

Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Peminatan Studi STMIK Amik Riau

Nora Lizarti ¹⁾*, Aniq Noviciatie Ulfah ²⁾

STMIK Amik Riau ^{1), 2)}

noralizarti@stmik-amik-riau.ac.id ¹⁾*, aniqnoviciatie@stmik-amik-riau.ac.id ²⁾

Abstrak

Peminatan studi pada STMIK Amik Riau merupakan pilihan minat berdasarkan kemampuan khusus dan ketertarikan mahasiswa. Program studi Teknik Informatika STMIK Amik Riau memiliki dua bidang peminatan, yaitu bisnis dan jaringan. Peminatan disesuaikan dengan kemampuan dan ketertarikan dari mahasiswa serta harus dipilih dengan baik dan tepat. Pengambilan peminatan sangat berpengaruh terhadap tugas akhir dan tingkat kelulusan mahasiswa. Pemilihan peminatan studi oleh mahasiswa saat ini hanya mengikuti teman dan tidak berdasarkan kemampuan, sehingga sebuah Sistem klasifikasi peminatan merupakan salah satu solusi untuk menyelesaikan permasalahan pemilihan peminatan pada program studi karena dianggap mampu memberikan rekomendasi peminatan yang baik dan tepat. Algoritma K-Nearest Neighbor (K-NN) merupakan algoritma klasifikasi yang dapat digunakan sebagai solusi dalam pengelompokan data. Pada penelitian ini data yang digunakan diperoleh dari nilai mata kuliah prasyarat selama semester satu hingga semester lima. Data diolah dengan membangun aplikasi yang menerapkan algoritma K-NN menggunakan PHP dan MySQL. Hasil keluaran sistem memiliki akurasi 100% dibandingkan hasil perhitungan manual menggunakan Ms. Excel. Pengujian menggunakan tools RapidMiner untuk mengukur performa algoritma. Hasil pengujian yang dilakukan terhadap 183 data latih dan 100 data uji menyatakan algoritma K-NN memiliki performa dengan hasil Accuracy, Recall, Precision, F Measure, dan Classification Error dengan nilai 98%, 100%, 100%, 91.67%, dan 2%. Penelitian ini dapat memberikan rekomendasi peminatan studi kepada mahasiswa Teknik Informatika STMIK Amik Riau.

Kata kunci: Peminatan studi, klasifikasi, algoritma K-NN

Abstract

[Implementation of K-Nearest Neighbor Algorithm For Determining Concentration of Study at Informatics Engineering Program of STMIK AMIK RIAU] Concentration of study at STMIK Amik Riau is a choice of interests based on special abilities and student interests. Informatics Engineering Program at STMIK Amik Riau has two subjects of interest, namely business and networking. The study concentration is tailored to the abilities and interests of students and must be chosen properly and correctly because it is very influential on the final assignment and graduation level of students. The classification system of interest is one of the solutions to solve the problem of choosing a concentration in the study program because it is considered capable of providing good and appropriate spinning recommendations. K-Nearest Neighbor (K-NN) is one of classification algorithm that can be used as a solution in classifying data. In this study, the data used was obtained from the value of prerequisite courses during semester one to semester five. Data is processed by building applications that implement the K-NN algorithm using PHP and MySQL. The output of the system has 100% accuracy compared to the results of manual calculations using Microsoft Excel. The Testing process used RapidMiner software to measure algorithm performance. The results of the tests carried out on 183 training data and 100 test data stated that the K-NN algorithm had a performance with the results of Accuracy, Recall, Precision, Measure, and Classification Error with values of 98%, 100%, 100%, 91.67%, and 2%. This study can provide a system that can help to give some study concentration recommendations to the student of Informatics Engineering Program at STMIK Amik Riau.

