

Penentuan Kriteria Calon Penerima Beasiswa Dengan Algoritma C4.5 (Studi Kasus : Universitas Potensi Utama)

Determination of Criteria for Scholarship Recipients Using the C4.5 C Algorithm (Case Study : Universitas Potensi Utama)

Elsa Aditya¹, Rahmadani Siregar², Sofyan Rahmad³

^{1,2,3}Fakultas Teknik dan Ilmu Komputer, Universitas Potensi Utama

Jl.K.L. Yos Sudarso KM.6,5 No. 3A, Tj. Mulia, Medan

E-mail: elsaaditya15@gmail.com¹, rahmadanisrgr@gmail.com², sofyandrahmad99@gmail.com³

Abstrak

Beasiswa pada dasarnya adalah pemberian berbentuk bantuan finansial yang diberikan bagi setiap orang yang memiliki kriteria tertentu demi tetap berlangsungnya pendidikan yang di tempuh. Metode Decision Tree sering digunakan untuk dijadikan pengambilan keputusan dalam suatu kasus. Pada penelitian ini, algoritma C4.5 digunakan untuk menentukan kriteria calon penerima beasiswa dengan cara menghitung jumlah entropy dan gain pada setiap atribut. Permasalahan yang terjadi pada penentuan kriteria penerima beasiswa adalah kurang efektifnya penentuan penerima beasiswa dikarenakan terdapat banyak faktor yang mempengaruhi apakah seseorang dapat memperoleh beasiswa atau tidak. Tujuan pembuatan penelitian ini adalah untuk mengetahui kriteria penentuan calon penerima beasiswa secara tepat. Penelitian ini dilakukan dengan menggunakan 30 (tiga puluh) sampel data mahasiswa calon penerima beasiswa. Berdasarkan hasil perhitungan persentase Correctly Classified Instance adalah sebesar 90% dan persentase Incorrectly Classified Instance adalah sebesar 10%. Pada perhitungan dengan algoritma C4.5 ini, terdapat keputusan jika orang tua mahasiswa berpenghasilan rendah, maka mahasiswa dapat menerima beasiswa.

Kata Kunci—Decision Tree, C4.5, Beasiswa

Abstract

Scholarships are basically the provision of financial assistance given to everyone who has certain criteria for the continuation of the education taken. The Decision Tree method is often used to make decisions in a case. In this study, the C4.5 algorithm is used to determine the criteria for prospective scholarship recipients by calculating the amount of entropy and gain on each attribute. The problem that occurs when planning a scholarship is the ineffectiveness of prospective scholarship recipients because there are factors that affect whether someone can get a scholarship or not. The purpose of making this research is to find out when the prospective scholarship recipients are right. This research was conducted using 30 (thirty) data samples of prospective scholarship recipients. the result based on the calculation of Correctly Classified Instances is 90% and the proportion of Correctly Classified Instances is 10%. In the calculation with the C4.5 algorithm, there is a decision if the student's parents are low, then the student can receive a scholarship.

