



ALGORITMA K-NEAREST NEIGHBOR (K-NN) DENGAN NORMALISASI MAX MIN UNTUK MENENTUKAN CALON MAHASISWA YANG LAYAK MENERIMA KIP KULIAH MERDEKA

Efori Bu'ulolo

buuloloefori21@gmail.com

Teknik Informatika, Universitas Budi Darma

Abstrak

Calon mahasiswa penerima KIP Kuliah Merdeka pada Universitas Budi Darma tidak sebanding jumlahnya dengan kuota yang tersedia, sehingga kadang pengelola KIP Kuliah mengalami kesulitan dalam menentukan mahasiswa yang benar-benar layak menerima KIP Kuliah Merdeka atau yang tidak layak. Faktor lain adalah nilai ujian, prestasi, keadaan ekonomi dan lain sebagainya yang memiliki kemiripan. Untuk itu sangat diperlukan suatu teknik untuk menentukan calon mahasiswa yang layak menerima KIP Kuliah Merdeka, dan salah satu teknik dalam data mining yang sering digunakan adalah algoritma K-Nearest Neighbor (K-NN). Penggunaan normalisasi data untuk menghilangkan ketimpangan nilai antar kriteria dalam data. Perhitungan jarak dengan algoritma K-NN dengan normalisasi max min dan pengujian data dengan bahasa pemrograman python hasilnya adalah nilai *class* data yang baru berdasarkan nilai kedekatan nilai *attribute* data yang lama dengan data yang baru yang terdiri atas prestasi sekolah, hasil ujian, total penghasilan orang tua, mempunyai KKS dan kepemilikan rumah.

Kata kunci: KIP Kuliah Merdeka, K-NN, Max Min

Abstract

The number of prospective students receiving KIP Kuliah Merdeka at Budi Darma University is not proportional to the available quota, so sometimes the KIP Kuliah administrators have difficulty in determining which students are truly worthy of receiving KIP Kuliah Merdeka or those who are not. Other factors are test scores, achievements, economic conditions and so on which are similar. For this reason, a technique is needed to determine prospective students who are worthy of receiving KIP Kuliah Merdeka, and one of the techniques in data mining that is often used is the K-Nearest Neighbor (K-NN) algorithm. Use of data normalization to eliminate differences in values between criteria in the data. Distance calculation using the K-NN algorithm with max min normalization and data testing using the Python programming language. The result is a new data class value based on the closeness of the old data attribute values to the new data consisting of school achievement, exam results, total parent income, have KKS and home ownership.

Keywords: KIP Kuliah Merdeka, K-NN, Max Min

1. Pendahuluan

Kartu Indonesia Pintar (KIP) kuliah Merdeka merupakan program beasiswa untuk membantu calon mahasiswa yang berpotensi secara akademik, namun tidak memiliki kemampuan ekonomi. Program KIP kuliah digagas oleh presiden Joko Widodo yang merupakan kelanjutan dari program bidikmisi. Penyelenggara program KIP Kuliah Merdeka diselenggarakan oleh Perguruan Tinggi Negeri (PTN) dan Perguruan Tinggi Swasta (PTS) yang memenuhi persyaratan, salah satunya adalah program studi terakreditasi Baik Sekali atau terakreditasi Baik untuk wilayah tertentu[1][2].

Mahasiswa penerima KIP Kuliah Merdeka diberikan biaya hidup sesuai dengan standar biaya hidup tempat kuliah mahasiswa dan biaya Pendidikan sehingga mahasiswa penerima KIP Kuliah Merdeka tidak dibebankan dengan biaya apapun di Perguruan Tinggi selama menjalankan perkuliahan dengan masa studi 4 tahun untuk Sarjana dan 3 tahun untuk Diploma tiga. Syarat utama mahasiswa penerima KIP Kuliah Merdeka yaitu mahasiswa punya potensi akademik yang baik dan berasal dari keluarga miskin atau rentan miskin, sedangkan syarat lainnya berdasarkan panduan penerimaan KIP Kuliah dan aturan Perguruan Tinggi masing-masing penerima kuota.

Penerimaan Mahasiswa KIP Kuliah Merdeka atau namanya dulu Bidikmisi dimulai tahun 2018, setelah salah satu program studi yaitu Teknik Informatika STMIK Budi Darma memperoleh akreditasi B (Baik). Penerimaan mahasiswa baru tersebut dimulai pada tahun ajaran 2018/2019 dan sampai pada saat ini, kecuali tahun 2021 karena proses migrasi data perubahan bentuk dari STMIK menjadi Universitas Budi Darma.

