

# Pemetaan Sasaran *Marketing* Calon Mahasiswa Baru UBHARA Jaya Menggunakan Algoritma C4.5

Efrianto Nugroho Putro <sup>1</sup>, Rafika Sari <sup>1,\*</sup>

\* Korespondensi: e-mail: [rafika.sari@dsn.ubharajaya.ac.id](mailto:rafika.sari@dsn.ubharajaya.ac.id)

<sup>1</sup> Informatika; Universitas Bhayangkara Jakarta Raya; Jl. Raya Perjuangan, Margamulya, Bekasi, Jawa Barat, No.Telp. (021) 88955882; e-mail: [efriantonugrohoputro@gmail.com](mailto:efriantonugrohoputro@gmail.com), [rafika.sari@dsn.ubharajaya.ac.id](mailto:rafika.sari@dsn.ubharajaya.ac.id)

Submitted : 8 April 2023  
Revised : 21 April 2023  
Accepted : 15 Mei 2023  
Published : 30 Mei 2023

## Abstract

*Information technology developed by higher education institutions must be able to solve problems and be the right solution for the effectiveness of the performance of a higher education institution. One of them is in predicting the acceptance of prospective new students which is an important process to test how much public interest is in becoming a student on a campus. In this study, a predictive model was built using a data mining approach. The purpose of this study is to show predictions of the acceptance of prospective new students regarding the attractiveness of prospective students at Bhayangkara Jakarta Raya University (Ubhara) when viewed from the location data of the prospective student's school of origin. So that Ubhara can adjust his marketing strategy in introducing the Ubhara to prospective students. This study uses a dataset of prospective new students who registered at Ubhara in 2019 and 2020 and selected 50 datasets as samples and applied data mining techniques using the C.45 algorithm. Data mining techniques are built through Python programming using the stages of data selection, checking for missing data values and data transformation. The results of this study are used as an evaluation and recommendation for marketing using a confusion matrix, an accuracy rate of 60% is obtained. This can be improved by using more datasets.*

**Keywords:** C4.5 Algorithm, New Students, Marketing Strategy

## Abstrak

Teknologi informasi yang dikembangkan oleh lembaga pendidikan perguruan tinggi harus dapat menyelesaikan permasalahan dan menjadi solusi yang tepat bagi efektifitas kinerja sebuah perguruan tinggi. Salah satunya dalam melakukan prediksi penerimaan calon mahasiswa baru yang merupakan suatu proses penting untuk menguji seberapa besar daya tarik masyarakat untuk menjadi mahasiswa disuatu kampus. Pada penelitian ini, dibangunlah model prediksi melalui teknik pendekatan data mining. Tujuan penelitian ini yaitu menunjukkan prediksi penerimaan calon mahasiswa baru mengenai daya tarik calon mahasiswa Universitas Bhayangkara Jakarta Raya (Ubhara) jika dilihat dari data lokasi sekolah asal calon mahasiswa tersebut. Sehingga Ubhara dapat menyesuaikan strategi marketingnya dalam memperkenalkan Ubahar kepada calon mahasiswa. Penelitian ini menggunakan dataset calon mahasiswa baru yang mendaftar ke Ubhara pada tahun 2019 dan 2020 dan dipilih 50 dataset sebagai sampel serta penerapan teknik data mining menggunakan algoritma C4.5. Teknik data mining dibangun melalui pemrograman Python dengan menggunakan tahapan-tahapan data seleksi, cek data *missing values* dan transformasi data. Hasil penelitian ini digunakan sebagai evaluasi dan rekomendasi untuk pemasaran menggunakan confusion matrix, diperoleh tingkat akurasi sebesar 60% hal ini

dapat ditingkatkan dengan penggunaan *dataset* yang lebih banyak.

**Kata kunci:** Algoritma C4.5, Mahasiswa Baru, Sasaran Marketing

---

## 1. Pendahuluan

Kemajuan teknologi yang berkembang pesat saat ini digunakan untuk membantu menyelesaikan pekerjaan dengan cepat dan efisien. Perkembangan tersebut mendorong kita untuk lebih membuka diri pada perubahan-perubahan yang ada di era globalisasi seperti sekarang ini (Hafiz, 2020). Manfaat teknologi ini sudah dirasakan oleh semua kalangan masyarakat dari masyarakat kota hingga pedesaan. Pengguna dari kalangan masyarakat di berbagai bidang, mulai dari bidang ekonomi, pemerintahan, sosial, budaya hingga bidang pendidikan. Penerapan teknologi dalam bidang pendidikan sangat sering memanfaatkan perkembangan teknologi informasi sebagai penunjang kegiatan belajar mengajar di dalam kampus. Salah satunya penerapan teknologi dalam bidang pendidikan untuk menghasilkan mutu pendidikan yang lebih baik termasuk dalam hal memberikan prediksi penerimaan calon mahasiswa baru di perguruan tinggi.

