

KLASIFIKASI KELAYAKAN PENERIMA KERINGANAN UANG KULIAH TUNGGAL MAHASISWA POLITEKNIK NEGERI BENGKALIS MENGGUNAKAN ALGORITMA DECISION TREE C5.0

Maisharoh, Elvi Rahmi

Teknik Informatika, Politeknik Negeri Bengkalis
Jalan Bathin Alam Sungai Alam, Indonesia
wakkmel108@gmail.com

ABSTRAK

Sistem penentuan keringanan Uang Kuliah Tunggal (UKT) di Politeknik Negeri Bengkalis saat ini masih dilakukan secara manual. Kondisi tersebut menyebabkan proses seleksi membutuhkan waktu yang lama dan berpotensi menimbulkan kesalahan subjektif dalam pengambilan keputusan. Permasalahan utama dalam penelitian ini adalah belum tersedianya sistem klasifikasi yang objektif dan terkomputerisasi untuk menentukan kelayakan penerima keringanan UKT. Penelitian ini bertujuan untuk mengembangkan model klasifikasi kelayakan penerima keringanan UKT menggunakan algoritma Decision Tree C5.0 serta mengimplementasikannya dalam aplikasi berbasis website. Dataset yang digunakan terdiri dari atribut tempat tinggal, status anak bagi orang tua, pekerjaan orang tua, penghasilan orang tua, jumlah tanggungan, kepemilikan DTKS/KIP/KKS/PKH, dan kepemilikan SKTM. Metode penelitian meliputi tahapan preprocessing data, pembagian dataset, pembangunan model Decision Tree C5.0, serta evaluasi kinerja model menggunakan confusion matrix. Hasil pengujian menunjukkan bahwa model menghasilkan tingkat akurasi sebesar 88,42%, yang menandakan kemampuan klasifikasi yang tinggi dan konsisten. Implementasi sistem berbasis web mampu mempercepat proses seleksi, mengurangi subjektivitas, serta meningkatkan efisiensi dalam penentuan kelayakan penerima keringanan UKT.

Kata kunci: Keringanan UKT, Klasifikasi, Decision Tree C5.0, Data Mining, Website

1. PENDAHULUAN

Perkembangan teknologi informasi pada era Revolusi Industri 4.0 telah mendorong digitalisasi dan otomatisasi di berbagai bidang, termasuk sektor pendidikan tinggi. Digitalisasi layanan akademik menghasilkan data dalam jumlah besar yang dapat dimanfaatkan untuk mendukung pengambilan keputusan. Salah satu kebijakan penting di perguruan tinggi negeri adalah penerapan Uang Kuliah Tunggal (UKT), yaitu sistem pembiayaan perkuliahan yang disesuaikan dengan kemampuan ekonomi mahasiswa. Pendidikan tinggi memiliki peran strategis dalam pembangunan suatu negara, sehingga kebijakan UKT perlu diterapkan secara adil dan tepat sasaran.

Politeknik Negeri Bengkalis sebagai perguruan tinggi negeri menyediakan fasilitas pengajuan keringanan UKT bagi mahasiswa yang mengalami keterbatasan ekonomi. Namun, proses pendataan dan seleksi mahasiswa penerima keringanan UKT saat ini masih dilakukan secara manual, sehingga kurang efisien dan berpotensi menimbulkan kesalahan subjektif.

Permasalahan utama dalam proses penentuan keringanan UKT adalah lamanya waktu seleksi akibat banyaknya data mahasiswa yang harus dianalisis, serta tingginya risiko terjadinya *human error* dalam pengambilan keputusan. Selain itu, belum tersedia sistem terkomputerisasi yang mampu mengklasifikasikan kelayakan mahasiswa secara objektif berdasarkan data ekonomi dan sosial yang dimiliki.

Penelitian ini bertujuan untuk membangun model klasifikasi kelayakan penerima keringanan UKT menggunakan algoritma Decision Tree C5.0 serta mengimplementasikannya dalam sebuah aplikasi berbasis website guna membantu panitia penyeleksi dalam pengambilan keputusan yang lebih cepat, objektif, dan akurat.