Prosedur penerimaan calon mahasiswa baru di Universitas Budi Darma melalui jalur KIP Kuliah Merdeka dimulai dari calon mahasiswa mendaftar melalui sistem Informasi KIP Kuliah Merdeka <https://kip-kuliah.kemdikbud.go.id/>, dan memilih program studi dan Perguruan Tinggi yang diinginkan, selanjutnya pihak pengelola Universitas Budi Darma menghubungi calon mahasiswa melalui *email* dan nomor kontak yang tersedia untuk diinformasi jadwal ujian tertulis dan wawancara. Layak atau tidak calon mahasiswa penerima KIP Kuliah Merdeka berdasarkan prestasi pada saat sekolah, tingkat ekonomi keluarga, hasil ujian tertulis, dan hasil wawancara[1].

Calon mahasiswa penerima KIP Kuliah Merdeka tidak sebanding dengan kuota yang tersedia, sehingga pengelola KIP Kuliah Merdeka di Universitas Budi Darma kadang mengalami kesulitan dalam menentukan mahasiswa yang benar-benar layak menerima KIP Kuliah Merdeka. Selain karena jumlah dan kuota faktor lain adalah calon penerima KIP Kuliah Merdeka kadang memiliki kriteria yang hampir bersamaan seperti nilai ujian, prasetasi, keadaan ekonomi dan lain sebagainya[3].

Untuk itu sangat diperlukan suatu teknik untuk menentukan calon mahasiswa yang layak menerima KIP Kuliah Merdeka, dan salah satu teknik dalam data mining yang sering digunakan adalah algoritma K-Nearest Neighbor (K-NN) [4]. Algoritma K-NN adalah teknik klasifikasi objek berdasarkan mayoritas tetangga terdekat, atau definisi lainnya adalah membandingkan kriteria data lama dengan data yang baru, hasil klasifikasi data yang baru berdasarkan nilai class mayoritas data yang lama[5][6]. Dalam proses perhitungannya menggunakan normalisasi max min. Normalisasi max min adalah salah bentuk normalisasi dengan untuk mempertahankan bentuk dan nilai pasti dari data minimum dan maksimum[7][8].

Penggunaan normalisasi data untuk menghilangkan ketimbangan nilai antar kriteria dalam data, contohnya kriteria gaji dengan kriteria umur, dimana penulisan nilai gaji menggunakan banyak digit angka sedangkan umur maksimal tiga digit angka sehingga jika dilakukan perhitungan tanpa normalisasi maka angka nilai gaji akan mereduksi nilai umur dan tidak berdampak sama sekali[9]. Maka, berdasarkan uraian diatas, maka penulis mengangkat judul penelitian yaitu algoritma K-Nearest Neighbor (K-NN) dengan normalisasi max min untuk menentukan calon mahasiswa yang layak menerima KIP Kuliah Merdeka.

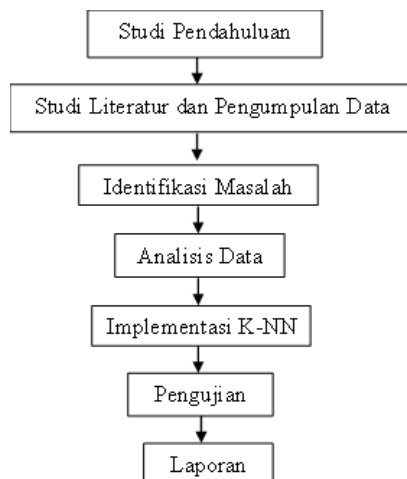
Penelitian yang dilakukan oleh M. Syukri Mustafa dan I Wayan Simpen, tentang penerapan algoritma K-NN untuk prediksi pasien terkena penyakit diabetes. Hasil penelitiannya menggunakan 104 data, dengan prediksi yang benar sebanyak 71 dan salah atau ragu-ragu sebesar 33 dengan tingkat akurasi sebesar 68.3%[10]. Penelitian yang dilakukan oleh Muhammad Sholeh dkk, tentang klasifikasi dengan K-NN untuk prediksi penyakit diabetes, hasil penelitiannya adalah dengan dataset penyakit diabetes, proses pengujian K=1 sampai dengan K=30, Hasil pengujian menghasilkan nilai akurasi yang terbaik adalah dengan melakukan normalisasi Min-Max dan diperoleh nilai k=3 dengan nilai akurasi 74%[11].

Penelitian yang dilakukan oleh Lia Farokhah, tentang Klasifikasi bunga dengan ekstraksi fitur warna RGB dengan K-NN, hasil penelitiannya adalah metode ekstraksi fitur warna RGB kurang cocok untuk mengklasifikasi label objek dengan kemiripan warna walaupun bentuknya berbeda namun masih bisa dipakai dalam mengklasifikasikan objek dengan perbedaan warna yang mencolok di salah satu sisi walaupun bentuk objeknya memiliki struktur bentuk yang mirip[12]. Penelitian yang dilakukan oleh Nora Lizarti dan Aniq Noviciatie Ulfah, tentang penerapan K-NN untuk penentuan peminatan studi, hasilnya adalah 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[13].