Keywords: Study Interest, Classification, K-Nearest Neighbor (K-NN) Algorithm

1. PENDAHULUAN

STMIK Amik Riau adalah perguruan tinggi komputer pertama di Riau yang selalu menjaga eksistensi dan kualitas lulusan. Hal tersebut ditunjang oleh STMIK Amik Riau dengan memberikan mata kuliah peminatan yang bertujuan untuk memberikan ciri khas dan keahlian khusus kepada mahasiswanya. Peminatan studi pada perguruan tinggi STMIK Amik Riau merupakan pilihan minat berdasarkan kemampuan khusus dan ketertarikan mahasiswa terhadap pilihan konsentrasi yang ditawarkan oleh program studi. Program studi Teknik Informatika STMIK Amik Riau memiliki dua bidang peminatan, yaitu peminatan bisnis dan peminatan jaringan.

Peminatan studi yang diambil sebaiknya disesuaikan dengan kemampuan dan ketertarikan dari mahasiswa. Pengaruh pengambilan peminatan studi harus dipilih dengan baik dan tepat. Berdasarkan penelitian [1] dari 49 responden yang digunakan menyatakan bahwa keterbatasan informasi mengenai bidang peminatan menjadi salah satu faktor sulitnya mahasiswa menentukan peminatan yang sesuai, hal berpengaruh terhadap tugas akhir dan tingkat kelulusan mahasiswa. Pengambilan peminatan yang salah dapat mengakibatkan kegagalan dan kemunduran mahasiswa dalam menyelesaikan tugas akhirnya, dari 500 lulusan ada 16% lulusan yang kurang sesuai dalam pemilihan peminatan[2]. Pemilihan peminatan studi oleh mahasiswa saat ini hanya mengikuti teman dan tidak berdasarkan kemampuan, sehingga sebuah sistem klasifikasi peminatan merupakan salah satu solusi untuk menyelesaikan permasalahan pemilihan peminatan pada program studi Teknik Informatika.

Algoritma klasifikasi dianggap mampu mempertimbangkan peminatan mahasiswa dalam mengambil peminatan yang baik dan tepat. Algoritma klasifikasi merupakan pengelompokan data atau objek baru berdasarkan variabel yang diamati dengan tujuan untuk memprediksi suatu objek dari yang masih belum diketahui kelas atau kategorinya [3]. Hasil perhitungan yang menerapkan algoritma ini dapat memberikan informasi kemampuan akademik mahasiswa melalui hasil keluaran sistem berdasarkan data nilai mahasiswa yang berupa skor nilai pada setiap peminatan.

Data yang digunakan pada penelitian ini menggunakan data numerik yang didapat dari mahasiswa melalui nilai mata kuliah prasyarat selama semester satu hingga semester lima. Data yang dipelajari adalah data mata kuliah prasyarat dan mata kuliah pilihan program studi Teknik Informatika. Data tersebut dibutuhkan untuk mempelajari pola-pola hubungan dari satu mata kuliah prasyarat dengan mata kuliah lainnya sehingga membentuk satu peminatan yang utuh. Data tersebut yang kemudian diolah menggunakan algoritma klasifikasi guna mendapatkan hasil keluaran peminatan mahasiswa.

Dalam penelitian ini, akan dilakukan proses sampling dengan objek, iterasi, data uji dan data latihan

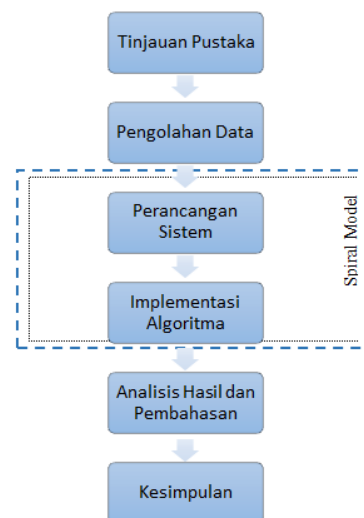
yang sama, sehingga nantinya menghasilkan tingkat akurasi dari algoritma. Penelitian ini menggunakan PHP dan MySQL sebagai tools dalam proses *scripting/programming* dalam pembuatan sistem yang mengimplementasikan algoritma knn, dan kemudian hasil dari implementasi algoritma tersebut akan dibandingkan dengan hasil perhitungan secara manual. Selain itu, dalam melakukan analisa perbandingan *performance* dari algoritma tersebut digunakan *tools* RapidMiner. Berdasarkan hal diatas tujuan penelitian ini akan menerapkan algoritma klasifikasi K-NN guna mengetahui apakah algoritma K-NN dapat menyelesaikan permasalahan peminatan studi mahasiswa STMIK Amik Riau serta diharapkan penelitian ini mampu memberikan rekomendasi kepada mahasiswa dalam memilih peminatan studi yang benar dan tepat.