Rencana penyelesaian masalah dalam penelitian ini dilakukan dengan menerapkan teknik data mining, khususnya metode klasifikasi menggunakan algoritma Decision Tree C5.0. Dataset yang digunakan meliputi atribut tempat tinggal, status anak, pekerjaan orang tua, penghasilan orang tua, jumlah tanggungan, kepemilikan DTKS/KIP/KKS/PKH, serta kepemilikan SKTM. Sistem dikembangkan menggunakan bahasa pemrograman PHP dengan framework Laravel dan diharapkan mampu meningkatkan efisiensi proses seleksi serta mengurangi potensi kesalahan manusia dalam penentuan keringanan UKT.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining adalah metode untuk menemukan pola tertentu dari kumpulan data yang berjumlah besar. Data mining tidak hanya dipelajari pada bidang ilmu komputer dan statistik saja, tetapi juga bisa diterapkan dan memudahkan pekerjaan di bidang lainnya juga. Data mining juga adalah metode dalam ilmu komputer yang biasa digunakan dalam proses pencarian knowledge. Tahapan di dalamnya berguna untuk mencari pola-pola tertentu dari data yang ada pada

database. Biasanya, metode ini banyak ditemukan pada bidang machine learning dan statistika[12].

2.2. Klasifikasi

Klasifikasi adalah salah satu teknik pada data mining yang memetakan data ke dalam kelompok atau kelas yang telah ditentukan. Klasifikasi merupakan metode supervised learning yang membutuhkan data training berlabel untuk menghasilkan sebuah aturan yang mengklasifikasikan data uji ke dalam kelompok atau kelas yang telah ditentukan. Beberapa teknik klasifikasi yang digunakan adalah decision tree, rule-based classifier, neural-network, support machine dan naïve Bayes classifier. Setiap teknik menggunakan algoritme pembelajaran untuk mengidentifikasi model yang memberikan hubungan yang paling sesuai antara himpunan atribut dan label kelas dari data input[15].

2.3. Decision Tree

Decision Tree adalah suatu metode yang biasa dipakai untuk melaksanakan klasifikasi pada beberapa data. Metode ini terjadi pada gabungan decision node, disambungkan dengan cabang, hidup melalui dasar pada root node hingga selesai di leaf node. Decision tree adalah prosedur diskriminasi nonlinear yang memakai beberapa faktor independen agar memberi percontohan pada suatu kelompok yang lebih kecil dengan cara berangsur-angsur. Tata cara ini ialah tata cara yang amat terkenal dilakukan sebab hasil pada model yang dibangun lebih mudah untuk dimengerti. Konsep pohon keputusan ialah menukar informasi sebagai pohon keputusan serta ketentuan keputusan. Ada beberapa algoritma yang sering digunakan dalam membangun decision tree, diantaranya adalah[14]

2.4. C5.0

Algoritma C5.0 merupakan metode data mining dengan algoritma klasifikasi yang berbasis pada teknik decision tree. Algoritma ini disempurnakan dari algoritma ID3 dan C4.5 oleh Ross Quinlan pada tahun 1987. Algoritma ini dinilai lebih baik dibanding algoritma sebelumnya dalam hal akurasi dan memori. Kecepatan dalam membuat model dinilai sangat cepat dibanding algoritma lainnya. Algoritma ini juga dapat menangani atribut yang bernilai diskrit maupun kontinu. Kelebihan inilah yang membuat algoritma C5.0 dinilai unggul dibanding algoritma lainnya. Algoritma ini bermula dari semua atribut yang dijadikan akar dari pohon keputusan. Kemudian dipilih atribut yang memiliki nilai gain tertinggi untuk dijadikan root node. Selanjutnya atribut lainnya akan dievaluasi dengan cara yang serupa untuk mendapatkan akar node selanjutnya. Adapun persamaan untuk menghitung information gain dari keseluruhan kasus adalah sebagai berikut[2]:

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (1)$$

dengan:

$I(S_1, S_2, \dots, S_m)$: Keseluruhan kasus pada kelas i
 m : Banyaknya kelas
 p_i : Proporsi dari S_i terhadap S
 S_i : Jumlah kasus pada kelas i
 S : Jumlah Kasus

Sedangkan untuk menghitung nilai entropy ditunjukkan pada persamaan berikut:

$$Entropy(S) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (3)$$

dengan:

S : Nilai kasus
 m : Banyaknya kelas
 p_i : Proporsi dari S_i terhadap S

Untuk mendapatkan nilai gain selanjutnya dihitung dengan formula sebagai berikut:

$$Gain(A) = I(S_1, S_2, \dots, S_m) - Entropy(A) \quad (4)$$

dengan:

A : Variabel
 m : Banyaknya kelas
 $I(S_1, S_2, \dots, S_m)$: Keseluruhan kasus pada kelas i