Keywords—3-5 Decision Tree, C4.5, Scholarships

1. PENDAHULUAN

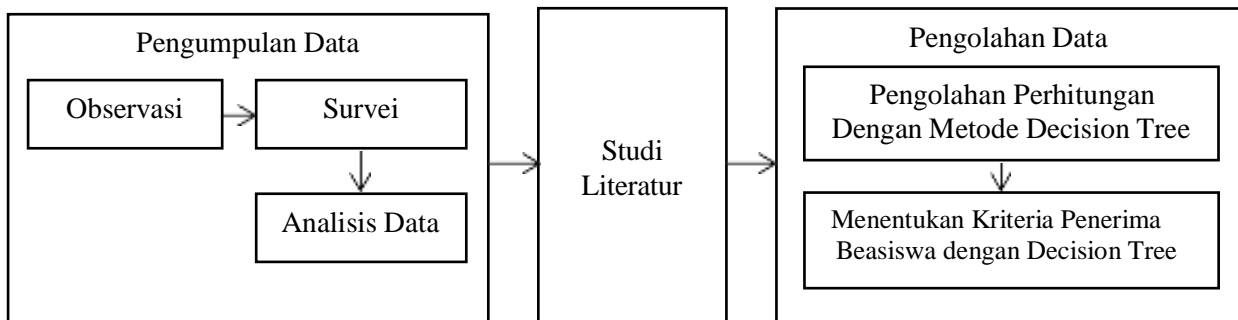
Beasiswa pada dasarnya adalah pemberian berbentuk bantuan finansial yang diberikan bagi setiap orang yang memiliki kriteria tertentu demi tetap berlangsungnya pendidikan yang di tempuh. Beasiswa bisa berasal dari mana saja, baik dari instansi pemerintahan, perusahaan maupun yayasan. Beasiswa merupakan suatu pemberian dengan harapan dapat meningkatkan semangat untuk terus belajar. Pada setiap kegiatan seleksi pendaftar beasiswa terdapat banyak faktor yang mempengaruhi apakah seseorang dapat memperoleh beasiswa. Permasalahan yang sering terjadi pada saat pengambilan keputusan untuk menentukan mahasiswa yang layak mendapat beasiswa adalah terkendala dikarenakan kuota beasiswa yang terbatas akan tetapi pendaftar beasiswa yang banyak dan juga perlu seleksi yang panjang untuk menentukan faktor yang dapat mendukung untuk memperoleh beasiswa. Penelitian ini diharapkan agar pemberian beasiswa dapat dilakukan efektif sehingga beasiswa yang diberikan dapat dimanfaatkan dengan sebaik mungkin. Untuk mendapatkan hasil seleksi yang tepat, calon penerima beasiswa haruslah di seleksi dengan sebaik mungkin agar beasiswa yang diberikan tepat sasaran.

Pada penelitian ini dilakukan penentuan kriteria penerima beasiswa dengan menggunakan metode *Decision Tree* algoritma C4.5, algoritma ini digunakan berdasarkan oleh penelitian terdahulu. Pemberian beasiswa diklasifikasikan menggunakan algoritma C4.5 dan menciptakan aturan baru untuk mendapatkan indikator penilaian yang sesuai dengan penentuan pemberian rekomendasi beasiswa[1]. Decision tree banyak digunakan untuk melakukan perhitungan sebuah prediksi dikarenakan decision tree memiliki kelebihan yaitu dapat menganalisa data yang besar secara lebih efektif dan lebih mudah di fahami oleh pengguna [2].

2. METODE PENELITIAN

2.1. Teknik Pengumpulan Data

Adapun metodologi penelitian dari Penentuan Kriteria Calon Penerima Beasiswa Dengan Algoritma C4.5 adalah sebagai berikut :



Gambar 1. Alur Analisa Penelitian

Pada tahap awal penelitian dilakukan proses pengumpulan data melalui observasi ke bagian yang bersangkutan dan melakukan survei kepada pegawai yang menangani bagian kemahasiswaan lalu menganalisa data yang di berikan oleh bagian tersebut. Tahap selanjutnya adalah melakukan studi literatur dan melakukan pengolahan data dengan melakukan perhitungan dengan menggunakan metode Decision Tree algoritma C4.5 dan memperoleh data Kriteria Penerima Beasiswa.

2.2. Metode Decision Tree

Algoritma *Decision Tree* adalah algoritma yang kuat, populer, berbasis logika, mudah ditafsirkan, lurus ke depan, dan dapat diterapkan secara luas untuk beberapa masalah dalam penambangan data [3]. *Decision Tree* adalah struktur yang mencakup simpul akar, cabang, dan simpul daun. Setiap simpul internal menunjukkan tes pada atribut, setiap cabang menunjukkan hasil tes, dan setiap simpul daun memegang label kelas. *Decision Tree* adalah struktur pohon seperti diagram alur, di mana setiap simpul internal mewakili tes pada atribut, setiap cabang mewakili hasil tes, label kelas diwakili oleh setiap simpul daun (atau simpul terminal) [3]. *Decision Tree* merupakan metode pengambilan keputusan yang paling banyak diminati dikarenakan proses perhitungannya relatif cepat juga hasil dari pohon keputusan tersebut lebih mudah dipahami. Pohon keputusan adalah struktur pohon seperti diagram alur, di mana setiap simpul internal (bukan simpul daun) menunjukkan tes pada atribut, setiap cabang mewakili hasil tes, dan setiap simpul daun (atau simpul terminal) memegang label kelas[4]. *Decision tree* merupakan sebuah diagram alir yang berbentuk seperti struktur sebuah pohon yang mana setiap node menyatakan tentang pengujian terhadap suatu atribut, setiap cabang menyatakan sebuah output dari pegujian tersebut serta leaf node menyatakan kelas atau distribusi kelas. Node yang teratas disebut dengan root node (node akar) yang memiliki edge keluar akan tetapi tidak memiliki edge masuk [5].