Untuk keterbaruan pada penelitian ini dibandingkan dengan penelitian sebelumnya adalah penerapan algoritma K-NN pada klasifikasi calon mahasiswa yang layak menerima KIP Kuliah Merdeka, dimana data yang digunakan harus dinormalisasikan terlebih dahulu dengan tujuan untuk menghilangkan ketimbangan antar kriteria dalam data. Selain itu, pengujian menggunakan bahasa pemrograman Python dengan *library sklearn.neighbors import KNeighborsClassifier*. Library tersebut sebagai *library* algoritma dari K-NN.

2. Metode

Berikut adalah metode penelitian yang dilakukan untuk menentukan calon mahasiswa yang layak menerima KIP Kuliah Merdeka, sebagai berikut:



Gambar 1. Tahapan / Kerangka Kerja Penelitian

1. Studi Pendahuluan

Langkah yang paling awal dalam penelitian ini adalah studi pendahuluan, yaitu mempelajari masalah yang diteliti. Agar penelitian ini fokus pada bidang yang diteliti maka ruang lingkup masalah yang akan diteliti harus ditentukan terlebih dahulu, karena tanpa menentukan serta mendefinisikan ruang lingkup masalah yang akan diteliti, maka tidak pernah ada solusi dari masalah tersebut.

2. Studi Literatur

Untuk mencapai tujuan yang akan diinginkan, maka perlu dipelajari beberapa literatur yang akan digunakan. Kemudian literatur yang dipelajari tersebut dipilih untuk dapat ditentukan sebagai literatur mana yang akan digunakan dalam sebagai referensi untuk penelitian. Dari studi literatur yang telah tersedia, maka dipelajari teori yang berhubungan dengan algoritma K-NN, normalisasi max min dan data mining yang akan dibahas. Teori-teori tersebut diatas berdasarkan sumber berupa buku, paper, jurnal / *prosiding* maupun situs internet yang berhubungan dengan data mining.

3. Pengumpulan Data

1) Wawancara

Wawancara adalah percakapan antara minimal 2 (dua) orang atau lebih dengan tujuan tertentu. Wawancara dilakukan oleh penulis untuk bertanya secara langsung tentang apa saja masalah yang sedang / sering dihadapi oleh Universitas Budi Darma khusus tentang dalam menentukan calon mahasiswa penerima beasiswa bidikmisi dan data tersebut dapat digunakan sebagai salah satu bahan pertimbangan dalam pengambilan keputusan.

2) Observasi

Pengamatan dapat diartikan sebagai teknik/metode ilmiah observasi. Penyelidikan yang dilakukan secara sistematis merupakan observasi. Untuk pengelompokkan data mahasiswa berpotensi maka penulis melakukan pengamatan terhadap data-data calon mahasiswa baru selama 3 (tiga) tahun terakhir.

3) Dokumentasi

Dokumentasi merupakan kegiatan yang dilakukan untuk mencari data mengenai hal-hal atau variabel berupa catatan notulen, buku, transkrip, surat, majalah, prasasti, rapat, lengger, agenda dan sebagainya baik yang berupa file atau dokumen disebut dalam dokumentasi. Variabel yang digunakan untuk pengelompokkan calon mahasiswa berpotensi adalah nilai ujian seleksi digabung dengan nilai ujian nasional.

4. Identifikasi Masalah

Agar penulis mengetahui masalah langsung yang dihadapi, maka penulis melakukan identifikasi masalah, dalam tahap ini penulis melakukan identifikasi masalah yang berhubungan dengan kesulitan yang dihadapi panitia dalam menentukan calon mahasiswa penerima KIP Kuliah, kemudian menggambarkan/mendeskripsikan masalah tersebut untuk diperoleh penyelesaiannya.

5. Analisis Data

Tahap selanjutnya yang dilakukan oleh penulis adalah analisis data, dengan tujuan agar penulis mengetahui dan memperoleh gambaran yang jelas bagaimana bentuk penyelesaian data dan algoritma apa yang dapat digunakan untuk penyelesaiannya. Dalam penelitian ini penulis menggunakan algoritma data mining K-NN.

6. Implementasi K-NN

Implementasi merupakan suatu cara untuk menerapkan algoritma yang digunakan terhadap data yang telah tersedia. Pada penelitian ini penulis menggunakan algoritma K-NN dan data calon mahasiswa penerima KIP Kuliah atau beasiswa bidikmisi sebelumnya.

7. Pengujian

Selanjutnya melakukan pengujian data yang digunakan dengan menggunakan pemrograman Python untuk menguji akurasi algoritma K-NN terhadap data yang digunakan[14].

8. Laporan

Hasil dari proses analisis data dan implementasi dituangkan kedalam laporan hasil penelitian dan publikasi dalam bentuk prosiding atau jurnal.

