

# Penerapan Data Mining Menggunakan Metode Algoritma Naive Bayes Classifier untuk Mendukung Strategi Promosi

Aulia Yahyah Sari<sup>1</sup>, Encep Supriatna<sup>2</sup>

<sup>1,2</sup>Sistem Informasi, Universitas Ma'soem, Indonesia  
auliyahyasari20@gmail.com

---

*Received : Nov' 2023 Revised : Nov' 2023 Accepted : Dec' 2023 Published : Dec' 2023*

---

## ABSTRACT

The purpose of this study is to make it easier to find out the prediction results with the classification method from PPDB data for 2021-2022 to support promotion strategies at the Skye Foundation with the Naive Bayes Classifier Algorithm and to find out the level of accuracy in classification in determining promotion strategies. In this study the authors used 210 data as training data and 53 data as testing data for manual testing in excel and for testing using RapidMiner tools. Based on the results of the analysis in the research conducted, the results of predicting registration and non-registration quickly and accurately from tests carried out by comparing training data with data testing using RapidMiner tools obtained an accuracy rate of 92.45%.

**Keywords:** Classification; Naives Bayes Classifier; Prediction; Promotion Strategy; RapidMiner.

## ABSTRAK

Adapun tujuan dari penelitian ini adalah untuk memudahkan mengetahui hasil prediksi dengan metode klasifikasi dari data PPDB tahun 2021-2022 untuk mendukung strategi promosi di Yayasan Skye dengan Algoritma Naive Bayes Classifier dan untuk mengetahui tingkat akurasi pada klasifikasi dalam menentukan strategi promosi. Pada penelitian ini menggunakan 210 data sebagai data *training* dan 53 data sebagai data *testing* untuk pengujian manual di excel maupun pada pengujian menggunakan *tools RapidMiner*. Berdasarkan hasil analisa pada penelitian yang dilakukan, didapatkan hasil prediksi registrasi dan tidak registrasi dengan cepat dan akurat dari pengujian yang dilakukan dengan membandingkan data training dengan data testing menggunakan *tools RapidMiner* didapat tingkat akurasi sebesar 92,45%.

**Kata Kunci :** Klasifikasi; *Naives Bayes Classifier*; Prediksi; *RapidMiner*; Strategi Promosi.

---

## PENDAHULUAN

Banyak perusahaan dan instansi pemerintah yang menggunakan data mining untuk menggali informasi. Selain perusahaan dan instansi lainnya, penerapan data mining juga dilakukan dalam dunia pendidikan seperti halnya dalam sekolah negeri maupun swasta [1]. Penggunaan teknik data *mining* diharapkan dapat membantu mempercepat proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola informasi yang terkandung didalam data transaksi menjadi sebuah pengetahuan (*knowledge*) yang baru [2].

Dalam dunia pendidikan baik sekolah negeri maupun swasta Penerimaan Peserta Didik Baru (PPDB) merupakan kegiatan ujung tombak sekolah dalam

mendapatkan siswa [3]. Dalam hal perencanaan kegiatan promosi guna peningkatan minat calon siswa untuk mendaftar, Unit Yayasan Skye membutuhkan masukan-masukan guna meningkatkan minat lulusan sekolah yang ada di wilayah Rancaekek dan luar wilayah Rancaekek untuk menjadi Siswa di Yayasan Skye. Penilaian kinerja pegawai merupakan suatu upaya untuk mengukur tingkat produktivitas pegawai. Kriteria yang sangat mempengaruhi penilaian kinerja antara lain: masa kerja, usia, jenis kelamin, divisi, status, pendidikan, pekerjaan, perilaku kerja, aspek kepribadian, aspek teknis pekerjaan, aspek non teknis, dan pencapaian tujuan. Standar penilaian yang diberlakukan serta sikap pegawai dan atasan saat menghadapi penilaian.

Dalam peranannya yang sangat penting maka strategi promosi perlu dikelola secara baik. Strategi promosi perlu dikelola upaya untuk meminimalisir biaya operasional yang keluar dan membantu dalam meningkatkan jumlah siswa mendaftar pada setiap tahunnya dengan menggali atau mencari informasi dari data pendaftaran tahun-tahun sebelumnya sebagai bahan evaluasi dan analisa untuk mendukung strategi promosi tahun berikutnya.