2.3. Algoritma C4.5

Algoritma C4.5 adalah algoritma yang digunakan untuk menghasilkan pohon keputusan dikembangkan oleh Ross Quinlan. C4.5 adalah perpanjangan dari Algoritma ID3 Quinlan sebelumnya. Pohon keputusan yang dihasilkan oleh C4.5 dapat digunakan untuk klasifikasi dan untuk alasan ini C4.5 sering disebut sebagai pengklasifikasi statistik. Seperti membelah kriteria, algoritma C4.5 menggunakan information gain. Itu bisa menerima data dengan nilai kategorikal atau numerik. Ambang adalah dihasilkan untuk menangani nilai kontinu dan kemudian membagi atribut dengan nilai di atas ambang batas dan nilai sama dengan atau di bawah ambang batas. Algoritma C4.5 dapat dengan mudah menangani nilai yang hilang, karena nilai atribut yang hilang tidak digunakan dalam mendapatkan perhitungan dengan C4.5[3]. Algoritma ini merupakan algoritma yang digunakan untuk membentuk sebuah pohon keputusan. Pohon keputusan dapat di artikan sebagai suatu cara untuk memprediksi atau mengklasifikasi yang sangat kuat [6]. Sebuah pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan record yang lebih kecil dengan menerapkan serangkaian keputusan. Terdapat dua rumus algoritma C4.5 yaitu rumus untuk mencari nilai gain dan nilai entropy, adapun rumus-rumus tersebut seperti dibawah ini :

Rumus pertama :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan :

S : Himpunan Kasus

A : Atribut

|S_i| : Jumlah Partisi Atribut A

|S| : Jumlah Kasus dalam S

Rumus kedua :

$$Entropy (S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

Keterangan :

S : Himpunan Kasus

n : Banyaknya Partisi S

p_i : Probabilitas yang didapat dari kasus i dibagi total kasus.

3. HASIL DAN PEMBAHASAN

3.1. Perhitungan Metode C4.5

Pada tahap ini dilakukan proses perhitungan menggunakan metode C4.5, terdapat beberapa atribut yang digunakan untuk mencari kriteria calon penerima beasiswa yaitu jenis kelamin, asal kota, pendapatan orang tua, IPK, dimana pada masing-masing atribut tersebut dapat mempengaruhi hasil dari entropy dan gain. Dalam data tersebut, terdapat 18 mahasiswa yang dapat menerima beasiswa dan 12 mahasiswa yang tidak dapat menerima beasiswa, dimana data keseluruhan jumlah mahasiswa tersebut menjadi data awal untuk memulai perhitungan menggunakan rumus metode C4.5.

3.2. Data Latih

Berikut adalah data latih yang akan digunakan untuk menentukan kriteria calon penerima beasiswa dengan algoritma c4.5 :

Tabel 1. Data Pengajuan Beasiswa Mahasiswa

No.	Jenis Kelamin	Asal Kota	Penghasilan Orang Tua	IPK	Beasiswa
1	L	Medan	Tinggi	Sedang	Tidak Menerima
2	L	Luar Medan	Sedang	Rendah	Tidak Menerima
3	L	Medan	Sedang	Tinggi	Menerima
4	L	Medan	Sedang	Sedang	Tidak Menerima
5	L	Medan	Sedang	Sedang	Menerima
6	L	Medan	Sedang	Sedang	Tidak Menerima
7	P	Medan	Tinggi	Sedang	Tidak Menerima
8	L	Medan	Tinggi	Sedang	Tidak Menerima
9	L	Luar Medan	Sedang	Sedang	Menerima
10	L	Medan	Sedang	Sedang	Tidak Menerima
11	L	Medan	Sedang	Tinggi	Tidak Menerima
12	P	Luar Medan	Sedang	Tinggi	Menerima
13	L	Luar Medan	Sedang	Tinggi	Menerima
14	P	Medan	Rendah	Tinggi	Menerima
15	L	Medan	Sedang	Tinggi	Menerima
16	L	Medan	Rendah	Sedang	Menerima