3. Hasil dan Pembahasan

3.1 Hasil

Dalam penentuan calon mahasiswa yang layak memperoleh KIP Kuliah Merdeka diperlukan data calon mahasiswa yang mengikuti seleksi calon mahasiswa KIP Kuliah sebelumnya. Data yang digunakan harus mempunyai *attribute* kriteria dan *attribute* keputusan. Adapun datanya ditunjukkan pada Tabel 1 berikut ini:

Tabel 1. Data Seleksi KIP Kuliah Merdeka Sebelumnya

No	Nama	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Keterangan
1	A. Pretiwi	Ada	78	2,8 Jt	Tidak	Keluarga	Tidak
2	Enggi J. Sinaga	Tidak	40	2,1 Jt	Ada	Keluarga	Tidak
3	Parasia Leo	Tidak	75	1,1 Jt	Ada	Keluarga	Diterima
4	Lenny A. Daulay	Tidak	95	1,5 Jt	Ada	Keluarga	Diterima
5	Howu Laia	Ada	80	2,2 Jt	Ada	Keluarga	Diterima
6	Amanda Z	Tidak	58	2,5 Jt	Tidak	Menumpang	Tidak
7	Fransiska Dita	Tidak	70	2,6 Jt	Tidak	Keluarga	Tidak
8	Putri Erly	Tidak	92	1,8 Jt	Ada	Keluarga	Diterima
9	Amy Ahmad	Tidak	47	1,7 Jt	Ada	Keluarga	Tidak
10	Hidayat Ritonga	Tidak	59	2,9 Jt	Tidak	Menumpang	Tidak
11	Afdillah Hafiz	Tidak	88	3,8 Jt	Tidak	Keluarga	Tidak
12	M. Siarani	Tidak	87	3,9 Jt	Tidak	Keluarga	Tidak
13	Kurmi Lasmi	Ada	55	3,5 JT	Ada	Sewa	Tidak
14	Dodo Angraini	Tidak	75	2,5 Jt	Tidak	Keluarga	Tidak
15	Suci Loila	Ada	77	2,2 Jt	Ada	Sewa Tahunan	Diterima
16	Dodi Harun	Ada	50	2,3 Jt	Tidak	Keluarga	Tidak
17	Duli Sarman	Ada	91	2,0 Jt	Ada	Keluarga	Diterima
18	Rusiaman	Ada	80	2,5 Jt	Ada	Keluarga	Diterima
19	Oky Setiawan	Tidak	45	1,2 Jt	Tidak	Keluarga	Tidak
20	Boloidan	Tidak	60	2,9 Jt	Tidak	Sewa	Tidak
21	Susi Amini	Tidak	65	3,1 Jt	Tidak	Keluarga	Tidak
22	Tuti Armin Zega	Tidak	80	3,4 Jt	Tidak	Keluarga	Tidak

Data yang dibutuhkan pada algoritma K-NN adalah data dalam bentuk numerik, oleh karena itu data dalam bentuk kategori harus dirubah dalam bentuk numerik yaitu kriteria prestasi sekolah, mempunyai KIS, dan status kepemilikan rumah harus dirubah dalam bentuk numerik.

Tabel 2. Kategori Prestasi Sekolah

Kategori	Nilai
Ada	1
Tidak	2

Tabel 3. Mempunyai KKS

Kategori	Nilai
Ada	1
Tidak	2

Tabel 4. Status Kepemilikan Rumah

Kategori	Nilai
Keluarga	1
Sewa	2
Menumpang	3

Setelah beberapa kriteria nilai kategorinya dirubah dalam bentuk numerik, maka data yang digunakan berubah seperti pada Tabel 5 berikut ini:

Tabel 5. Data Seleksi KIP Kuliah Merdeka Setelah Beberapa Nilai Kategori Berubah dalam Bentuk Numerik

No	Nama	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Ket.
1	A. Pretiwi	1	78	2,8 Jt	2	1	Tidak
2	Enggi J. Sinaga	2	40	2,1 Jt	1	1	Tidak
3	Parasia Leo	2	75	1,1 JT	1	1	Diterima
4	Lenny A. Daulay	2	95	1,5 Jt	1	1	Diterima
5	Howu Laia	1	80	2,2 Jt	1	1	Diterima
6	Amanda Z	2	58	2,5 Jt	2	3	Tidak
7	Fransiska Dita	2	70	2,6 Jt	2	1	Tidak
8	Putri Erly	2	92	1,8 Jt	1	1	Diterima
9	Amy Ahmad	2	47	1,7 Jt	1	1	Tidak
10	Hidayat Ritonga	2	59	2,9 Jt	2	3	Tidak
11	Afdillah Hafiz	2	88	3,8 Jt	2	1	Tidak
12	M. Siarani	2	87	3,9 Jt	2	1	Tidak
13	Kurmi Lasmi	1	55	3,5 JT	1	2	Tidak
14	Dodo Angraini	2	75	2,5 Jt	2	1	Tidak
15	Suci Loila	1	77	2,2 Jt	1	2	Diterima
16	Dodi Harun	1	50	2,3 Jt	2	1	Tidak
17	Duli Sarman	1	91	2,0 Jt	1	1	Diterima
18	Rusiaman	1	80	2,5 Jt	1	1	Diterima
19	Oky Setiawan	2	45	1,2 Jt	2	1	Tidak
20	Boloidan	2	60	2,9 Jt	2	2	Tidak
21	Susi Amini	2	65	3,1 Jt	2	1	Tidak
22	Tuti Armin Zega	2	80	3,4 Jt	2	1	Tidak