Teknologi informasi yang dikembangkan oleh lembaga pendidikan perguruan tinggi harus mampu menyelesaikan permasalahan dan menjadi solusi yang tepat bagi efektifitas kinerja sebuah perguruan tinggi. Salah satunya dalam proses melakukan prediksi penerimaan calon mahasiswa baru yang merupakan suatu proses penting untuk menguji seberapa besar daya tarik masyarakat untuk menjadi mahasiswa disuatu perguruan tinggi. Dari hasil prediksi yang akurat akan menentukan strategi yang tepat untuk melakukan proses *marketing* serta pengenalan kampus kepada calon pendaftar disuatu Perguruan Tinggi Negeri.

Proses prediksi penerimaan calon mahasiswa baru di Universitas Bhayangkara Jakarta Raya (UBHARA Jaya) setiap tahunnya dilakukan dengan cara manual, hanya menggunakan perhitungan minat dan daya tarik berdasarkan rata-rata penerimaan mahasiswa baru setiap tahunnya. Tentunya proses prediksi yang dilakukan secara manual belum ada pengujian tingkat akurasi perhitungannya maka dari itu diperlukan suatu metode yang baik untuk membantu proses prediksi penerimaan calon mahasiswa baru di Ubhara supaya mendapatkan hasil yang lebih akurat dan dapat dijadikan landasan untuk rekomendasi peningkatan sasaran *marketing* khususnya bagi Biro Humas dan Pemasaran di Ubhara.

Teknologi yang paling tepat dalam proses prediksi salah satunya dengan menerapkan data mining dalam melakukan proses prediksi. Tujuan dari data mining adalah untuk mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat (Kusrini & Taufiq Emha, 2009; Nofriansyah, 2015). Dengan sumber data yang akurat yang kemudian diolah dengan algoritma tertentu sehingga menghasilkan suatu hasil prediksi dan rekomendasi. Dalam penerapannya data mining memerlukan berbagai perangkat lunak analisis data untuk menemukan pola dan relasi data agar dapat digunakan untuk membuat prediksi dan klasifikasi

dengan akurat (Budiman et al., 2015). Metode yang dapat digunakan untuk membantu proses prediksi penerimaan calon mahasiswa baru di UBHARA adalah algoritma C4.5.

Berdasarkan uraian permasalahan, penelitian ini akan berfokus pada tingkat akurasi dari perhitungan algoritma C4.5 dalam proses prediksi penerimaan calon mahasiswa baru di UBHARA. Sehingga hasil perhitungan tersebut dapat dijadikan landasan rekomendasi untuk meningkatkan strategi *marketing* kedepannya, lokasi mana yang menjadi prioritas UBHARA dalam pemasarannya.

Beberapa penelitian terdahulu yang dijadikan rujukan diantaranya penelitian yang dilakukan oleh (Sari et al., 2022, 2023), pada 2023 mengenai sistem pendukung keputusan penentu bidang peminatan tugas akhir mahasiswa menggunakan algoritma C4.5 memberikan hasil delapan rules model rekomendasi untuk mahasiswa tingkat akhir dalam mengambil bidang peminatan skripsi dengan tingkat akurasi 70%, dan pada 2022 mengenai sistem pendukung keputusan penerima bantuan langsung tunai (BLT) mengolah data secara tepat sasaran sesuai dengan kriteria yang sudah ditetapkan memberikan hasil calon penerima sebesar 25% data lolos sebagai penerima bantuan BLT. Selanjutnya penelitian (Rufiyanto et al., 2020) terkait penggunaan metode C4.5 untuk memprediksi tingkat kepuasan mahasiswa terhadap layanan akademik dengan melibatkan lima aspek kualitas pelayanan yaitu aspek tangibles, reliability, responsiveness, assurance, dan empathy. Kemudian penelitian (Rahmadani et al., 2020) dalam memetakan mata kuliah dan keterlambatan kelulusan mahasiswa menggunakan algoritma C4.5 dan *K-Nearest Neighbors* dengan mengolah data sehingga diperoleh faktor yang menyebabkan mahasiswa terhambat dalam kelulusan dan memberikan rekomendasi terhadap factor yang sudah ditemukan dari hasil data mining menggunakan metode C4.5 dan KNN.