17	L	Medan	Rendah	Sedang	Menerima
18	L	Medan	Rendah	Sedang	Menerima
19	L	Medan	Tinggi	Sedang	Tidak Menerima
20	L	Medan	Sedang	Tinggi	Menerima
21	L	Medan	Sedang	Tinggi	Menerima
22	L	Medan	Sedang	Tinggi	Menerima
23	L	Medan	Rendah	Sedang	Menerima
24	L	Medan	Sedang	Sedang	Tidak Menerima
25	P	Medan	Rendah	Tinggi	Menerima
26	P	Medan	Rendah	Tinggi	Menerima
27	L	Luar Medan	Rendah	Sedang	Menerima
28	L	Medan	Rendah	Sedang	Menerima
29	L	Medan	Sedang	Sedang	Tidak Menerima
30	P	Medan	Sedang	Sedang	Tidak Menerima

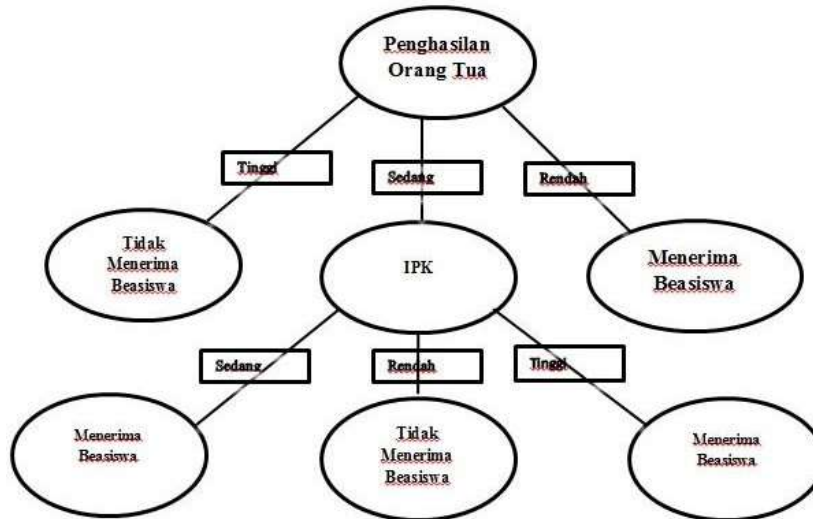
Data yang ada lalu di proses menggunakan rumus untuk mencari nilai Entropy dan nilai Gain tertinggi, dengan rumus yang telah di tetapkan. Berikut adalah perhitungan untuk mencari simpul akar sebuah pohon keputusan :

Tabel 2. Hasil Perhitungan Node 1

Node	Atribut	Kelas	Jumlah Kasus	Menerima Beasiswa	Tidak Menerima Beasiswa	Entropy	Gain
1	Total		30	18	12	0.970951	
	IPK						0.2151561
		Rendah	1	0	1	0	
		Sedang	18	8	10	0.991076	
		Tinggi	11	10	1	0.439497	
	Jenis Kelamin						0.0033964
		Laki-laki	24	14	10	0.979869	
		Perempuan	6	4	2	0.918296	
	Asal Kota						0.025973
		Medan	25	14	11	0.989588	
		Luar Medan	5	4	1	0.721928	
	Penghasilan Orang Tua						0.405699
		Rendah	9	9	0	0	
		Sedang	17	9	8	0.997503	
		Tinggi	4	0	4	0	

Pada tabel 2, dapat dilihat yaitu dalam atribut IPK, terdapat 3 nilai atribut yaitu atribut Rendah, Sedang dan Tinggi dimana nilai dari atribut IPK “Tinggi” telah mengklasifikasikan kasus menjadi sebuah keputusan yaitu dapat “Menerima Beasiswa”, serta nilai dari atribut IPK “Rendah” juga telah mengklasifikasikan yaitu “Tidak Menerima Beasiswa”, untuk nilai dari atribut IPK “Sedang” telah mengklasifikasikan kasus menjadi sebuah keputusan yaitu dapat “Menerima Beasiswa”, maka dapat dibentuk sebuah node baru dan juga menjadi node terakhir.