Langkah selanjutnya adalah melakukan normalisasi data tersebut diatas dengan model max min dengan tujuan untuk menghindari ketimbangan pengaruh antar kriteria attribute pada data.

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}}$$

Setelah selesai datanya dinormalisasikan maka terjadi perubahan pada nilai attribut seperti pada Tabel 6 berikut ini:

Tabel 6. Data Seleksi KIP Kuliah Merdeka Setelah Normalisasi

No	Nama	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Keterangan
1	A. Pretiwi	0	0,69	0,61	1	0	Tidak
2	Enggi J. Sinaga	1	0,00	0,36	0	0	Tidak
3	Parasia Leo	1	0,64	0,00	0	0	Diterima
4	Lenny A. Daulay	1	1,00	0,14	0	0	Diterima
5	Howu Laia	0	0,73	0,39	0	0	Diterima
6	Amanda Z	1	0,33	0,50	1	1	Tidak
7	Fransiska Dita	1	0,55	0,54	1	0	Tidak
8	Putri Erly	1	0,95	0,25	0	0	Diterima
9	Amy Ahmad	1	0,13	0,21	0	0	Tidak
10	Hidayat Ritonga	1	0,35	0,64	1	1	Tidak
11	Afdillah Hafiz	1	0,87	0,96	1	0	Tidak
12	M. Siarani	1	0,85	1,00	1	0	Tidak
13	Kurmi Lasmi	0	0,27	0,86	0	0,5	Tidak
14	Dodo Angraini	1	0,64	0,50	1	0	Tidak
15	Suci Loila	0	0,67	0,39	0	0,5	Diterima
16	Dodi Harun	0	0,18	0,43	1	0	Tidak
17	Duli Sarman	0	0,93	0,32	0	0	Diterima
18	Rusiaman	0	0,73	0,50	0	0	Diterima
19	Oky Setiawan	1	0,09	0,04	1	0	Tidak
20	Boloidan	1	0,36	0,64	1	0,5	Tidak
21	Susi Amini	1	0,45	0,71	1	0	Tidak
22	Tuti Armin Zega	1	0,73	0,82	1	0	Tidak

Pada tahun 2023, dua orang calon penerima KIP Kuliah dengan data sebagai berikut:

Tabel 7. Data Mahasiswa Calon Perima KIP Kuliah Merdeka 2023

No	Nama Calon	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Keterangan
1	Yuni L. H	1	0,87	0,96	1	0,5	?
2	Mhd. Jowi	0	0,45	0,50	0	1	?

Langkah-langkah untuk memperoleh nilai keterangan (Layak/Tidak) pada Tabel 7 adalah sebagai berikut:

- Jumlah tetangga terdekat adalah (K=3)
- Proses perhitungan jarak menggunakan persamaan Euclidean Distance dengan

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

Untuk calon Yuni L. H, hasil perhitungan jaraknya sebagai berikut:

Tabel 8. Rekapitulasi Hasil Perhitungan Jarak Yuni L. H

No	Nama	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Keterangan	Jarak Calon 1
11	Afdillah Hafiz	1	0,87	0,96	1	0	Tidak	0,50
12	M. Siarani	1	0,85	1	1	0	Tidak	0,50
22	Tuti Armin Zega	1	0,73	0,82	1	0	Tidak	0,54
20	Boloidan	1	0,36	0,64	1	0,5	Tidak	0,60
21	Susi Amini	1	0,45	0,71	1	0	Tidak	0,70
14	Dodo Angraini	1	0,64	0,5	1	0	Tidak	0,72
7	Fransiska Dita	1	0,55	0,54	1	0	Tidak	0,73
10	Hidayat Ritonga	1	0,35	0,64	1	1	Tidak	0,79
6	Amanda Z	1	0,33	0,5	1	1	Tidak	0,87
1	A. Pretiwi	0	0,69	0,61	1	0	Tidak	1,19
19	Oky Setiaman	1	0,09	0,04	1	0	Tidak	1,31
8	Putri Erly	1	0,95	0,25	0	0	Diterima	1,33
4	Lenny A. Daulay	1	1	0,14	0	0	Diterima	1,39
16	Dodi Harun	0	0,18	0,43	1	0	Tidak	1,42
3	Parasia Leo	1	0,64	0	0	0	Diterima	1,49
9	Amy Ahmad	1	0,13	0,21	0	0	Tidak	1,54
15	Suci Loila	0	0,67	0,39	0	0,5	Diterima	1,54
2	Enggi J. Sinaga	1	0	0,36	0	0	Tidak	1,54
13	Kurmi Lasmi	0	0,27	0,86	0	0,5	Tidak	1,54
18	Rusiaman	0	0,73	0,5	0	0	Diterima	1,58
5	Dodo Angraini	0	0,73	0,39	0	0	Diterima	1,61
17	Duli Sarman	0	0,93	0,32	0	0	Diterima	1,63