2. BAHAN DAN METODE

Kerangka kerja dalam penelitian ini sebanyak 6 tahapan yang digambarkan pada Gambar 1 berikut ini:

2.1 Tinjauan Pustaka

Tinjauan pustaka dilakukan terhadap penelitian yang terkait dengan teori-teori tentang *data mining*, metode klasifikasi, algoritma K-NN, sistem pendukung keputusan, peminatan studi, Spiral model, PHP, MySQL dan perhitungan validasi.



Gambar 1. Kerangka Kerja Penelitian

2.2 Pengolahan Data

Merupakan tahapan dalam mempersiapkan data yang diperoleh (data mentah) diubah menjadi data siap diolah disimpan dalam format .xlsx (Ms. Excel). Data yang digunakan pada penelitian ini diperoleh dari Pangkalan Data Perguruan Tinggi (PDPT) STMIK Amik Riau berupa data nilai mahasiswa angkatan 2014 dan 2015[4].

Data Set

Data yang digunakan nilai mahasiswa pada semester 1 sampai dengan 5. Nilai mahasiswa yang diambil terdiri dari 12 matakuliah yang terkait dengan peminatan yang ada yaitu peminatan jaringan dan

bisnis. Data yang diperoleh dibagi menjadi dua bagian yaitu data latih dan data uji.

Preprocessing Data

Pembersihan data: pada tahap ini akan dilakukan proses menghilangkan data-data yang tidak sesuai, data-data yang tidak lengkap, menghaluskan noisy data dan memecahkan redundansi yang disebabkan oleh integrasi data. Data yang diperoleh sebelum dilakukan pembersihan total data mahasiswa sebesar 435, setelah dilakukan pembersihan data diperoleh data mahasiswa sebanyak 283.

Transformasi data: dari data nilai mahasiswa yang berupa data angka yang ditampilkan pada Tabel 1 dibawah ini.

Tabel 1. Tranformasi Nilai

Nilai Huruf	Nilai Angka
A	4,00
A-	3,75
B+	3,50
B	3,00
B-	2,75
C+	2,50
C	2,00
D	1,00
E	0

2.3 Perancangan Sistem

Perancangan sistem merupakan tahap dimana akan dilakukan identifikasi kebutuhan sistem dalam hal ini adalah *software* dan *hardware* serta *desain interface* yang nantinya akan digunakan dalam tahap implementasi algoritma klasifikasi agar mempermudah perhitungan dan mempercepat perolehan hasil klasifikasi. Model Perancangan perangkat lunak spiral model adalah model proses peranti lunak *evolusioner* yang berpasangan dengan iterasi *prototype* dengan aspek yang terkontrol dan sistematis yang diadaptasi dari model *waterfall* [5]. Adapun tahapan dari Spiral model adalah *Requirement, Designing, Develop, dan Testing*[6]. Tahap perancangan sistem merupakan tahap *Requirement Planning* pada model Spiral untuk pengembangan perangkat lunak yang dibangun.

