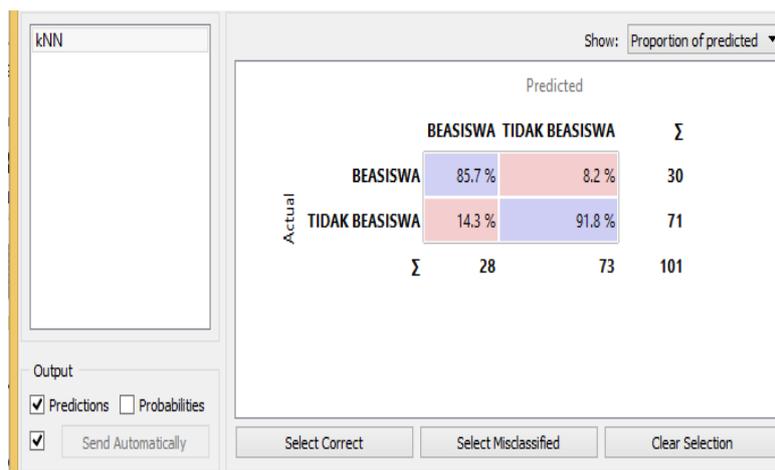
Gambar 3. Aplikasi KNN



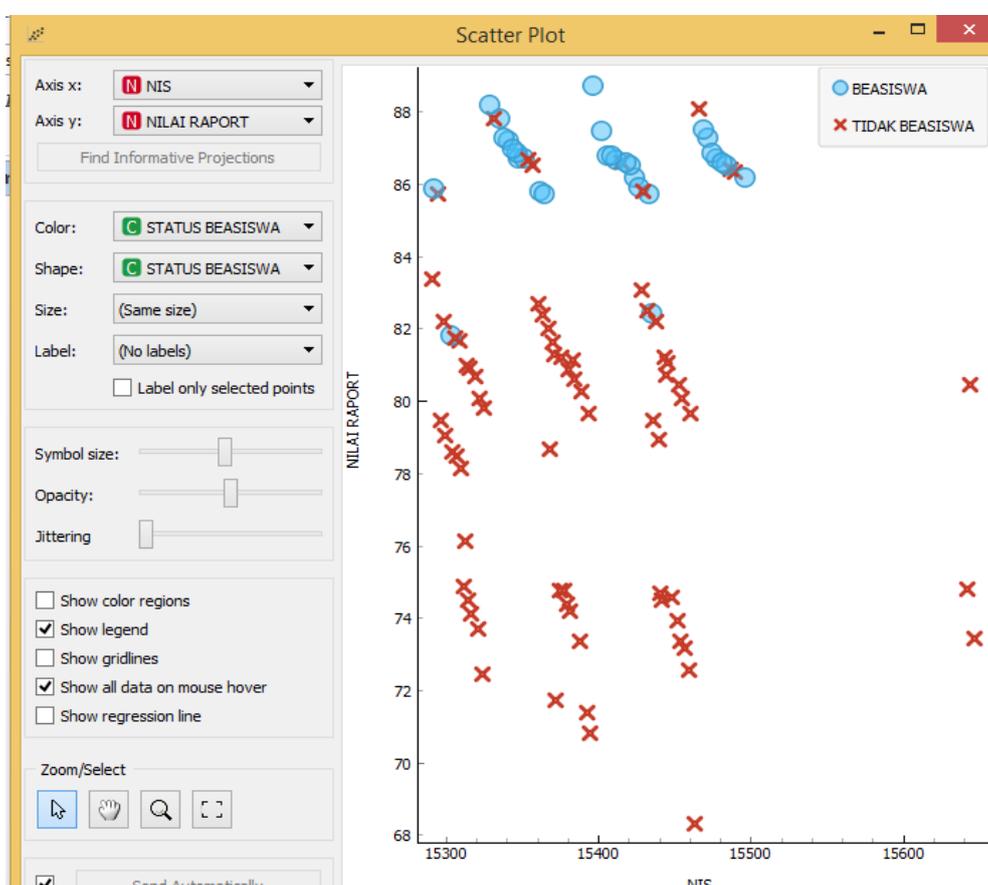
Gambar 5. Evaluasi Dengan *Confusion Matrix* Pada Data Training