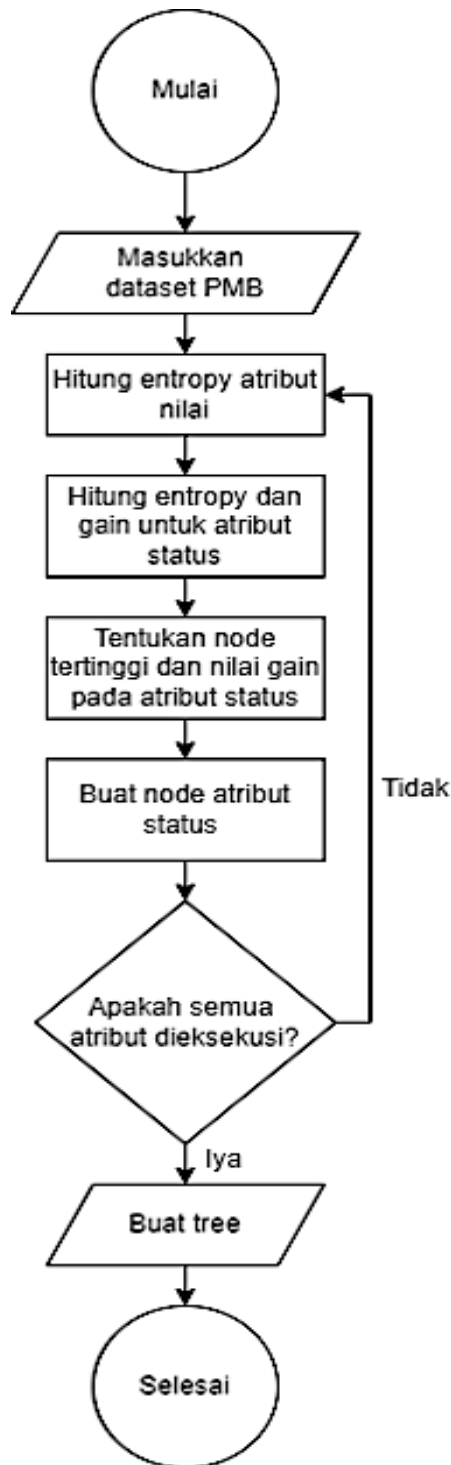
## **2. Metode Penelitian**

Pada penelitian ini, data yang digunakan berupa data penerimaan mahasiswa baru UBHARA pada tahun 2019 dan 2020. Tahapan pertama *data mining* yaitu *data preparation* sebelum masuk kedalam pengolahan dan evaluasi model menggunakan algoritma C4.5. Pada tahapan *data preparation*, dilakukan pemilihan data, pembersihan data dan transformasi data. Selanjutnya akhir dari tahapan *data mining* yaitu analisis pada data. Penelitian menggunakan algoritma C4.5 untuk memperoleh model pohon keputusan. Selanjutnya pengujian model dilakukan dengan menggunakan *confusion matrix*. Hal ini dapat memberikan hasil tingkat akurasi dataset mengenai aktivitas *marketing* yang akan dilakukan UBHARA kedepannya. Adapun tahapan perhitungan algoritma C4.5 dalam penelitian ini disajikan pada Gambar 1.

Secara umum algoritma C4.5 digunakan untuk membangun pohon keputusan dengan tahapan sebagai berikut (Kusrini & Taufiq Emha, 2009):

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap-tiap nilai.
- c. Bagi kasus dalam cabang.

- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.



Sumber: Hasil Penelitian (2022)

Gambar 1. Flowchart Penelitian Menggunakan Algoritma C4.5

Didalam algoritma C4.5 terdapat dua konsep dalam pohon keputusannya. Adapun dua konsep tersebut diformulasikan sebagai: (i) Entropy ( $S$ ), merupakan jumlah bit yang diperkirakan dan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada

ruang sampel  $S$ . Entropy dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. C4.5 adalah algoritma berbasis *gain ratio* maka dalam membentuk pohon keputusan diperlukan tahapan perhitungan sebagai berikut (Chauhan & Chauhan, 2013; Xiaoliang et al., 2009)

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

Kemudian konsep (ii) Gain ( $S, A$ ) merupakan perolehan informasi dari atribut  $A$  relatif terhadap *output* data  $S$ . Perolehan informasi didapat dari *output data* atau *variable dependent S* yang dikelompokkan berdasarkan atribut  $A$ , dinotasikan dengan gain ( $S, A$ ). Untuk rumus yang digunakan pada konsep pencarian nilai *gain* pada persamaan 2.