Yayasan Skye Media Cipta Teknologi yang terdiri dari SMP-SMK ini merupakan salah satu sekolah yang baru didirikan pada tahun 2020 berlokasi di Jalan Walini No.24, Bojongloa, Kecamatan Rancaekek, Kabupaten Bandung, Jawa Barat 40394. Yayasan ini baru memiliki dua angkatan dimana sekolah swasta ini juga menghadapi persaingan yang cukup ketat. Berdasarkan data perolehan PPDB di Yayasan Skye Media Cipta Teknologi pada dua tahun terakhir yaitu:

**Tabel 1. Time Student Capacity**

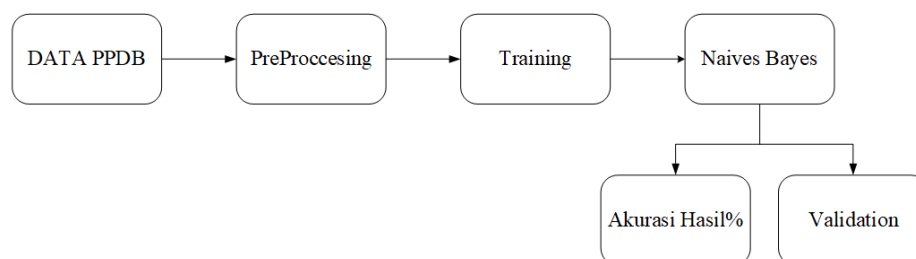
	Tahun	Pendaftar	Terdaftar	Kapasitas
SMP	2021	32	23	96
	2022	80	72	96
SMK	2021	65	52	96
	2022	86	77	96

Data PPDB dihimpun berdasarkan tahun angkatan yang jumlahnya bisa mencapai ratusan record data dengan atribut yang beragam. Jika diamati, tumpukan data ini bisa menjadi aset berharga dengan cara diolah menggunakan teknik yang tepat, guna menemukan informasi atau pengetahuan berharga yang bisa dimanfaatkan untuk menentukan strategi promosi sekolah. Indikator variabel dalam penelitian ini antara lain: Alamat, Asal sekolah, Gelombang, Informasi PPDB dan Pendapatan Orang Tua.

Kegiatan promosi yang sudah dilakukan menghasilkan calon pendaftar yang masih rendah dilihat dari data terdaftar sebelumnya bahwa jumlah siswa yang terdaftar belum mendekati daya tampung atau kapasitas yang sekolah sediakan. Maka dari itu, perlu dilakukan pengklasifikasian data berdasarkan data PPDB untuk mengetahui strategi promosi seperti apa yang perlu ditingkatkan, serta belum mengetahui tingkat akurasi dalam menentukan strategi promosi yang akan digunakan. Penelitian ini bertujuan klasifikasi untuk rekomendasi strategi promosi sesuai kategori dengan menggunakan metode *Naïves Bayes Classifier*.

## METODE

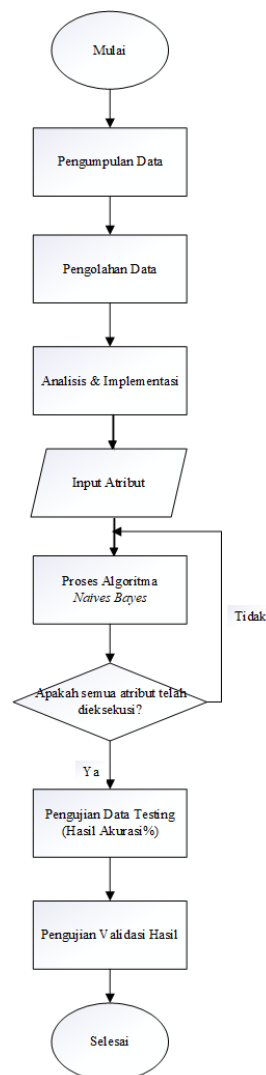
Metode yang digunakan dalam penelitian ini yaitu *Naïve Bayes Classifier*. *Naïve Bayes Classifier* ialah salah satu pengklasifikasi statistik, dimana pengklasifikasi ini dapat memprediksi probabilitas keanggotaan kelas suatu data *tuple* yang akan masuk ke dalam kelas tertentu, sesuai dengan perhitungan probabilitas [4]. Metode data mining pada pengklasifikasian data PPDB dilakukan dengan beberapa tahapan. Adapun alur proses pengklasifikasian dapat dilihat dibawah ini:



**Gambar 1. Alur Proses Klasifikasi *Naïve Bayes***

Adapun alur *data mining* menggunakan metode *Naïve Bayes Classifier* sebagai berikut:

1. **Data PPDB**  
Data PPDB merupakan data-data yang dimiliki oleh calon siswa yang telah mendaftar. Ada beberapa atribut yang dimiliki oleh data PPDB ini yaitu, Alamat, Asal sekolah, Gelombang, Informasi PPDB dan Pendapatan Orang Tua.
2. ***Pre-Proccesing***  
Pada tahap *pre-processing* bertujuan untuk membersihkan data mentah menjadi data siap diolah. Pada tahap *pre-processing* bertujuan untuk menghapus data yang terindikasi memiliki duplikasi data menghilangkan *missing value* pada data [5]. Sehingga pada saat diproses tidak muncul kesalahan atau masalah. *Preprocessing* selanjutnya melakukan penghapusan atribut yang tidak digunakan pada tahap selanjutnya. Sehingga atribut yang tersisa hanya atribut yang nantinya digunakan pada tahap modeling.
3. ***Naïves Bayes Classifier***  
Pada tahapan ini, dilakukan proses pengklasifikasian data *training* menggunakan *tools RapidMiner studio* terhadap data PPDB yang sudah di testing menggunakan model *Classifier* yang dibentuk.
4. **Akurasi Hasil**  
Akurasi hasil menunjukkan kedekatan hasil pengukuran dengan nilai sesungguhnya.
5. ***Validation***  
*Cross Validation* merupakan salah satu teknik untuk menilai/memvalidasi keakuratan sebuah model yang dibangun berdasarkan dataset tertentu [6]. Salah satunya metode *Cross Validation* yang populer adalah *K-fold Cross Validation*.



Gambar 2. Prosedur Kerja

Adapun alur data mining menggunakan metode *Naïves Bayes Classifier* sebagai berikut:

1. Pengumpulan Data

Data yang digunakan dalam penelitian menggunakan data yang diambil langsung dari objek penelitian. Data yang digunakan dalam penelitian ini adalah data dari Yayasan Skye sebanyak 263 data.

2. Mengolah Data

Data yang telah diperoleh akan diolah dengan memilah data atau membersihkan data jika terdapat *missing value*. *Missing value* adalah data yang tidak lengkap dikarenakan atribut tidak tercatat maupun atribut memang tidak dimiliki [7]. Lalu penghapusan atribut yang tidak akan digunakan.

### 3. Analisis dan Implementasi

Setelah melakukan semua tahap pemrosesan data, maka akan dihasilkan data *training*. Data *training* merupakan data yang digunakan untuk melakukan proses perhitungan menggunakan teknik data mining klasifikasi dengan menggunakan metode Algoritma *Naives Bayes* dengan cara perhitungan probabilitas dari setiap atribut dan melakukan proses pengujian dari hasil yang telah diperoleh [8].

### 4. Pengujian Validasi Hasil

Penelitian ini membutuhkan tes validasi agar model yang dibuat merupakan model yang terbaik dengan tingkat akurasi yang tinggi dan maksimal [9]. Hasil pengujian akan di validasi menggunakan software *RapidMiner*. *RapidMiner* digunakan untuk membandingkan hasil akurasi yang didapat pada *confusion matrix* [10]. Lalu didapat hasil tingkat akurasi dari model yang telah digunakan dalam penerapan implementasi dalam mendukung strategi promosi di Yayasan Skye Media Cipta Teknologi Rancaekek.

## HASIL DAN PEMBAHASAN

### Analisis Data

Analisis dilakukan berdasarkan data PPDB dari Yayasan Skye Digipreneur tahun ajaran 2021-2022 dengan total sebanyak 263 data.

**Tabel 2. Atribut Penelitian**

No	Indikator	Penjelasan
1	Alamat	Wilayah menjadi acuan dalam melakukan promosi agar akses pendidikan dapat tersebar ke daerah-daerah tertentu.
2	Asal Sekolah	Asal sekolah indentiknya terdapat dari negeri maupun swasta.
3	Gelombang	Gelombang merupakan jalur pendaftaran bagi calon siswa.
4	Informasi PPDB	Informasi PPDB menjadi media promosi yang dapat berpengaruh sebagian besar untuk keberhasilan dalam menarik minat calon pendaftar dari beberapa media yang digunakan.
5	Pendapatan Orang Tua	Pendapatan merupakan aspek pendukung dalam penerimaan calon siswa baru.