Dari perhitungan diatas, ternyata 3 (tiga) tetangga terdekat dari calon Yuni L. H, Keterangannya semua menyatakan “Tidak”, maka dapat disimpulkan calon 1 mahasiswa KIP Kuliah Merdeka atas nama “Yuni L. H”, “Tidak” direkomendasikan.

Untuk calon Mhd. Jowi, hasil perhitungan jaraknya sebagai berikut:

Tabel 9 Rekapitasi Hasil Perhitungan Jarak Calon Mhd. Jowi

No	Nama	Prestasi Sekolah	Hasil Ujian	Total Penghasilan Orang Tua	Mempunyai KKS	Status Kepemilikan Rumah	Keterangan	Jarak Calon 2
15	Suci Loila	0	0,67	0,39	0	0,5	Diterima	0,56
13	Kurmi Lasmi	0	0,27	0,86	0	0,5	Tidak	0,64
18	Rusiaman	0	0,73	0,5	0	0	Diterima	1,04
5	Howu Laia	0	0,73	0,39	0	0	Diterima	1,04
17	Duli Sarman	0	0,93	0,32	0	0	Diterima	1,12
6	Amanda Z	1	0,33	0,5	1	1	Tidak	1,42
10	Hidayat Ritonga	1	0,35	0,64	1	1	Tidak	1,42
1	A. Pretiwi	0	0,69	0,61	1	0	Tidak	1,44
16	Dodi Harun	0	0,18	0,43	1	0	Tidak	1,44
9	Amy Ahmad	1	0,13	0,21	0	0	Tidak	1,48
2	Enggi J. Sinaga	1	0	0,36	0	0	Tidak	1,49
20	Boloidan	1	0,36	0,64	1	0,5	Tidak	1,51
3	Parasia Leo	1	0,64	0	0	0	Diterima	1,51
8	Putri Erly	1	0,95	0,25	0	0	Diterima	1,52
4	Lenny A. Daulay	1	1	0,14	0	0	Diterima	1,56
7	Fransiska Dita	1	0,55	0,54	1	0	Tidak	1,74
14	Dodo Angraini	1	0,64	0,5	1	0	Tidak	1,74
21	Susi Amini	1	0,45	0,71	1	0	Tidak	1,74
22	Tuti Armin Zega	1	0,73	0,82	1	0	Tidak	1,78
19	Oky Setiaman	1	0,09	0,04	1	0	Tidak	1,83
11	Afdillah Hafiz	1	0,87	0,96	1	0	Tidak	1,84
12	M. Siarani	1	0,85	1	1	0	Tidak	1,85

Dari perhitungan diatas, ternyata 3 (tiga) tetangga terdekat dari calon Mhd. Jowi. Keteranganannya menyatakan 2 (dua) “Diterima” dan satu “Tidak”, maka dapat disimpulkan calon 2 mahasiswa KIP Kuliah Merdeka atas nama “Mhd. Jowi”, direkomendasikan untuk di “Terima”.

3.2 Pengujian

Selanjutnya dilakukan pengujian data dengan menggunakan bahasa Python dengan aplikasi Anaconda dan editor Jupyter Notebook.