$$Gain(S, A) = \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (2)$$

dengan  $S$  adalah himpunan kasus,  $A$  adalah atribut,  $N$  adalah jumlah partisi atribut  $A$ ,  $|S|$  adalah jumlah kasus pada data ke- $i$ ,  $|S_i|$  adalah jumlah kasus dalam  $S$ . Untuk menghitung *gain ratio* perlu diketahui suatu term baru yang disebut *Split Information* dengan persamaan 3 dan 4.

$$Split\ Information(S, A) = - \sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (3)$$

$$Gain\ ratio(S, A) = - \frac{Gain(S, A)}{Split\ Information(S, A)} \quad (4)$$

Kemudian untuk evaluasi dan analisis hasil perhitungan dilakukan dengan menghitung nilai *Confusion Matrix* dengan menggunakan persamaan 5.

$$Accuracy = \left( \frac{TP + TN}{TP + TN + FP + FN} \right) \times 100\% \quad (5)$$

$$Precision = \left( \frac{TP}{TP + FP} \right) \times 100\%$$

$$Recall = \left( \frac{TP}{TP + FN} \right) \times 100\%$$

dengan  $TP$  merupakan jumlah data *True Positive*,  $TN$  merupakan jumlah data *True Negative*,  $FP$  merupakan jumlah data *False Positive*,  $FN$  merupakan jumlah data *False Negative*.

Data primer yang diperoleh memiliki 82 atribut, namun pada *modelling* penelitian atribut yang digunakan melalui tahapan data *selection*. Adapun atribut yang digunakan setelah melalui tahap *selection* yaitu:

- a. Nama Peserta: Nama peserta yang menjadi calon mahasiswa Ubhara
- b. Nama Sekolah Asal: Nama sekolah asal calon mahasiswa Ubhara
- c. Alamat Sekolah: Alamat sekolah asal calon mahasiswa Ubhara
- d. Jarak: Jarak lokasi Ubhara ke Sekolah asal calon mahasiswa Ubhara
- e. Keterangan Prioritas: Keterangan yang dilihat dari jarak lokasi Ubhara ke Sekolah asal calon mahasiswa UBHARA.

Tahap terakhir pada penelitian adalah tahap analisis hasil dari hasil pengujian. Hasil dari pengujian dengan algoritma C4.5 untuk mengetahui sasaran *marketing* calon mahasiswa baru pada setiap tahunnya melalui pemetaan lokasi yang menjadi pola dan target *marketing* UBHARA menggunakan algoritma C4.5 dalam mengetahui klasifikasi pemetaan yang telah ditentukan melalui skala Tidak Prioritas, Prioritas dan Sangat Prioritas sehingga UBHARA akan memberikan promosi kepada sekolah yang memiliki keterangan sangat prioritas yang sudah dilakukan, kemudian hasil tersebut dianalisis menggunakan pendekatan atau teknik analisis data deskriptif.

### 3. Hasil dan Pembahasan

Dalam perhitungan Algoritma C4.5 dipaparkan perhitungan yang berhubungan dengan pencarian nilai *Entropy* dan *gain* pada dataset penerimaan mahasiswa baru sebanyak 50 data dengan 3 atribut. Berikut pemaparan perhitungan pada model C4.5 pada penelitian.

#### 3.1. Persiapan Data

Persiapan data pada tahapan ini bertujuan untuk menyiapkan data yang akan digunakan dalam perhitungan model C4.5. Adapun tujuannya yaitu untuk melihat atribut serta data yang ada sebelum melakukan pengolahan *dataset*. Pada perhitungan C4.5, data yang digunakan sejumlah 50 data dan menggunakan 5 atribut yaitu nama peserta, nama kecamatan sekolah, nama sekolah, jarak dan prioritas. Untuk mendapatkan jarak, penulis menentukan jarak antar Nama Sekolah ke Ubhara melalui nama kecamatan sekolahnya. Selanjutnya, akan ditentukan skala prioritasnya. Pada tahapan ini, atribut jarak tidak digunakan karna telah digunakan dalam proses penentuan skala prioritas. Adapun Tabel 1 merupakan data yang digunakan pada penelitian.