2.5. Uang Kuliah Tunggal (UKT)

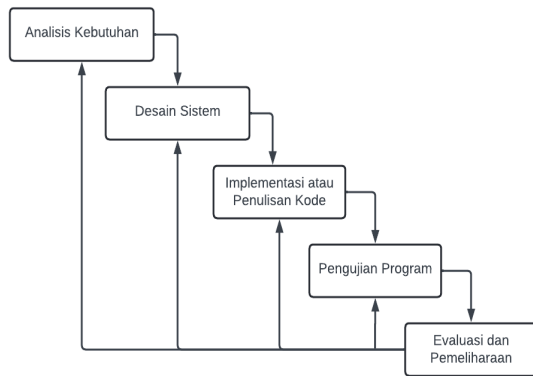
UKT adalah sistem dari beberapa biaya perkuliahan yang dibiayai untuk setiap mahasiswa dalam kemampuan materialnya. Biaya tersebut merupakan semua biaya operasional untuk mahasiswa berprestasi setiap semester pada prodi PTN dan UKT disepakati dari BKT yang dikurangi dengan biaya yang ditanggung oleh pemerintah untuk dilakukan oleh PTN, kebijakan tersebut bertujuan agar meringankan beban uang kuliah yang wajib dikeluarkan oleh mahasiswa dari awal masuk kuliah hingga lulus[17].

2.6. Website

Website (lebih dikenal dengan sebutan situs) adalah sejumlah halaman web yang memiliki topik saling terkait, terkadang disertai pula dengan berkas-berkas gambar, video atau jenis-jenis berkas lainnya. Website adalah kumpulan halaman-halaman yang digunakan untuk menampilkan informasi teks, gambar diam atau gerak, animasi, suara, dan atau gabungan dari semuanya, baik yang bersifat statis maupun dinamis yang membentuk satu rangkaian bangunan yang saling terkait, yang masing-masing dihubungkan dengan jaringan halaman[19].

2.7. Waterfall

Model waterfall (air terjun) sering juga disebut model sekuensial linier (sequential linear) atau alur hidup klasik (classic life cycle) merupakan model yang menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut dimulai dari analisis, desain, pengodean, pengujian, hingga tahap pendukung (support)[24].



Gambar 1. Metode waterfall

a. Analisis Kebutuhan

Analisis kebutuhan merupakan tahapan pertama dari metode SDLC model waterfall, pada tahapan ini dilakukan pengumpulan data, proses pengumpulan data dalam pembuatan sistem ini dilakukan dengan proses wawancara dan observasi lapangan.

b. Desain Sistem

Desain sistem merupakan proses perancangan mulai dari use case diagram, activity diagram, class diagram sampai pembuatan desain user interface dari sistem yang akan dibuat. Fungsi dari desain sistem ini adalah agar pada saat proses implementasi atau penulisan kode lebih mudah, karena yang akan dibuat sudah ada gambaran yang jelas.

c. Implementasi atau Penulisan Kode

Tahap implementasi ini merupakan proses dalam perubahan desain sistem menjadi bentuk kode program, karena dalam penelitian ini sistem dibuat dalam bentuk website.

d. Pengujian Program

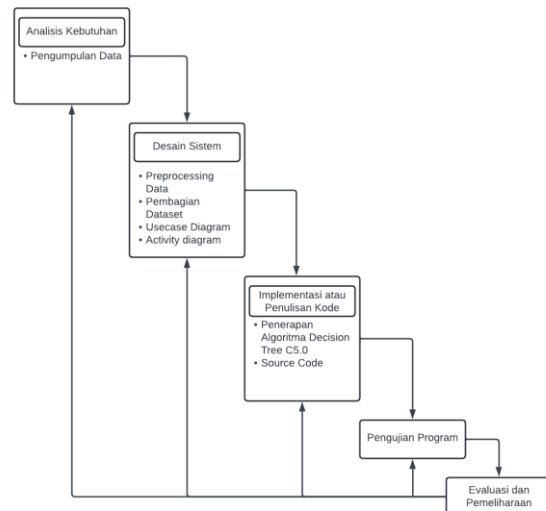
Tahapan pengujian program dilakukan setelah sistem selesai dibuat, pengujian program dilakukan untuk melihat dan memastikan bahwa sistem sudah berjalan seperti yang direncanakan.

e. Evaluasi dan Pemeliharaan

Tahapan terakhir pada metode SDLC waterfall yaitu evaluasi dan pemeliharaan, pada tahapan ini sistem sudah diterapkan secara langsung.