Atribut-atribut diatas akan menghasilkan atribut atau kolom baru yakni class prediction, apakah kriteria dari setiap atribut menghasilkan registrasi atau tidak registrasi. Berikut dataset dari data PPDB dari Yayasan Skye dibawah ini merupakan sampel sebanyak 26 data.

No	Responden	Daftar Skye	Alamat	Asal Sekolah	Pendapatan Orang Tua	Gelombang PPDB	Informasi PPDB	Tahun PPDB	Status
1	Siswa 1	SMK	Jl. Taman anggrek no.20 blok	Negeri	0 - 4.000.000	Gelombang 1	Teman	2021	Registrasi
2	Siswa 2	SMK	Grand niscorn rancaekek blok	Negeri	0 - 4.000.000	Gelombang 2	Kunjungan ke Sekolah	2021	Registrasi
3	Siswa 3	SMK	Buah dua Rancaekek	Negeri	> 8.000.000	Gelombang 3	Instagram	2021	Registrasi
4	Siswa 4	SMK	JL MELUR IX NO 12 RT 07 RA	Negeri	0 - 4.000.000	Gelombang 1	Spanduk	2021	Registrasi
5	Siswa 5	SMK	Kp. Babakan Asem Rt 01/03	Negeri	0 - 4.000.000	Gelombang 1	Instagram	2021	Registrasi
6	Siswa 6	SMK	Perumahan pesona Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Spanduk	2021	Registrasi
7	Siswa 7	SMK	Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Spanduk	2021	Registrasi
8	Siswa 8	SMK	Perum abdi negara blok N5 N	Negeri	0 - 4.000.000	Gelombang 2	Brosur	2021	Registrasi
9	Siswa 9	SMK	Jl. Melur IV No. 103 Rancaek	Negeri	> 8.000.000	Gelombang 1	Website	2021	Registrasi
10	Siswa 10	SMK	Jln cempaka 1 no 50 blok 6 RT	Negeri	0 - 4.000.000	Gelombang 1	Teman	2021	Registrasi
11	Siswa 11	SMK	Jln. Kamboja 1 No. 68 RT. 01	Negeri	5.000.000 - 8.000.000	Gelombang 1	Brosur	2021	Registrasi
12	Siswa 12	SMK	Pesona Rancaekek indah ked	Negeri	0 - 4.000.000	Gelombang 1	Brosur	2021	Registrasi
13	Siswa 13	SMK	jl.raya cikalang cileunyi no 82	Negeri	0 - 4.000.000	Gelombang 1	Brosur	2021	Registrasi
14	Siswa 14	SMK	Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Teman	2021	Registrasi
15	Siswa 15	SMK	Jln yasaadi rt 03 nr 11 desa s	Negeri	0 - 4.000.000	Gelombang 1	Teman	2021	Registrasi
16	Siswa 16	SMP	Griya utama kec Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Teman	2021	Registrasi
17	Siswa 17	SMK	Jatinangor	Swasta	0 - 4.000.000	Gelombang 2	Teman	2021	Tidak Registrasi
18	Siswa 18	SMK	Kp. sukaraja rt 4 nr 8 kecama	Negeri	0 - 4.000.000	Gelombang 2	Brosur	2021	Registrasi
19	Siswa 19	SMK	Jl. pesona ayu 1. kec. Ranca	Negeri	0 - 4.000.000	Gelombang 3	Teman	2021	Registrasi
20	Siswa 20	SMK	Komp GRITARA Jln Palembang	Negeri	5.000.000 - 8.000.000	Gelombang 2	Teman	2021	Registrasi
21	Siswa 21	SMK	Jalan anggrek bulan no 19 pe	Swasta	5.000.000 - 8.000.000	Gelombang 2	Spanduk	2021	Registrasi
22	Siswa 22	SMK	rancaekek kencana blok 7 rt 1	Negeri	0 - 4.000.000	Gelombang 1	Teman	2021	Registrasi
23	Siswa 23	SMK	Pangalengan	Negeri	0 - 4.000.000	Gelombang 3	Instagram	2022	Registrasi
24	Siswa 24	SMK	Rancakendal dua	Swasta	5.000.000 - 8.000.000	Gelombang 1	Instagram	2022	Registrasi
25	Siswa 25	SMK	Jatinangor kampung geluis	Swasta	> 8.000.000	Gelombang 1	Brosur	2021	Registrasi
26	Siswa 26	SMK	Kp. Pintu Rancaekek RT. 02 I	Negeri	> 8.000.000	Gelombang 1	Brosur	2022	Registrasi