1. Siapkan data dalam bentuk csv, sehingga data berubah seperti berikut dengan nama file kipkuliah.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	No,"Nama","Prestasi Sekolah","Hasil Ujian","Total Penghasilan Orang Tua","Mempunyai KKS","Status Kepemilikan Rumah","Keterangan"													
2	1,"A. Pretiwi","0","0.69","0.61","1","0","0"													
3	2,"Enggi J. Sinaga","1","0","0.36","0","0","0"													
4	3,"Parasia Leo","1","0.64","0","0","0","1"													
5	4,"Lenny A. Daulay","1","1","0.14","0","0","1"													
6	5,"Howu Laia","0","0.73","0.39","0","0","1"													
7	6,"Amanda Z","1","0.33","0.5","1","1","0"													
8	7,"Fransiska Dita","1","0.55","0.54","1","0","0"													
9	8,"Putri Erly","1","0.95","0.25","0","0","1"													
10	9,"Amy Ahmad","1","0.13","0.21","0","0","0"													
11	10,"Hidayat Ritonga","1","0.35","0.64","1","1","0"													
12	11,"Afdillah Hafiz","1","0.87","0.96","1","0","0"													
13	12,"M. Siarani","1","0.85","1","1","0","0"													
14	13,"Kurmi Lasmi","0","0.27","0.86","0","0.5","0"													
15	14,"Dodo Angraini","1","0.64","0.5","1","0","0"													
16	15,"Suci Loila","0","0.67","0.39","0","0.5","1"													
17	16,"Dodi Harun","0","0.18","0.43","1","0","0"													
18	17,"Duli Sarman","0","0.93","0.32","0","0","1"													
19	18,"Rusiaman","0","0.73","0.5","0","0","1"													
20	19,"Okky Setiawan","1","0.09","0.04","1","0","0"													
21	20,"Boloidan","1","0.36","0.64","1","0.5","0"													
22	21,"Susi Amini","1","0.45","0.71","1","0","0"													
23	22,"Tuti Armin Zega","1","0.73","0.82","1","0","0"													

Gambar 2. Data Dalam Bentuk CSV

2. Ketik kode program Python pada editor Jupyter Notebook

```
#library
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import datasets
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report

#membaca file
kipkuliah=pd.read_csv('kipkuliah.csv')
df=pd.DataFrame(kipkuliah,columns=["No","Nama","Prestasi Sekolah","Hasil Ujian","Total Penghasilan Orang Tua","Mempunyai KKS","Status Kepemilikan Rumah","Keterangan"])
X=np.asarray(kipkuliah)
x_train=X[:,2:7]
y_train=X[:,7:8]
```



```
#merubah tipe data semua ke float
x_train=x_train.astype('float')
y_train=y_train.astype('float')
#menampilkan algoritma knn
knn=KNeighborsClassifier(n_neighbors=3,
weights='uniform',
algorithm='auto',
metric='euclidean')
knn.fit(x_train,y_train)

#memprediksi hasil knn
y_pred=knn.predict(x_train)
print((y_train).T)
print(y_pred)
print(classification_report(y_train,y_pred))

#uji 1 dengan data(1,0.87,0.96,1,0.50)
uji=np.asarray([[1.00,0.87,0.96,1.00,0.50]])
y_pred=knn.predict(uji)
print(y_pred)

#uji 2 dengan data(0,0.45,0.50,0,1)
uji=np.asarray([[0.00,0.45,0.50,0.00,1.00]])
y_pred=knn.predict(uji)
print(y_pred)
```

#Hasil Pengujian

```
In [7]: #uji 1 dengan data(1,0.87,0.96,1,0.50)
uji=np.asarray([[1.00,0.87,0.96,1.00,0.50]])
y_pred=knn.predict(uji)
print(y_pred)

[0.]

In [8]: #uji 2 dengan data(0,0.45,0.50,0,1)
uji=np.asarray([[0.00,0.45,0.50,0.00,1.00]])
y_pred=knn.predict(uji)
print(y_pred)

[1.]
```

Gambar 3. Hasil Pengujian

Setelah data kipkuliah.csv terbaca pada python, maka langkah selanjutnya melakukan pengujian data baru terhadap data lama, dimana pengujian pertama atas nama Yuni L. H dengan kriteria yang ada, hasilnya adalah [0] artinya adalah tidak direkomendasikan. Untuk pengujian kedua atas nama Mhd. Jowi dengan hasil [1] artinya adalah diterima / direkomendasikan untuk menerima KIP Kuliah Merdeka. Hasil perhitungan algoritma K-NN secara manual dan python hasilnya sama dan tidak terdapat perbedaan hasil sehingga akurasi 100%.

4. Kesimpulan dan Saran

4.1 Kesimpulan

Kesimpulan yang diperoleh dari hasil penelitian ini adalah:

1. Algoritma K-Nearest Neighbor(K-NN) dapat mengklasifikasikan calon mahasiswa yang layak menerima KIP Kuliah Merdeka atau yang tidak.
2. Dengan normalisasi max min perhitungan pencarian jarak dengan K-NN lebih sederhana, karena nilai nominal attribute mempunyai digit yang sama dan lebih sedikit.

3. Perhitungan secara manual algoritma K-NN mempunyai hasil yang sama dengan pengujian dengan menggunakan bahasa Python pada data yang sama.
4. Dengan implementasi algoritma K-NN dengan normalisasi max min pada klasifikasi mahasiswa yang layak memperoleh KIP Kuliah Merdeka, maka mempermudah pihak pengelola dalam proses seleksi.