3. METODE PENELITIAN

Penelitian ini menggunakan metode waterfall, yaitu metode pengembangan sistem yang dilakukan secara bertahap dan berurutan, di mana setiap tahapan harus diselesaikan terlebih dahulu sebelum melanjutkan ke tahap berikutnya. Metode ini dipilih karena memiliki alur kerja yang sistematis dan terstruktur sehingga sesuai untuk pengembangan sistem klasifikasi keringanan UKT. Prosedur penelitian dalam metode waterfall bertujuan untuk memperoleh informasi dan solusi yang tepat dalam menjawab permasalahan penelitian.



Gambar 2. Prosedur Penelitian

3.1. Analisis Kebutuhan (Pengumpulan Data)

Tahap ini bertujuan untuk mengidentifikasi kebutuhan sistem dan data yang diperlukan. Data yang digunakan dalam penelitian ini adalah data mahasiswa pengajuan keringanan UKT di Politeknik Negeri Bengkalis. Atribut data meliputi tempat tinggal, status anak bagi orang tua, penghasilan orang tua, pekerjaan orang tua, jumlah tanggungan, kepemilikan DTKS/KIP/KKS/PKH, serta kepemilikan SKTM. Selain itu, dilakukan analisis kebutuhan fungsional sistem, seperti proses input data, proses klasifikasi, serta penyajian hasil keputusan.

3.2. Desain Sistem

Tahap desain sistem dilakukan untuk menggambarkan rancangan sistem sebelum diimplementasikan. Pada tahap ini dilakukan beberapa aktivitas, yaitu:

- Preprocessing data, meliputi pembersihan data, penghapusan data duplikat, serta transformasi data numerik dan kategorik agar dapat diproses oleh algoritma Decision Tree C5.0.
- Pembagian dataset, yaitu membagi data menjadi data latih dan data uji untuk membangun dan menguji model klasifikasi.
- Perancangan sistem, yang dituangkan dalam bentuk usecase diagram untuk menggambarkan interaksi pengguna dengan sistem serta activity diagram untuk menunjukkan alur proses sistem dari awal hingga akhir.

3.3. Implementasi Sistem

Tahap implementasi merupakan tahap penerapan desain sistem ke dalam bentuk program. Sistem dikembangkan menggunakan bahasa pemrograman PHP dengan framework Laravel. Pada tahap ini dilakukan penerapan algoritma Decision Tree C5.0 untuk membangun model klasifikasi kelayakan penerima keringanan UKT berdasarkan data latih yang telah disiapkan.

3.4. Pengujian Program

Pengujian dilakukan untuk memastikan sistem berjalan sesuai dengan kebutuhan dan tujuan penelitian. Model klasifikasi diuji menggunakan data uji dan dievaluasi menggunakan confusion matrix untuk memperoleh nilai akurasi, precision, recall, dan F1-score. Selain itu, pengujian fungsional sistem dilakukan untuk memastikan setiap fitur website dapat berjalan dengan baik.

3.5. Evaluasi dan Pemeliharaan

Tahap evaluasi dilakukan berdasarkan hasil pengujian sistem dan model klasifikasi. Jika

ditemukan kesalahan atau kekurangan, maka dilakukan perbaikan dan penyempurnaan sistem agar kinerja sistem menjadi lebih optimal. Tahap pemeliharaan bertujuan untuk menjaga sistem tetap berjalan dengan baik dan siap digunakan dalam jangka waktu yang panjang.

4. HASIL DAN PEMBAHASAN

Penelitian ini akan menggunakan 80 data training dan 20 data testing. Selanjutnya peneliti akan memastikan setiap atribut pada dataset telah siap untuk diproses. Rangkuman mengenai dataset terangkum pada Tabel 1.

Tabel 1. Dataset

Tempat Tinggal	Status Anak	Pekerjaan	Penghasilan	Tanggungan	DTKS	SKTM
Layak	Lengkap	Petani & IRT	1.000.000	3	Punya	Punya
Layak	Lengkap	Buruh Harian Lepas & IRT	2.000.000	2	Punya	Punya
Layak	Lengkap	Petani & IRT	1.000.000	2	Punya	Punya
Layak	Lengkap	PNS & IRT	3.600.000	4	Tidak Punya	Tidak Punya
Layak	Lengkap	Buruh Harian Lepas & IRT	2.000.000	4	Punya	Punya
Layak	Lengkap	PNS & IRT	. 4.916.700	4	Tidak Punya	Tidak Punya
Layak	Lengkap	Wiraswasta & IRT	. 1.500.000	4	Tidak Punya	Punya
Layak	Lengkap	Petani & IRT	. 2.000.000	4	Tidak Punya	Punya
Layak	Lengkap	Petani & IRT	. 1.000.000	3	Punya	Punya
Layak	Lengkap	Petani & Petani	. 1.000.000	3	Punya	Punya