Gambar 3. Data PPDB 2021-2022

### Perhitungan Naives Bayes

Tahapan-tahapan untuk melakukan penerapan Algoritma *Naives Bayes* adalah sebagai berikut:

#### 1. Cleaning Data

Pembersihan data bertujuan untuk menghapus data yang terindikasi memiliki duplikasi data menghilangkan *missing value* pada data. Sehingga pada saat diproses tidak muncul kesalahan atau masalah. Tahap ini menghapus data yang duplikat. Awalnya data berjumlah 263 baris serta 9 kolom selanjutnya dilakukan penghapusan duplikasi yang ternyata hasilnya tidak berubah sehingga dapat terlihat bahwa data tersebut tidak memiliki duplikasi data. Begitu pula bahwa setelah data dicek tidak memiliki *missing value*.

#### 2. Penghapusan Atribut

Penghapusan atribut yang tidak digunakan sehingga, atribut yang tersisa hanya atribut yang nantinya digunakan pada tahap modeling. Berikut data setelah dilakukan penghapusan atribut sebagai berikut :

No	Alamat	Asal Sekolah	Pendapatan Orang Tua	Gelombang PPDB	Informasi PPDB	Status
1	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
2	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Kunjungan ke Sekolah	Registrasi
3	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 3	Instagram	Registrasi
4	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Spanduk	Registrasi
5	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Instagram	Registrasi
6	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Spanduk	Registrasi
7	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Spanduk	Registrasi
8	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Brosur	Registrasi
9	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Website	Registrasi
10	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
11	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Brosur	Registrasi
12	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Brosur	Registrasi
13	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Brosur	Registrasi
14	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Teman	Registrasi
15	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
16	Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Teman	Registrasi
17	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 2	Teman	Tidak Registrasi
18	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Brosur	Registrasi
19	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Teman	Registrasi
20	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 2	Teman	Registrasi
21	Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 2	Spanduk	Registrasi
22	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
23	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Instagram	Registrasi
24	Luar Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Instagram	Registrasi
25	Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 1	Brosur	Registrasi
26	Luar Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Brosur	Registrasi
27	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 1	Website	Tidak Registrasi
28	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 3	Instagram	Tidak Registrasi
29	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Kerabat	Registrasi
30	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 2	Spanduk	Registrasi

Gambar 4. Penghapusan Atribut

### 3. Data Training

Data *training* ini menggunakan data dengan jumlah 210 data, dibawah ini merupakan sampel sebanyak 30 data.

No	Alamat	Asal Sekolah	Pendapatan Orang Tua	Gelombang PPDB	Informasi PPDB	Status
1	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
2	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Kunjungan ke Sekolah	Registrasi
3	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 3	Instagram	Registrasi
4	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Spanduk	Registrasi
5	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Instagram	Registrasi
6	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Spanduk	Registrasi
7	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Spanduk	Registrasi
8	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Brosur	Registrasi
9	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Website	Registrasi
10	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
11	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Brosur	Registrasi
12	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Brosur	Registrasi
13	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Brosur	Registrasi
14	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Teman	Registrasi
15	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
16	Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Teman	Registrasi
17	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 2	Teman	Tidak Registrasi
18	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Brosur	Registrasi
19	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Teman	Registrasi
20	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 2	Teman	Registrasi
21	Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 2	Spanduk	Registrasi
22	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
23	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Instagram	Registrasi
24	Luar Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Instagram	Registrasi
25	Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 1	Brosur	Registrasi
26	Luar Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Brosur	Registrasi
27	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 1	Website	Tidak Registrasi
28	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 3	Instagram	Tidak Registrasi
29	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Kerabat	Registrasi
30	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 2	Spanduk	Registrasi

Gambar 5. Data Training

### 4. Data Testing

Data *testing* ini menggunakan 53 data, dibawah ini merupakan sampel sebanyak 35 data.