Gambar 6. Hasil *Scatter Plot* Data Training



Gambar 7. Evaluasi Dengan Confusion Matrix Pada Data Testing



Gambar 8. Hasil scatter plot data testing

Tabel 7. Perbandingan hasil evaluasi data training dan data testing

Method	AUC	CA	F1	Precision	Recall	Sampling	Data
KNN	1	97%	95%	97%	93%	Cross Validation	Training
	97%	89%	82%	81%	83%		Training

Hasil dan perbandingan dengan penelitian lain adalah metode ini telah banyak digunakan dalam menyelesaikan suatu masalah. Misalnya pada penelitian yang melibatkan KNN untuk merekomendasikan laptop yang menunjukkan kepuasan terhadap hasil rekomendasi itu sebesar 85% (Raharja, Juardi, & Agung, 2019). Hal ini menunjukkan bahwa metode KNN dapat memberikan hasil yang akurat dan dapat dipertanggungjawabkan.

4. KESIMPULAN

Kesimpulan dari penelitian ini terciptanya aplikasi untuk seleksi data penerima beasiswa dengan menggunakan algoritma KNN. Evaluasi algoritma KNN menggunakan metode *confusion matrix* menunjukkan hasil dari perhitungan rata-rata akurasi dari metode KNN ini sebesar 90,5%. Evaluasi perbandingan data training dan testing pada metode KNN dengan *cross validation sampling* memperlihatkan hasil perhitungan rata – rata *precision* sebesar 89%.

5. REFERENSI

- Anshori, L., Regasari, R., & Putri, M. (2018). Implementasi Metode K-Nearest Neighbor untuk Rekomendasi Keminatan Studi (Studi Kasus : Jurusan Teknik Informatika Univ *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(7), 2745–2753.
- N. L. G. P. Suwirmayanti. (2017) "Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil," *Techno. Com*, vol. 16, no. 2, pp. 120–131.
- Hutami, R., & Astuti, E. Z. (2016). Implementasi Metode K-Nearest Neighbor Untuk Prediksi Penjualan Furniture Pada CV.Octo Agung Jepara. *Universitas Dian Nuswantoro Semarang*.
- Mustakim, G. O. F. (2016). Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa, *13(2)*, 195–202.
- Sumarlin, S. (2016). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM. *Jurnal Sistem Informasi Bisnis*, 5(1), 52–62. <https://doi.org/10.21456/vol5iss1pp52-62>
- Suwirmayanti, N. L. G. P. (2017). Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil. *Techno. Com*, 16(2), 120–131.
- Tang, Y., Jing, L., Li, H., & Atkinson, P. M. (2016). A multiple-point spatially weighted k-NN method for object-based classification. *International Journal of Applied Earth Observation and Geoinformation*, 52, 263–274. <https://doi.org/10.1016/j.jag.2016.06.017>
- Tharwat, A., Mahdi, H., Elhoseny, M., & Hassanien, A. E. (2018). Recognizing human activity in mobile crowdsensing environment using optimized k-NN algorithm. *Expert Systems with Applications*, 107, 32–44. <https://doi.org/10.1016/j.eswa.2018.04.017>
- Y. Yahya and W. Puspita Hidayanti (2020). "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada 'Lombok Vape On,'" *Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 104–114, 2020, doi: 10.29408/jit.v3i2.2279.
- A. Q. Sesilia Novita R, Prihastuti Harsani. (2018) . "Penerapan K-Nearest Neighbor (KNN) untuk Klasifikasi Anggrek Berdasarkan Karakter Morfologi Daun dan Bunga," vol. 15, no. 1, pp. 118–125.
- C. A. Rahardja, T. Juardi, and H. Agung. (2019). "Implementasi Algoritma K-Nearest Neighbor Pada Website Rekomendasi Laptop," *J. Buana Inform.*, vol. 10, no. 1, p. 75, 2019, doi: 10.24002/jbi.v10i1.1847.