Tabel 1. Data Penelitian

No	Nama Peserta	Kecamatan Sekolah	Nama Sekolah	Prioritas
1.	Megan Lituhayu	Bekasi Utara	SMA MUTIARA 17 AGUSTUS	Sangat Prioritas
2.	Nur Aini Dwi Rahmawati	Bekasi Utara	SMA MUTIARA 17 AGUSTUS	Tidak Prioritas
3.	Sela Dwi Putri	Bekasi Utara	SMA MUTIARA 17 AGUSTUS	Tidak Prioritas
4.	Raymond Ramadany Kalla Mellim	Bekasi Utara	SMA MUTIARA 17 AGUSTUS	Prioritas
5.	Royhan Naharuddin	Bekasi Utara	SMA MUTIARA 17 AGUSTUS	Sangat Prioritas
⋮	⋮	⋮	⋮	⋮
50.	Restu Rafly Anggara	Bekasi Selatan	SMAN 8 BEKASI	Prioritas

Sumber: Hasil Penelitian (2022)

Pada proses normalisasi data pada sub bab ini, dataset yang digunakan pada perhitungan manual dan model akan dilakukan transformasi data. Untuk mendapatkan label data pada atribut prioritas, data yang digunakan sebagai acuan ialah data sekolah dengan mencari jarak Sekolah

Asal calon mahasiswa baru dengan Universitas Bhayangkara (UBHARA). Adapun tabel normalisasi tiap-tiap atribut dapat dilihat pada Tabel 2.

Tabel 2. Normalisasi Data

Atribut	Nilai Bobot	Kategori
Jarak Sekolah	<= 50 Km	Tidak Prioritas
	51 Km – 250 Km	Prioritas
	251 Km – 450 Km	Sangat Prioritas
Nama Sekolah	SMA MUTIARA 17 AGUSTUS	0
	SMK TARUNA BANGSA	1
	SMA MANDALAHAYU BEKASI	2
	SMA MUHAMMADIYAH 09 BEKASI	3
	SMAN 2 KOTA BEKASI	4
	SMAN 3 BEKASI	5
	SMAN 8 BEKASI	6
Kecamatan Sekolah	Kec. Bekasi Selatan	0
	Kec. Bekasi Timur	1
	Kec. Bekasi Utara	2

Sumber: Hasil Penelitian (2022)

### 3.2. Perhitungan Jumlah Kasus

Setelah dilakukan normalisasi, dilakukan perhitungan jumlah kasus. Pada tahapan ini, perhitungan jumlah kasus dilakukan terhadap dua atribut melalui skala prioritas. Dua atribut tersebut ialah Nama Kecamatan Sekolah serta Nama Sekolah.

Tabel 3. Jumlah Kasus

Atribut	Kelas	Jumlah Kasus	Tidak Prioritas	Prioritas	Sangat Prioritas
Kecamatan Sekolah	Kec. Bekasi Utara	15	3	3	9
	Kec. Bekasi Timur	17	2	9	6
	Kec. Bekasi Selatan	18	3	8	7
Jumlah Kasus		50	8	20	22
Nama Sekolah	SMA MUTIARA 17 AGUSTUS	7	2	1	3
	SMK TARUNA BANGSA	8	1	2	5
	SMK MANDALAHAYU	9	1	5	3
	SMA MUHAMMADIYAH 09 BEKASI	8	1	4	3
	SMAN 2 KOTA BEKASI	6	2	3	1
	SMAN 3 BEKASI	2	1	0	1
	SMAN 8 BEKASI	10	0	5	5
Jumlah Kasus		50	8	20	22

Sumber: Hasil Penelitian (2022)

### 3.3. Perhitungan Algoritma C4.5

Dari data hasil transformasi selanjutnya dianalisis untuk menghasilkan sebuah pohon keputusan dengan menggunakan Algoritma C4.5, yaitu dengan melakukan formulasi perhitungan *entropy* dan *gain* yang dapat dilihat pada Tabel 4.

Tabel 4. Perhitungan Algoritma C4.5

Atribut	Kelas	Jumlah Kasus	Tidak Prioritas	Prioritas	Sangat Prioritas	Entropy	Information Gain
Kecamatan	Kec. Bekasi Utara	15	3	3	9	1,4	1,413
Sekolah	Kec. Bekasi Timur	17	2	9	6	1,4	
	Kec. Bekasi Selatan	18	3	8	7	1,5	
Nama Sekolah	SMA MUTIARA 17 AGUSTUS	7	2	1	3	1,441	1,052
	SMK TARUNA BANGSA	8	1	2	5	1,298	
	SMK MANDALAHAYU	9	1	5	3	1,351	
	SMA MUHAMMA-DIYAH 09 BEKASI	8	1	4	3	1,405	
	SMAN 2 KOTA BEKASI	6	2	3	1	1,459	
	SMAN 3 BEKASI	2	1	0	1	0	
	SMAN 8 BEKASI	10	0	5	5	0	