Perancangan Sistem Machine Learning

Pada tahap ini merupakan tahap *Design planning* pada metode Spiral. Proses yang dilakukan pada tahap ini adalah merancang kebutuhan *design program*, data dan *behaviour* yang digunakan oleh program. Adapun perancangan sistem adalah sebagai berikut: perancangan *database relasional*: merancang penyimpanan data dalam konseptual model. Database yang dibangun terdiri atas 3 tabel; perancangan *usecase diagram*: digambarkan akan perilaku serta tugas-tugas dari setiap elemen atau actor; perancangan struktur menu: perancangan menu-menu pada aplikasi sesuai dengan fungsi masing-masing; Perancangan antarmuka (*interface*) merupakan tahapan merancang atau mendesain tampilan antarmuka (*interface*) sistem yang akan membuat interaksi antara pengguna (*user*) dengan sistem. Tahap ini pada model Spiral

merupakan tahapan *Develop*. Dimana pada tahap ini dilakukan proses *coding* sistem yang akan dibangun berdasarkan tahap *design* sistem yang sudah dibuat sebelumnya. Terdapat dua menu yang disajikan dalam sistem yang dibangun yaitu menu Home dan Menu implementasi Algoritma K-NN.

2.4 Implementasi Algoritma

Pada tahap ini dilakukan proses implementasi algoritma K-NN dalam klasifikasi peminatan studi mahasiswa teknik informatika STMIK Amik Riau pada sistem yang dibangun. Proses implementasi algoritma merupakan tahap *testing* pada model Spiral. Pada tahap ini sistem sudah menerapkan algoritma K-NN dimana merupakan algoritma yang digunakan untuk melakukan klasifikasi terhadap suatu objek, berdasarkan k buah data latih yang jaraknya paling dekat dengan objek tersebut. Syarat nilai k adalah tidak boleh lebih besar dari jumlah data latih, dan nilai k harus ganjil dan lebih dari satu. Nilai k yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antar setiap klasifikasi menjadi lebih kabur [7]. Langkah-langkah untuk menghitung metode Algoritma K-NN[8]:

- Menentukan parameter K (Jumlah tetangga paling dekat).
- Menghitung kuadrat jarak *euclid (queri instance)* masing-masing objek terhadap data sampel yang diberikan
- Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak Euclid terkecil.
- Mengumpulkan kategori Y (Klasifikasi *Nearest Neighbor*)
- Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *queri instance* yang telah dihitung.

Hasil dari penerapan algoritma K-NN pada sistem kemudian akan dibandingkan dengan pergitungan manual algoritma K-NN menggunakan Ms. Excel.

Berikut Gambar 2 tahapan implementasi algoritma K-NN .

Pada proses ini pengguna memilih jumlah k yang ini digunakan pada proses implementasi algoritma K-NN. Proses klasifikasi diproses setelah pengguna menekan button "klasifikasi" pada sistem. Penerapan algoritma K-NN dijalankan dengan menerapkan perhitungan jarak tetangga terdekat dengan persamaan

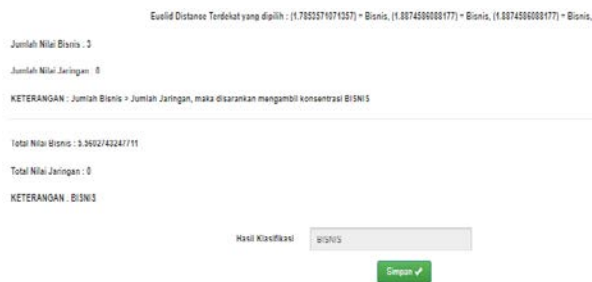
$$d(x, y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (1)$$

dimana d (x, y) adalah jarak antara titik pada data latih x dan titik data uji y yang akan diklasifikasi, dimana $x=x_1, x_2, \dots, x_i$ dan $y=y_1, y_2, \dots, y_i$ dan i merepresentasikan nilai atribut serta n merupakan dimensi atribut, serta menghasilkan klasifikasi dari data set yang diinputkan. Contoh hasil klasifikasi

penerapan algoritma K-NN dengan $k = 3$ dan data set yang dapat dilihat pada Gambar 3.