Berikut adalah gambar pohon keputusan dengan node terakhir :



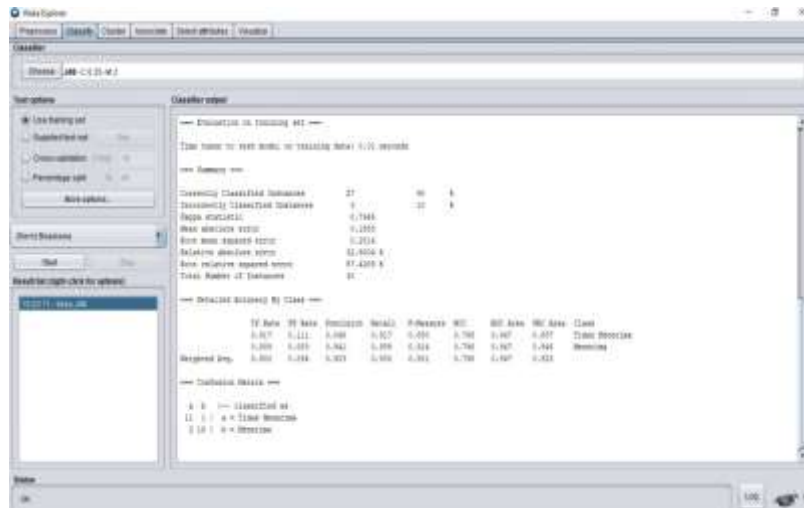
Gambar 2. Decision Tree

Dari pohon keputusan diatas maka dapat di ambil kesimpulan sebagai berikut :

1. IF “Penghasilan Orang Tua” mahasiswa bernilai “Tinggi” THEN “Tidak Menerima Beasiswa”.
2. IF “Penghasilan Orang Tua” mahasiswa bernilai “Rendah” THEN “Menerima Beasiswa”.
3. IF “Penghasilan Orang Tua” mahasiswa bernilai “Sedang” AND “IPK” adalah “Rendah” THEN “Tidak Menerima Beasiswa”.
4. IF “Penghasilan Orang Tua” mahasiswa bernilai “Sedang” AND “IPK” mahasiswa bernilai “Sedang” THEN “Menerima Beasiswa”.
5. IF “Penghasilan Orang Tua” mahasiswa bernilai “Sedang” AND “IPK” mahasiswa bernilai “Tinggi” THEN “Menerima Beasiswa”.

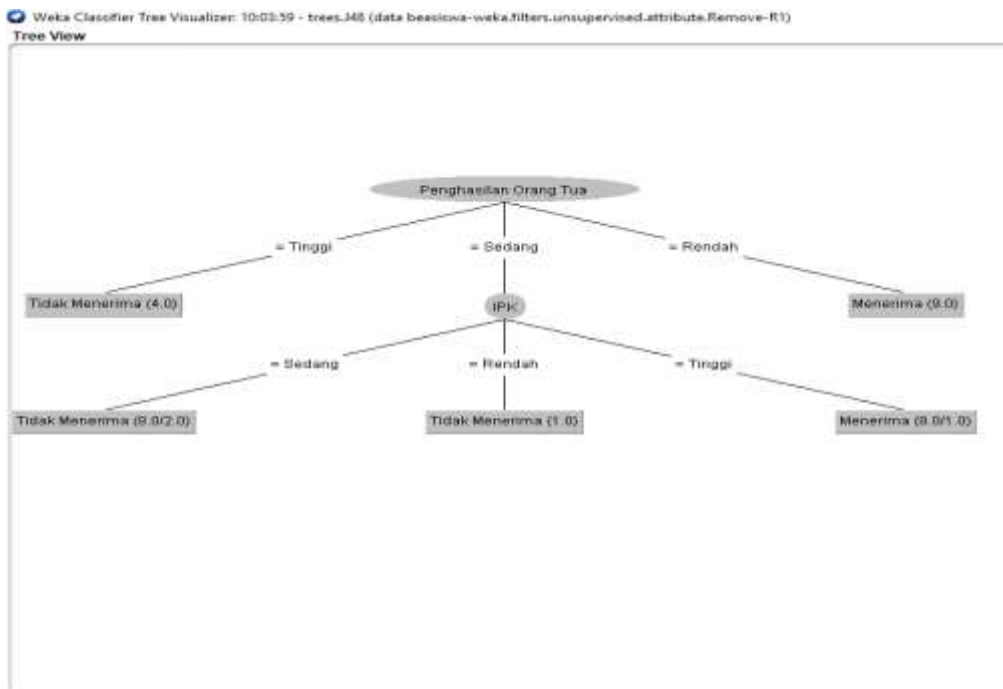
3.3. Pengujian Metode C4.5

Berdasarkan data yang sudah diperoleh dari hasil pengujian dengan metode C4.5, maka data hasil pengujian tersebut akan di uji kedalam tools WEKA yaitu sebanyak 30 data mahasiswa sehingga mendapatkan hasil prediksi kelulusan mahasiswa dimana untuk persentase *Correctly Classified Instance* adalah sebesar 90% dan *persentase Incorrectly Classified Instance* adalah sebesar 10%. Dari 30 data, terdapat 18 data mahasiswa yang dapat menerima beasiswa dan 12 data mahasiswa yang tidak mendapatkan beasiswa yang berhasil diklasifikasikan dengan benar.