Sumber: Hasil Penelitian (2022)

Selanjutnya menghitung nilai gain dengan menggunakan persamaan 2, yaitu:

$$a. \text{Gain (Kecamatan Sekolah)} = \text{entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{entropy}(S_i)$$

$$\begin{aligned} \text{Gain (Kecamatan Sekolah)} &= 1,47293504 - \left( \left( \frac{15}{50} \right) \times 1,4 + \left( \frac{17}{50} \right) \times 1,4 + \left( \frac{18}{50} \right) \times 1,5 \right) \\ &= 2,063 \end{aligned}$$

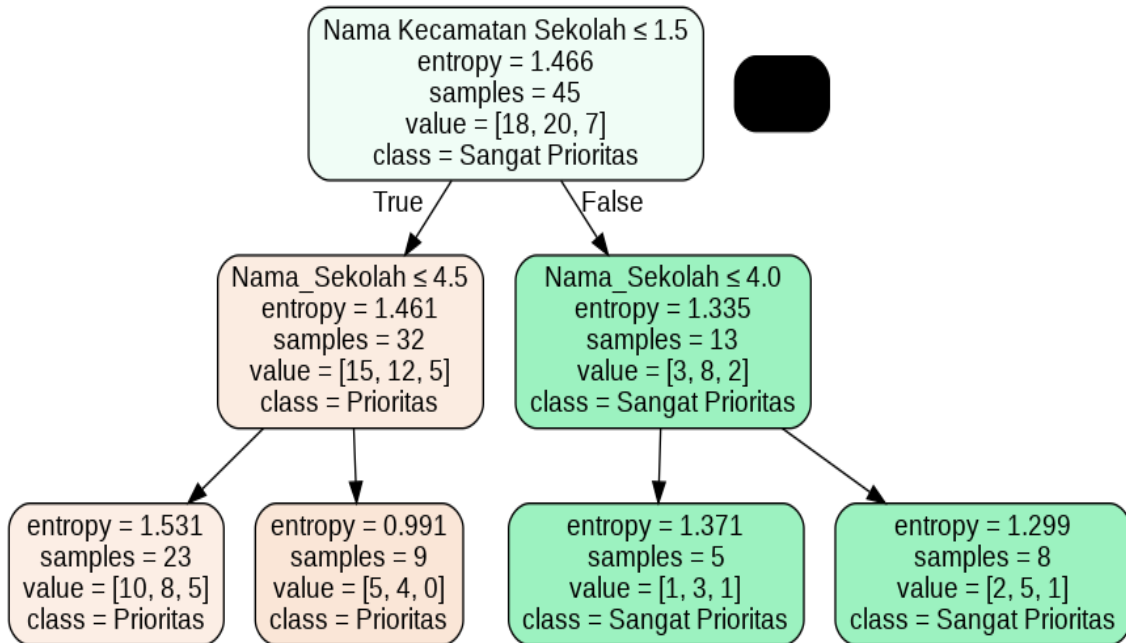
$$b. \text{Gain (Nama Sekolah)} = \text{entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{entropy}(S_i)$$

$$\begin{aligned} \text{Gain (Nama Sekolah)} &= 1,47293504 - \left( \left( \frac{7}{50} \right) \times 1,441 + \left( \frac{8}{50} \right) \times 1,298 + \right. \\ &\quad \left. \left( \frac{9}{50} \right) \times 1,351 + \left( \frac{8}{50} \right) \times 1,405 + \left( \frac{6}{50} \right) \times 1,459 + \right. \\ &\quad \left. \left( \frac{2}{50} \right) \times 0 + \left( \frac{10}{50} \right) \times 0 \right) = 0,42 \end{aligned}$$

Dari data hasil perhitungan pada Tabel 4 diperoleh nilai Gain terbesar terbesar pada atribut Kecamatan Sekolah yaitu sebesar 2,063. Maka untuk node awal yang terpilih ialah



Kecamatan Sekolah. *Output* pohon keputusan dari algoritma C4.5 menggunakan pemrograman Python ditampilkan oleh Gambar 2. *Output* ini bertujuan untuk memperlihatkan kelas dan nilai dataset yang terdapat di algoritma C4.5 pada sebuah *tree*.



