

**Penentuan Jurusan Kuliah Siswa Sekolah Menengah Kejuruan disesuaikan  
Dengan Minat menggunakan Algoritma K-Nearest Neighbor di SMKN1  
Gadingrejo Tahun Pelajaran 2018/2019**

---

# Penentuan Jurusan Kuliah Siswa Sekolah Menengah Kejuruan disesuaikan Dengan Minat menggunakan Algoritma K-Nearest Neighbor di SMKN1 Gadingrejo Tahun Pelajaran 2018/2019

Margi Prasojo, Yogi Saputra

STMIK IBI Darmajaya<sup>1), 2)</sup>  
[noralizarti@stmik-amik-riau.ac.id](mailto:noralizarti@stmik-amik-riau.ac.id)<sup>1)</sup>\*, [anignovicatie@stmik-amik-riau.ac.id](mailto:anignovicatie@stmik-amik-riau.ac.id)<sup>2)</sup>

## Abstrak

Peminatan studi di SMKN1 Gadingrejo Pringsewu merupakan pilihan minat berdasarkan kemampuan khusus dan ketertarikan siswa. SMKN1 Gadingrejo memiliki tujuh program studi peminatan, yaitu prodi Teknik Kendaraan Ringan, Prodi Desain Permodelan dan Informasi Bangunan, Teknik Kontruksi Bangunan, Teknik Elektronika Industri, Teknik Audio Video, Teknik Multi Media dan Teknik Animasi. Penentuan peminatan program studi ini disesuaikan dengan kemampuan dan ketertarikan siswa. Pengambilan peminatan sangat berpengaruh terhadap tingkat kelulusan siswa. Pemilihan program studi oleh siswa saat ini hanya mengikuti teman dan tidak berdasarkan kemampuan, sehingga diperlukan sebuah Sistem klasifikasi peminatan sebagai alternative 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 salahsatu algoritma klasifikasi yang dapat digunakan sebagai solusi dalam pengelompokan data. Pada penelitian ini data yang digunakan diperoleh dari nilai mata pelajaran selama semester satu hingga semester lima. Selanjutnya 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 Acuracy, Recall, Precision, F Measure, dan Clasificassion Error dengan nilai 98%, 100%, 100%, 91.67%, dan 2%. Penelitian ini dapat memberikan rekomendasi peminatan studi kepada calon siswa SMKN1 Gadingrejo.

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

## Abstract

[Determination of Vocational High School Student Lectures according to Interests using the K-Nearest Neighbor Algorithm at SMKN1 Gadingrejo Academic Year 2018/2019] The study interest at SMKN 1 Gadingrejo Pringsewu is a choice of interest based on the special abilities and interests of students. SMKN1 Gadingrejo has seven specialization study programs, namely Light Vehicle Engineering Study Program, Modeling Design and Building Information Study Program, Building Construction Engineering, Industrial Electronics Engineering, Audio Video Engineering, Multi Media Engineering and Animation Engineering. The determination of the specialization of this study program is adjusted to the abilities and interests of the students. The choice of specialization greatly influences the passing rate of students. Selection of study programs by students currently only follows friends and is not based on ability, so a specialization classification system is needed as an alternative solution to solving the problem of choosing a specialization in a study program because it is considered capable of providing good and appropriate specialization recommendations. The K-Nearest Neighbor (K-NN) algorithm is a classification algorithm that can be used as a solution in data grouping. In this study, the data used were obtained from the value of the subjects during the first semester to the fifth semester. Furthermore, the data is processed by building applications that apply the K-NN algorithm using PHP and MySQL. The system output has an accuracy of 100% compared to the results of manual calculations using Ms. Excel. Testing uses Rapid Miner tools to measure algorithm performance. The results of tests carried out on 183 training data and 100 test data indicate that the K-NN algorithm has performance with the results of Acuracy, Recall, Precision, F Measure, and Classification Error with a value of 98%, 100%, 100%, 91.67%, and 2%. . This study can provide recommendations for specialization of study to prospective students of SMKN 1 Gadingrejo.

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

---

---

## 1. PENDAHULUAN

SMKN1 Gadingrejo adalah sekolah menengah kejuruan pertama di Pringsewu yang berdiri di Kabupaten Pringsewu. Sekolah Menengah kejuruan ini merupakan salah satu sekolah yang memiliki kualitas terbaik dengan akreditasi A. Sekolah menengah kejuruan ini memiliki tujuh program studi peminatan, yaitu prodi Teknik Kendaraan Ringan, Prodi Desain Permodelan dan Informasi Bangunan, Teknik Kontruksi Bangunan, Teknik Elektronika Industri, Teknik Audio Video, Teknik Multi Media dan Teknik Animasi.