Gambar 2. Alur Proses Implementasi Algoritma



Gambar 3. Hasil Klasifikasi Penerapan Algoritma K-NN

Dari gambar diatas dapat dilihat bahwa algoritma K-NN menerapkan perhitungan jarak *Ecludian Distance* yang menghasilkan tiga nilai tetangga terdekat dengan data uji dengan nilai secara urut sebagai berikut 1.79, 1.89, dan 1.89 dengan label data secara urut adalah bisnis, bisnis, dan bisnis. Dari data diperoleh hasil tiga data yang kesemuanya masuk pada kelas bisnis dan tidak ada (nol) satupun yang masuk kelas jaringan. Jumlah kelas bisnis dengan jumlah sama dengan tiga lebih banyak daripada kelas jaringan dengan jumlah nol sehingga hasil klasifikasi data uji tersebut yaitu “BISNIS”.

3. HASIL DAN PEMBAHASAN

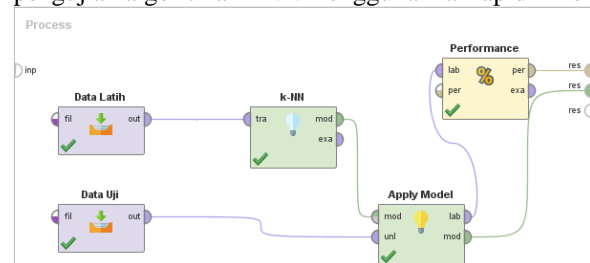
Dataset yang digunakan dalam penelitian berdasarkan nilai mahasiswa dari semester 1 samapai 5 dari 12 mata kuliah yang terkait peminatan jaringan dan bisnis. Dari 283 dibagi menjadi 183 sebagai data uji dan 100 sebagai data latih. Data uji dan data latih yang ada kemudiakan dihitung dan diselesaikan menggunakan algortima K-NN. Perhitungan dan pengelompokkan data peminatan dilakukan dengan dua metode yaitu dengan manual dengan

menggunakan Ms. Excel dan sistem berbasis Web. Hasil klasifikasi keluaran sistem dibandingkan dengan hasil perhitungan algoritma K-NN secara manual menggunakan Ms. Excel. Dari hasil perbandingan perhitungan manual dan sistem didapatkan kemiripan atau akurasi sebesar 100% dengan hasil keluaran sistem yang menerapkan algoritma K-NN.

Berdasarkan hasil pengujian algoritma K-NN. Tahap selanjutnya adalah proses menghitung validasi data yang diperoleh. Validasui data menggunakan operator *data access, modeling, scoring, dan validation*. *Data access* digunakan untuk membaca data yang akan diujikan yaitu data latih dan data uji. *Modeling* merupakan algoritma pengujian pada aplikasi yang akan kita gunakan untuk menguji data yang ada. Setelah data diterapkan pada algoritma K-NN maka perlu adanya penerapan model dan melakukan *scoring*.

Tahap selanjutnya yaitu perhitungan validasi data yaitu akurasi merupakan persentase ketepatan *record* data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi[7] dan semakin tinggi level akurasi, maka dapat dikatakan semakin efektif model algoritma klasifikasi tersebut[9]; *recall* adalah perbandingan jumlah dokumen relevan yang terambil sesuai dengan *query* yang diberikan dengan total kumpulan dokumen yang relevan dengan *query*[10]; *precision* adalah perbandingan jumlah dokumen yang relevan terhadap *query* dengan jumlah dokumen yang terambil hasil pencarian[10]; *f measure* nilai yang didapatkan dari pengukuran *precision* dan *recall* antara *class* hasil *cluster* dengan *class* sebenarnya yang terdapat pada data masukan[11]; *classification error* nilai *error* yang diperoleh dari data set[9] dari algoritma klasifikasi. Pengujian algoritma dilakukan menggunakan aplikasi RapidMiner. Pengujian algoritma dapat dijelaskan sebagai berikut:

Pengujian algoritma K-NN menggunakan aplikasi RapidMiner dengan menggunakan data latih dan data uji yang ada. Gambar 4berikut merupakan proses pengujian algoritma K-NN menggunakan RapidMiner.