Sumber: Hasil Penelitian (2022)

Gambar 2. *Output* Pohon Keputusan dari Algoritma C4.5

### 3.4. Perhitungan Nilai Akurasi

Pada tahapan ini, setelah model dilakukan perhitungan, maka akan dihitung akurasi yang dihasilkan dari 50 data sampel yang digunakan. Terdapat beberapa hal yang diketahui dari proses running program menggunakan Python yaitu Nilai *True Positif* berturut-turut yaitu 15,8,0. Maka dapat dihitung akurasi data uji dari model menggunakan persamaan 3.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Akurasi = \frac{2 + 1 + 0}{5}$$

$$Akurasi = \frac{3}{5}$$

$$Akurasi = 0,6$$

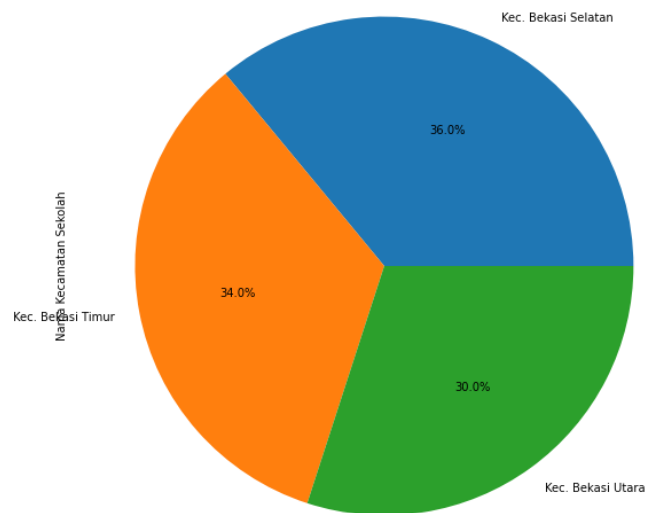
$$Akurasi = 60\%$$

Nilai *entropy* dan *information gain* merupakan fitur pada algoritma C4.5 untuk memberikan informasi mana yang menjadi atribut utama dalam menetapkan atribut yang akan menjadi pedoman dalam aktivitas *marketing* yang akan dilakukan. Pada penelitian ini diperoleh nilai *entropy* dan *information gain* terbesar terdapat pada atribut kecamatan nama sekolah. Jadi kecamatan nama sekolah merupakan atribut yang pertama kali dilihat dalam proses atau aktivitas *marketing* yang akan dilakukan. Sementara nilai akurasi merupakan keakuratan model dalam

memilih atribut yang menjadi atribut utama serta keakuratan model dalam mengolah dataset yang digunakan.

### 3.2. Visualisasi Data

Visualisasi data berbentuk diagram *pie* ini disertakan dengan tujuan untuk memudahkan pembaca dalam memperoleh informasi penting dari hasil perhitungan dan memahami data tersebut. Hasil visualisasi dari *dataset* hasil penelitian menampilkan nama kecamatan asal sekolah yang paling meminati mahasiswa Ubhara pada tahun 2019 dan tahun 2020 disajikan pada gambar 3. Infografis ini dapat digunakan juga sebagai bahan pertimbangan dan rekomendasi dalam menentukan sasaran *marketing* calon mahasiswa UBHARA Jaya pada periode tahun ajaran berikutnya.



Sumber: Hasil Penelitian (2022)

Gambar 3. Visualisasi Data Hasil Perhitungan untuk Rekomendasi Daerah Sasaran Marketing Calon Mahasiswa Baru UBHARA untuk Periode Tahun Ajaran Selanjutnya.

### 4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan oleh penulis serta pengujian model Algoritma *C4.5* dalam pemetaan sasaran marketing calon mahasiswa baru di Universitas Bhayangkara Jakarta Raya (UBHARA Jaya), didapatkan kesimpulan bahwa telah dapat dilakukan perhitungan prediksi pemetaan sasaran marketing calon mahasiswa baru di UBHARA dengan menggunakan algoritma *C4.5* dan bahasa pemrograman *Python*. Dari evaluasi menggunakan *confusion matrix* menggunakan 50 data sampel penerimaan calon mahasiswa baru, tingkat akurasi yang diberikan oleh model dengan menggunakan dataset penerimaan mahasiswa baru yaitu sebesar 60%. Serta akurasi dalam perhitungan manual menggunakan model *C4.5* diperoleh hasil 60%. Dari hasil visualisasi, adapun 50 data sampel nama kecamatan asal sekolah yang paling meminati Ubhara pada data penerimaan calon mahasiswa baru sekolah

tahun 2019 dan tahun 2020 ialah Kecamatan Bekasi Selatan. Serta pemetaan Universitas Bhayangkara (UBHARA) terhadap daerah yang menjadi target marketingnya yaitu pada skala sangat prioritas berada pada Kecamatan Bekasi Utara.