Gambar 5. Pengujian Data Menggunakan Tools Weka
 Sumber : Aplikasi WEKA 3.9

Dari pengujian data menggunakan tools Weka, maka dapat diperoleh sebuah pohon keputusan sebagai berikut :



Gambar 6. Pohon Keputusan Menggunakan Tools Weka
 Sumber : Aplikasi WEKA 3.9

4. KESIMPULAN

Berdasarkan dari penelitian yang telah dilakukan, maka dapat diambil kesimpulan beberapa kesimpulan dari Penentuan Kriteria Calon Penerima Beasiswa Dengan Algoritma C4.5, diantaranya sebagai berikut :

1. Dari perhitungan menggunakan metode *Decision Tree* algoritma C4.5 maka terdapat menghasilkan keputusan bahwa mahasiswa yang dapat menerima beasiswa adalah mahasiswa yang orang tua berpenghasilan “Rendah”.
2. Berdasarkan data latih penentuan kriteria calon penerima beasiswa dengan algoritma C4.5 sebanyak 30 data mahasiswa yang dihitung menggunakan Tools Weka terdapat hasil persentase *Correctly Classified Instance* adalah sebesar 90% dan persentase *Incorrectly Classified Instance* adalah sebesar 10%.
3. Kelebihan metode *Decision tree* ini memiliki hasil perhitungan yang lebih akurat dan juga dapat menghilangkan atribut perhitungan yang tidak diperlukan untuk membentuk sebuah pohon.
4. Kelemahan metode *Decision Tree* ini yaitu sulit untuk merancang pohon keputusan yang optimal.

5. SARAN

Sebaiknya dilakukan penelitian lebih lanjut mengenai perbandingan perhitungan antara metode decision tree dengan metode dan algoritma lainnya sehingga dapat dilihat hasil perbandingan yang lebih akurat dalam penentuan kriteria calon penerima beasiswa dengan metode lainnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh teman-teman yang sudah memberikan kontribusi sehingga penulis dapat menyelesaikan penelitian ini, dan serta para Pengajar juga para pegawai di Universitas Potensi Utama yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] E. Hasmin and S. Aisa, "Penerapan Algoritma C4.5 Untuk Penentuan Penerima Beasiswa Mahasiswa," *CogITO Smart J.*, vol. 5, no. 2, p. 308, 2019, doi: 10.31154/cogito.v5i2.219.308-320.
- [2] H. Hafizan and A. N. Putri, "Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi Balita Di Kabupaten Simalungun," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 1, no. 2, pp. 68–72, 2020, doi: 10.30645/kesatria.v1i2.23.
- [3] H. Sharma and S. Kumar, "A Survey on Decision Tree Algorithms of Classification in Data Mining," *Int. J. Sci. Res.*, vol. 5, no. 4, pp. 2094–2097, 2016, doi: 10.21275/v5i4.nov162954.
- [4] L. A. Gaurav and G. Hitesh, "Improved C4.5 decision tree," *Int. Conf. Internet Technol. Appl. ITAP 2010 - Proc.*, vol. 3, no. 3, pp. 341–345, 2010, doi: 10.1109/ITAPP.2010.5566133.
- [5] P. Kasih, "Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara," *Innov. Res. Informatics*, vol. 1, no. 2, pp. 63–69, 2019, doi: 10.37058/innovatics.v1i2.918.
- [6] F. A. Sianturi, "Analisa Decision Tree Dalam Pengolahan Data Siswa," *MEANS (Media Inf. Anal. dan Sist.*, vol. 3, no. 2, pp. 166–172, 2018, [Online]. Available: http://ejournal.ust.ac.id/index.php/Jurnal_Means/.
- [7] Hartono, H., Sadikin, M., Sari, D. M., Anzelina, N., Lestari, S., & Dari, W. (2020). Implementation of Artificial Neural Networks with Multilayer Perceptron for Analysis of Acceptance of Permanent Lecturers. *Jurnal Mantik*, 4(2), 1389-1396.