Tabel 2. Nilai Entropy

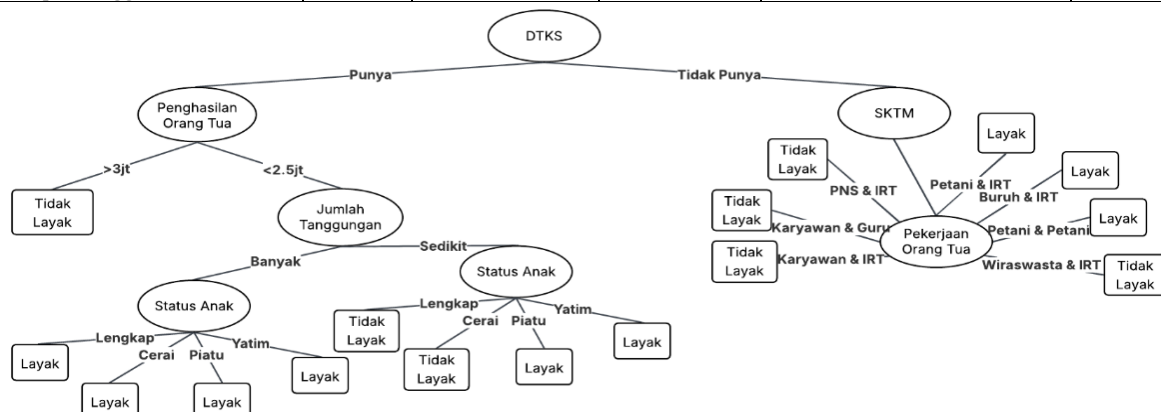
Perhitungan Entropy	10 Data
Punya	6
Tidak Punya	4
Total	10
Probabilitas Punya	0,6
Probabilitas Tidak Punya	0,4
Entropy	0,970950594

Kemudian dilanjutkan dengan data tranformation yaitu mengubah data yang masih mentah dijadikan data yang memiliki value dalam menentukan keputusan, ditunjukkan pada tabel 2.

Perhitungan untuk mencari nilai menggunakan Algoritma C5.0 dilakukan untuk mencari nilai entropy dan nilai gain, untuk selanjutnya dipakai dalam pembuatan pohon keputusan. Data yang digunakan untuk perhitungan dan hasil sebagaimana pada tabel 3.

Tabel 3. Perhitungan Gain

Atribut	Punya	Tidak Punya	Total Data	Entropy Setiap Atribut	Gain
Pekerjaan Orang Tua	12	8	20	0.4825	0.4883
Penghasilan Orang Tua	12	8	20	0.3427	0.6282
Jumlah Tanggungan	11	9	20	0.8091	0.1617
Status Anak	12	8	20	0.8919	0.0789
DTKS	12	8	20	0	0.9709
SKTM	17	3	20	0	0.9709
Tempat Tinggal	12	8	20	0.9709	0



Gambar 3. Pohon Keputusan

4.1. Pengujian

Untuk membuktikan sistem yang dibuat mampu memprediksi kelayakan penerima keringanan UKT, peneliti menggunakan pengujian confusion matrix pada hasil perhitungan yang dilakukan. Uji coba dilakukan dengan menggunakan 80:20 pada data training dan data testing yang didapatkan. Tabel 4 merupakan hasil dari True Positif (TP), True Negatif (TN), False Positif (FP) dan False Negatif (FN) uji coba sistem.

Tabel 4. Hasil Perhitungan

TP	84
TN	0
FP	0
FN	11

Berdasarkan tabel 4, maka akan dilakukan perhitungan akurasi, presisi, recall, dan f1-score dari sistem yang dibuat. Rangkuman mengenai hasil perhitungan tersebut terdapat pada tabel 5.