### **Ucapan Terima Kasih**

Ucapan terima kasih kami sampaikan kepada Biro Humas dan Pemasaran Universitas Bhayangkara Jakarta Raya, yang telah memfasilitasi penelitian ini untuk mendapatkan data-data riil yang dibutuhkan.

### **Daftar Pustaka**

- Budiman, I., Irwan, & Ramadina, R. (2015). Penerapan Fungsi Data Mining Klasifikasi untuk Prediksi Masa Studi Mahasiswa Tepat Waktu pada Sistem Informasi Akademik Perguruan Tinggi. *JUPITER (Jurnal Penelitian Ilmu dan Teknik Komputer)*, 7(1), 39–50.
- Chauhan, H., & Chauhan, A. (2013). Implementation of Decision Tree Algorithm c4.5. *International Journal of Scientific and Research Publications*, 3(10). [www.ijsrp.org](http://www.ijsrp.org)
- Hafiz, A. (2020). Penerapan Sistem Pendukung Keputusan Pengangkatan Karyawan Tetap Dengan Menggunakan Metode Topsis. *Jurnal Teknologi dan Informatika (JEDA)*, 1(2), 1–11.
- Kusrini, & Taufiq Emha, L. (2009). Algoritma Data Mining Yogyakarta. In *Algoritma Data Mining: Vol. February*. Andi Offset. <https://books.google.co.id/books?id=-Ojclag73O8C&printsec=frontcover#v=onepage&q&f=false>
- Nofriansyah, D. (2015). Konsep Data Mining Vs Sistem Pendukung Keputusan. In *Deepublish*. Deepublish. [https://www.google.co.id/books/edition/Konsep\\_Data\\_Mining\\_Vs\\_Sistem\\_Pendukung\\_K/PoJyCAAQBAJ?hl=id&gbpv=1&dq=algoritma+c4.5&printsec=frontcover](https://www.google.co.id/books/edition/Konsep_Data_Mining_Vs_Sistem_Pendukung_K/PoJyCAAQBAJ?hl=id&gbpv=1&dq=algoritma+c4.5&printsec=frontcover)
- Rahmadani, E. V., Pane, S. F., & Harani, N. H. (2020). Algoritma C4.5 dan K-Nearest Neighbors (KNN) untuk Memetakan Matakuliah dan Keterlambatan Kelulusan Mahasiswa. Kreatif. [https://www.google.co.id/books/edition/Algoritma\\_C4\\_5\\_dan\\_K\\_Nearest\\_Neighbors\\_K/BGv9DwAAQBAJ?hl=en&gbpv=0](https://www.google.co.id/books/edition/Algoritma_C4_5_dan_K_Nearest_Neighbors_K/BGv9DwAAQBAJ?hl=en&gbpv=0)
- Rufiyanto, A., Rochcham, M., & Rohman, A. (2020). *Penerapan Algoritma C4.5 Untuk Prediksi Kepuasan Mahasiswa Tahun 2020* (A. D. Nabila & A. Y. Wati (eds.); 1st ed.). Deepublish. [https://www.google.co.id/books/edition/Penerapan\\_Algoritma\\_C4\\_5\\_Untuk\\_Prediksi/\\_bkgEAAAQBAJ?hl=id&gbpv=1](https://www.google.co.id/books/edition/Penerapan_Algoritma_C4_5_Untuk_Prediksi/_bkgEAAAQBAJ?hl=id&gbpv=1)
- Sari, R., Fatoni, H., & Ramdhania, K. F. (2023). Decision Support System Design for Informatics Student Final Projects Using C4.5 Algorithm. *PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic*, 11(1), 123–134. <https://doi.org/10.33558/piksel.v11i1.5954>
- Sari, R., Subarkah, Setiawati, S., & Fitri, D. A. (2022). Sistem Pendukung Keputusan Penerima Bantuan Langsung Tunai (BLT) Menggunakan Metode Simple Additive Weighting (SAW) Berbasis Website. *Jurnal Peradaban Sains, Rekayasa Dan Teknologi (RADIAL)*, 10(2), 292–308. <https://doi.org/10.37971/radial.v10i2.296>

Xiaoliang, Z., Hongcan, Y., Jian, W., & Shangzhuo, W. (2009). *Research and Application of the Improved Algorithm C4.5 on Decision Tree*. <https://doi.org/10.1109/ICTM.2009.5413078>