4.2 Saran

Adapun saran yang disampaikan pada penelitian ini adalah sebagai berikut:

1. Pada data yang sama agar digunakan algoritma yang lain sebagai bahan perbandingan.
2. Agar hasil lebih optimal maka data yang digunakan sebagai sampel diperbanyak.
3. Agar dilakukan pengujian hasil untuk mengetahui akurasi hasil yang diperoleh dengan aplikasi data mining.

Daftar Pustaka

- [1] Puslapdik, Pedoman Pendaftaran Kartu Indonesia Pintar Kuliah - Kip Kuliah Merdeka 2022. Jakarta: Sekretaris Jenderal Kemendikbudristek, 2022. [Online]. Available: https://kip-kuliah.kemdikbud.go.id/uploads/Pedoman-Pendaftaran-KIP-K-2022-ver-20220202---final_cd9b5e.pdf
- [2] A. E. Wicaksono, "Implementasi Data Mining Dalam Pengelompokan Data Peserta Didik Di Sekolah Untuk Memprediksi Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means (Studi Kasus SMAN 16 Bekasi)," *J. Ilm. Teknol. dan Rekayasa*, vol. 21, no. 3, 2017, [Online]. Available: <http://www.ejournal.gunadarma.ac.id/index.php/teknol/article/view/1599>
- [3] F. tinus Bu'ulolo;Efori;Fau, Alwin;Ndruru, Eferoni;Waruwu, "Classification of prospective students who are eligible to receive KIP Merdeka lecture using the cart algoritma approach," in *Melaka International Social Sciences, Science, and Technology (MIC3ST)*, Melaka, Malaysia: <https://www.mic3st.com/>, 2023. [Online]. Available: <https://ir.uitm.edu.my/id/eprint/81897/>
- [4] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Yogyakarta: deepublish, 2020.
- [5] I. G. Harsemadi, "Rekayasa Sistem Pengelompokan Suasana Hati Terhadap Musik Menggunakan Algoritma K-Nearest Neighbor," *Eksplora Inform.*, vol. 5, no. 2, pp. 166–176, 2016, [Online]. Available: <https://eksplora.stikom-bali.ac.id/index.php/eksplora/article/view/88/71>
- [6] I. Listiowarni and N. Puspa Dewi, "Pemanfaatan Klasifikasi Soal Biologi Cognitive Domain Bloom's Taxonomy Menggunakan KNN Chi-Square Sebagai Penyusunan Naskah Soal," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 11, no. 2, pp. 186–197, 2020, doi: 10.31849/digitalzone.v11i2.4798.
- [7] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, p. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [8] D. Fenny, "Analisis Perbandingan Cosine Normalization dan Min Max Pada Pengelompokan Terjemahan Ayat Al Quran Menggunakan Algoritma K-Means Clustering," Universitas Islam Negeri Syarif Hidayatullah Jakarta, 2019. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/bitstream/123456789/48588/1/DEWINTA FENNY-FST.pdf>
- [9] V. Rapika Sari, E. Buulolo, and K. Kunci ABSTRAK, "Implementasi Algoritma K-Means dengan Normalisasi Sigmoidal Untuk Klastering Data Ternak Sapi," *Jikteks*, vol. 02, no. 01, pp. 30–42, 2023.
- [10] M. S. Mustafa and I. W. Simpen, "Implementasi Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Pasien Terkena Penyakit Diabetes Pada Puskesmas Manyampa Kabupaten Bulukumba," *Pros. Semin. Ilm. Sist. Inf. Dan Teknol. Inf.*, vol. VIII, no. 1, pp. 1–10, 2019, [Online]. Available: <https://ejurnal.dipangegara.ac.id/index.php/sisiti/article/view/1-10>
- [11] M. Sholeh, D. Andayati, and R. Y. Rachmawati, "Data Mining Model Klasifikasi Menggunakan Algoritma K-Nearest Neighbor Dengan Normalisasi Untuk Prediksi Penyakit Diabetes," *TelKa*, vol. 12, no. 02, pp. 77–87, 2022, doi: 10.36342/teika.v12i02.2911.
- [12] L. Farokhah, "Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1129–1136, 2020, doi: 10.25126/jtiik.2020722608.

-
- [13] N. Lizarti and A. N. Ulfah, “Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Peminatan Studi (Studi Kasus : Program Studi Teknik Informatika STMIK Amik Riau),” *Fountain Informatics J.*, vol. 4, no. 1, p. 1, 2019, doi: 10.21111/fij.v4i1.2822.
- [14] R. Bu’ulolo, Efori;Mesran;Hasibuan, Nelly Astuti; Aripin, Sueb;Utomo, Dito Utomo;Syaputra, *Big Data Analysis dengan Phyton untuk Perguruan Tinggi*, I. Yogyakarta, 2023.