Tabel 5. Hasil Pengujian

Akurasi	88.42%
Presisi (Layak)	100 %
Recall (Layak)	88.42%
F1-Score	93.85%

Berdasarkan tabel 5, hasil pengujian menunjukkan bahwa nilai True Positive sebesar 84 dan False Negative sebesar 11, sementara tidak ditemukan data pada kelas Tidak Layak. Nilai accuracy yang diperoleh sebesar 88,42%, precision sebesar 100%, recall sebesar 88,42%, dan F1-score sebesar 93,85%. Hasil ini menunjukkan bahwa model mampu mengklasifikasikan data Layak dengan baik, namun evaluasi terhadap kelas Tidak Layak belum dapat dilakukan karena keterbatasan data.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian model klasifikasi menggunakan confusion matrix, diperoleh nilai True Positive sebesar 84 dan False Negative sebesar 11, sementara tidak ditemukan data pada kelas Tidak Layak. Nilai akurasi yang diperoleh sebesar 88,42%, precision sebesar 100%, recall sebesar 88,42%, dan F1-score sebesar 93,85%. Hasil tersebut menunjukkan bahwa model yang dibangun mampu mengklasifikasikan data pada kelas Layak dengan baik dan memiliki tingkat ketepatan prediksi yang cukup tinggi. Namun demikian, keterbatasan data yang hanya mencakup satu kelas menyebabkan evaluasi terhadap kelas Tidak Layak belum dapat dilakukan secara menyeluruh.

Berdasarkan keterbatasan yang ditemukan dalam penelitian ini, disarankan pada penelitian selanjutnya untuk menggunakan data dengan distribusi kelas yang lebih seimbang, khususnya dengan menambahkan data

pada kelas Tidak Layak. Hal ini bertujuan agar evaluasi kinerja model dapat dilakukan secara lebih optimal dan menyeluruh. Selain itu, penelitian selanjutnya juga dapat mempertimbangkan penggunaan metode klasifikasi lain atau pengujian dengan jumlah data yang lebih besar untuk meningkatkan akurasi dan keandalan sistem dalam menentukan kelayakan keringanan UKT.

DAFTAR PUSTAKA

- [1] T. Menggunakan, C. Algoritma, N. S. Ramadhan, and W. N. Nursafitri, "Analisis Klasifikasi UKT Mahasiswa Berdasarkan Tingkat Penghasilan Orang," vol. 2, no. 1, pp. 1–9, 2024.
- [2] R. N. Amalda, N. Millah, and I. Fitria, "Implementasi algoritma c5.0 dalam menganalisa kelayakan penerima keringanan ukt mahasiswa itk," vol. 7, no. 1, pp. 101–116, 2022.
- [3] L. A. Larytasari and Y. Susanti, "PENENTUAN UKT MAHASISWA UNS DENGAN ALGORITMA ITERATIVE," 2019.
- [4] U. Islam, N. Sumatera, and U. Medan, "IMPLEMENTASI DATA MINING DALAM MENGLASIFIKASIKAN UKT (UANG KULIAH TUNGGAL) PADA UIN SUMATERA UTARA MEDAN," vol. 4, no. 2, pp. 370–376, 2020.
- [5] H. Kurniawan and S. Defit, "JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST) Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," vol. 1, no. 2, pp. 80–89, 2020.
- [6] W. Ridwan *et al.*, "Penentuan Uang Kuliah Tunggal Mahasiswa Universitas Negeri Gorontalo Berbasis Logika Fuzzy," vol. 3, pp. 62–65, 2021.
- [7] I. N. Rois, A. Wijaya, and A. V. Prananingrum, "Implementasi Keringanan Uang Kuliah Tunggal Bagi Mahasiswa di Era New Normal," vol. 1, no. September, pp. 103–121, 2020.
- [8] F. Ida, S. Muhammad, and A. Calam, "Implementasi Algoritma K-Means Custering Dalam Pengelompokan Uang Kuliah Tunggal Mahasiswa pada STAIN Mandailing Natal," vol. 3, no. 9, pp. 1–13, 2020.
- [9] U. Kuliah and T. Ukt, "IMPLEMENTASI ALGORITMA C4.5 UNTUK MENENTUKAN UANG KULIAH TUNGGAL (UKT) MAHASISWA (Studi Kasus : Mahasiswa UMRAH Tahun 2018) Farida," pp. 88–97, 2020.
- [10] N. Hajar, N. Y. Setiawan, and F. A. Bachtiar, "Pengelompokan Mahasiswa untuk Pengajuan Bantuan Uang Kuliah Tunggal menggunakan Metode K-Means Clustering (Studi Kasus BEM FILKOM UB)," vol. 6, no. 5, pp. 2353–2361, 2022.