Penentuan peminatan program studi di Sekolah Menengah Kejuruan disesuaikan dengan kemampuan dan ketertarikan siswa terhadap program studi yang ditawarkan oleh sekolah. Pengambilan keputusan penentuan peminatan program studi harus dipilih dengan baik dan tepat. Berdasarkan penelitian dari 49 responden yang digunakan menyatakan bahwa keterbatasan informasi mengenai bidang peminatan menjadi salah satu faktor sulitnya siswa untuk menentukan peminatan yang sesuai, sehingga hal ini berdampak terhadap nilai dan tingkat kelulusan siswa. Pengambilan peminatan yang salah dapat mengakibatkan berkurangnya mutu siswa dalam menyelesaikan studinya, dari 500 lulusan ada 16% lulusan yang kurang sesuai dalam pemilihan peminatan [2]. Pemilihan peminatan program studi oleh siswa saat ini hanya mengikuti teman serta mayoritas persepsi yang ada di masyarakat sekitar dan tidak berdasarkan kemampuan, sehingga sebuah sistem klasifikasi peminatan merupakan salah satu solusi untuk menyelesaikan permasalahan yang timbul pada saat pemilihan peminatan program studi di SMKN1 Gadingrejo.

Algoritma klasifikasi dianggap mampu mempertimbangkan peminatan siswa 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 siswa melalui hasil keluaran sistem berdasarkan data nilai siswa yang berupa skor nilai pada setiap peminatan.

Data yang digunakan pada penelitian ini menggunakan data numerik yang didapat dari siswa melalui nilai mata pelajaran prasyarat selama semester satu hingga semester lima. Data tersebut dibutuhkan untuk mempelajari pola-pola hubungan dari satu mata pelajaran satu dengan mata pelajaran lainnya sehingga membentuk satu peminatan yang utuh. Data tersebut yang kemudian diolah menggunakan algoritma klasifikasi guna mendapatkan hasil keluaran peminatan siswa .

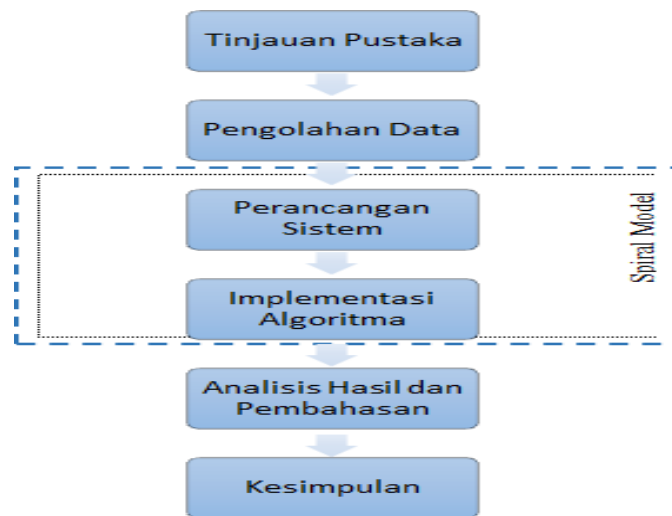
Dalam penelitian ini, akan dilakukan proses sampling dengan objek, iterasi, data uji dan data latih yang sama, sehingga nantinya menghasilkan tingkat akurasi dari algoritma. Penelitian ini menggunakan PHP dan MySQL sebagai tools dalam proses *scripting/programming* dalam pembuatan sistem yang mengimplementasikan 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 siswa di SMKN1 Gadingrejo serta diharapkan penelitian ini mampu memberikan rekomendasi kepada siswa 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:

### a. 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

### b. Pengolahan Data

Merupakan tahapan dalam mempersiapkan data yang diperoleh (data mentah) dirubah menjadi data siap diolah disimpan dalam format .xlsx (Ms. Excel). Data yang digunakan pada penelitian ini diperoleh dari Pangkalan Data Sekolah SMKN1 Gadingrejo berupa data nilai siswa angkatan 2014 dan 2015[4].

#### Data Set

Data yang digunakan nilai siswa pada semester 1 sampai dengan 5. Nilai siswa yang diambil terdiri dari 12 mata pelajaran yang terkait dengan program studi yang ada. 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 redudansi yang disebabkan oleh integrasi data. Data yang diperoleh sebelum dilakukan pembersihan total data siswa sebesar 435, setelah dilakukan pembersihan data diperoleh data siswa sebanyak 283.

Transformasi data: dari data nilai siswa 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

### c. 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

system 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.

#### d. Implementasi Algoritma

Pada tahap ini dilakukan proses implementasi algoritma K-NN dalam klasifikasi peminatan studi siswa 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 (query 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 *query instance* yang telah dihitung.

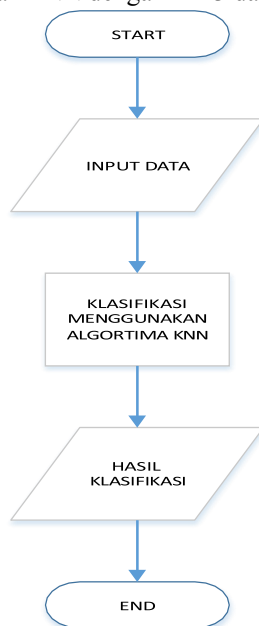
Hasil dari penerapan algoritma K-NN pada sistem kemudian akan dibandingkan dengan perhitungan manual algoritma K-NN menggunakan Ms. Excel.

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

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

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”.

- 
- [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 Siswa 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 Siswa 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 Siswa 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 Optimization Pada Metode K-Harmonic Means Untuk Klasterisasi Data,” *J. Ilmu Komput.*, vol. 5 No 1, 2012.
- [12]