Gambar 4. Pengujian Algoritma K-NN menggunakan RapidMiner

Gambar 4. diatas merupakan proses pengujian algoritma K-NN menggunakan RapidMiner. Dimana operator yang digunakan pada pegujian ini adalah *data access, modeling, scoring, dan validation*. *Data access* digunakan untuk membaca data yang akan diujikan yaitu data latih dan data uji. *Modeling* merupakan algoritma pengujian pada aplikasi yang akan kita

gunakan untuk menguji data yang ada. Setelah data diterapkan pada algoritma K-NN maka perlu adanya penerapan model dan melakukan *scoring*. Tahap selanjutnya yaitu perhitungan validasi data yaitu

Accuracy

Hasil persentase tingkat akurasi algoritma K-NN dapat dilihat pada Tabel 2. dibawah ini:

Tabel 2. Hasil Persentase Tingkat Akurasi Algoritma K-NN

K (iterasi)	Accuracy (%)
1	70
3	72
4	74
5	80
6	84
7	95
8	93
9	98
10	96
11	92
12	93
13	91
19	89
21	88
22	90
23	87
25	88
27	85
44	85

Dapat dilihat pada Tabel 2. diatas bahwa persentase tingkat akurasi algoritma K-NN yang paling rendah adalah 70% dengan $k = 1$ dan persentase akurasi tertinggi mencapai 98% dengan $k = 9$.

Recall

Hasil Persentase tingkat *recall* untuk algoritma K-NN dapat dilihat pada Tabel 4.5 sebagai berikut.

Tabel 3. Hasil Persentase Tingkat *Recall* Algoritma K-NN

K (iterasi)	Recall
1	69.23
3	61.54
4	69.23
5	84.62
6	100
7	92.31
8	84.62
9	84.62
10	76.92
11	61.54
12	61.54
13	53.85
19	38.46

K (iterasi)	Recall
21	30.77
22	46.15
23	38.46
25	23.08
27	7.69
44	7.69

Berdasarkan pada Tabel 3. diatas bahwa persentase tingkat *recall* algoritma K-NN yang paling rendah dengan nilai 7.69% adalah $k = 27$ dan 44 dan persentase tertinggi mencapai 100% dengan $k = 6$.

Precision

Hasil Persentase tingkat *precision* untuk algoritma K-NN dapat dilihat pada Tabel 4. sebagai berikut.

Tabel 4. Hasil Persentase Tingkat *Precision* Algoritma K-NN

K (iterasi)	Precision
1	25.71
3	25.81
4	29.03
5	37.93
6	44.83
7	75
8	68.75
9	100
10	90.91
11	72.73
12	80
13	70
19	62.5
21	57.14
22	66.67
23	50
25	60
27	25
44	25

Dapat dilihat pada Tabel 4 diatas bahwa persentase tingkat *precision* algoritma K-NN yang paling rendah dengan nilai 25% adalah $k = 27$ dan 44 dan persentase tertinggi mencapai 100% dengan $k = 9$.

F measure

Hasil Persentase tingkat *f measure* untuk algoritma K-NN dapat dilihat pada Tabel 5. sebagai berikut.

Tabel 5. Hasil Persentase Tingkat *F Measure* Algoritma K-NN

K (iterasi)	f measure
1	37.5
3	36.36
4	40.91
5	52.38
6	61.9

K (iterasi)	f measure
7	82.76
8	75.86
9	91.67
10	83.33
11	66.67
12	69.57
13	60.87
19	47.62
21	40
22	54.55
23	43.48
25	33.33
27	11.76
44	11.76

Dapat dilihat pada Tabel 5. diatas bahwa persentase tingkat *f measure* algoritma K-NN yang paling rendah adalah 11.76% dengan $k = 27$ dan 44 dan persentase tingkat *f measure* tertinggi mencapai 91.67% dengan $k = 9$

Classification Error

Hasil Persentase tingkat *classification error* untuk algoritma K-NN dapat dilihat pada Tabel 6. sebagai berikut.

Tabel 6. Hasil Persentase Tingkat *Classification Error* Algoritma K-NN

K (iterasi)	Classification Error
1	30
3	28
4	26
5	20
6	16
7	5
8	7
9	2
10	4
11	8
12	7
13	9
19	11
21	12
22	10
23	13
25	12
27	15
44	15

Dapat dilihat pada Tabel 6. diatas bahwa persentase tingkat *classification error* algoritma K-NN yang paling rendah adalah 2% dengan $k = 9$ dan persentase tingkat *classification error* tertinggi mencapai 30% dengan $k = 1$.

4. KESIMPULAN

Berdasarkan hasil dan proses analisis yang dilakukan dalam penelitian ini, maka dapat ditarik

simpulan yaitu sistem yang dibangun dapat melakukan klasifikasi data peminatan studi menggunakan algoritma K-NN mahasiswa Teknik Informatika STMIK Amik Riau dengan tingkat kesesuaian terhadap algoritma K-NN dengan hasil perhitungan secara manual, yang memiliki tingkat kesamaan sebesar 100%, dan validasi data dihitung menggunakan Rapid Miner dengan nilai *Acuracy*, *Recall*, *Precision*, *F Measure*, dan *Clasificassion Error* dengan nilai 98%, 100%, 100%, 91.67%, dan 2% serta sistem yang dibangun dapat digunakan untuk rekomendasi peminatan studi kepada mahasiswa Teknik Informatika STMIK Amik Riau.

5. DAFTAR PUSTAKA

- [1] A. C. Febryanti, I. Darmawan, and R. Andreswari, "Pemodelan Sistem Pendukung Keputusan Pemilihan Bidang Peminatan Menggunakan Metode Simple Additive Weighting Studi Kasus: Program Studi Sistem Informasi Universitas Telkom Modelling Of Decision Support System For Fields Of Interest Selection With Simple," in *e-Proceeding of Engineering*, 2017, vol. 4, no. 2, pp. 3114–3121.
- [2] D. L. Widyastomo and R. S. Perdana, "Implementasi Fuzzy K-Nearest Neighbour (FK-NN) Untuk Pemilihan Keminatan Mahasiswa Teknik Informatika (Studi Kasus: Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4414–4420, 2018.
- [3] T. H. Simanjuntak and W. F. Mahmudy, "Implementasi Modified K-Nearest Neighbor Dengan Otomatisasi Nilai K Pada Pengklasifikasian Penyakit Tanaman Kedelai," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, no. June, 2017.
- [4] PDPT, "Data Nilai Mahasiswa STMIK Amik Riau," Pekanbaru, 2015.
- [5] R. S. Pressman, *Software Engineering: A Practitioner's Approach*, 7th ed. New York, USA: Mc Graw Hill Higher Education, 2010.
- [6] P. Saxena and M. Kaushik, "Advantages and Limitation of Different SDLC Models.," *Int. J. Technol. Res. Eng.*, vol. I(3), pp. 117–121, 2013.
- [7] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*. Waltham: Morgan Kaufmann, 2012.
- [8] M. A. Banjarsari, H. I. Budiman, and A. Farmadi, "Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4,," *Kumpul. J. Ilmu Komput.*, vol. 2, no. 2, pp. 50–64, 2015.
- [9] P. Mittal and N. S. Gill, "A Comparative

- Analysis Of Classification Techniques On Medical Data Sets,” *IJRET Int. J. Res. Eng. Technol.*, pp. 454–460, 2014.
- [10] D. Kurniawan, “Evaluasi Sistem Temu Kembali Informasi Model Ruang Vektor,” *J. Sains MIPA*, vol. 16, no. 3, pp. 155–162, 2010.
- [11] I. M. K. Wicaksana and I. M. Widiartha, “Penerapan Metode Ant Colony Optimzation Pada Metode K-Harmonic Means Untuk Klasterisasi Data,” *J. Ilmu Komput.*, vol. 5 No 1, 2012.