1	Alamat	Asal Sekolah	Pendapatan Orang Tua	Gelombang PPDB	Informasi PPDB	Status
2	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 2	Kunjungan ke Sekolah	Registrasi
3	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Spanduk	Registrasi
4	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Spanduk	Registrasi
5	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
6	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 1	Teman	Registrasi
7	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 2	Teman	Tidak Registrasi
8	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 2	Teman	Registrasi
9	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Instagram	Registrasi
10	Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 1	Brosur	Registrasi
11	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 3	Instagram	Tidak Registrasi
12	Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Spanduk	Registrasi
13	Luar Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 1	Instagram	Registrasi
14	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Brosur	Registrasi
15	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 2	Teman	Registrasi
16	Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 1	Spanduk	Registrasi
17	Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 1	Kunjungan ke Sekolah	Registrasi
18	Luar Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Teman	Registrasi
19	Luar Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 3	Spanduk	Registrasi
20	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 1	Instagram	Registrasi
21	Luar Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 3	Website	Tidak Registrasi
22	Wilayah Rancaekek	Negeri	5.000.000 - 8.000.000	Gelombang 2	Teman	Registrasi
23	Luar Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 3	Brosur	Tidak Registrasi
24	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Teman	Tidak Registrasi
25	Luar Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Brosur	Tidak Registrasi
26	Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 3	Teman	Tidak Registrasi
27	Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 2	Kunjungan ke Sekolah	Registrasi
28	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Teman	Registrasi
29	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Spanduk	Registrasi
30	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 1	Instagram	Registrasi
31	Wilayah Rancaekek	Swasta	0 - 4.000.000	Gelombang 3	Teman	Registrasi
32	Luar Wilayah Rancaekek	Swasta	5.000.000 - 8.000.000	Gelombang 2	Teman	Tidak Registrasi
33	Luar Wilayah Rancaekek	Swasta	> 8.000.000	Gelombang 2	Kunjungan ke Sekolah	Registrasi
34	Wilayah Rancaekek	Negeri	> 8.000.000	Gelombang 1	Teman	Registrasi
35	Wilayah Rancaekek	Negeri	0 - 4.000.000	Gelombang 3	Teman	Registrasi

Gambar 6. Data Testing

### 5. Perhitungan Probabilitas Prior (P(Ci))

Tahap pertama perhitungan untuk menentukan strategi promosi dengan metode *Naives Bayes* adalah dengan mencari probabilitas dari masing-masing kelas. Dalam penentuan strategi promosi akan ditentukan dengan 2 kelas yaitu kelas "Registrasi" dan "Tidak Registrasi". Cara perhitungannya adalah dengan mencari berapa jumlah data yang ada pada kelas "Registrasi" dan berapa jumlah pada kelas "Tidak Registrasi" dari total keseluruhan data *training*, lalu membaginya

dengan total keseluruhan data. Hasil dari perhitungan dapat dilihat pada tabel dibawah ini:

**Tabel 3. Probabilitas Kelas**

- P(Ci)
- P(Efektif) = $177/210 = 0,843$ Jumlah data "Registrasi" pada kolom "Status" dibagi jumlah data
- P(Tidak Efektif) = $33/210 = 0,157$ Jumlah data "Tidak Registrasi" pada kolom "Status" dibagi jumlah data

#### 6. Menghitung Probabilitas Masing-Masing Atribut

Cara mencari probabilitas suatu atribut adalah dengan membandingkan atribut dari data *testing* dengan atribut dari data *training*.

##### a. Alamat

**Tabel 4. Atribut Alamat**

P(Alamat)	Registrasi	Tidak Registrasi
Wilayah Rancaekek	0,71	0,18
Luar Wilayah Rancaekek	0,29	0,82
Total	1,00	1,00

##### b. Asal Sekolah

**Tabel 5. Atribut Asal Sekolah**

P(Asal Sekolah)	Registrasi	Tidak Registrasi
Negeri	0,52	0,39
Swasta	0,48	0,61
Total	1,00	1,00

##### c. Pendapatan Orang Tua

**Tabel 7. Atribut Pendapatan Orang Tua**

P(Pendapatan Orang Tua)	Registrasi	Tidak Registrasi
0 - 4.000.000	0,53	0,61
5.000.000 - 8.000.000	0,28	0,27
> 8.000.000	0,19	0,12
Total	1,00	1,00

##### d. Gelombang PPDB

**Tabel 8. Atribut Gelombang PPDB**

P(Gelombang PPDB)	Registrasi	Tidak Registrasi
Gelombang 1	0,54	0,03
Gelombang 2	0,24	0,33
Gelombang 3	0,22	0,64
Total	1,00	1,00

e. Informasi PPDB

**Tabel 9. Atribut Informasi PPDB**

P(Informasi PPDB)	Registrasi	Tidak Registrasi
Brosur	0,15	0,09
Spanduk	0,11	0,03
Website	0,01	0,18
Instagram	0,15	0,06
Teman	0,31	0,55
Kerabat	0,10	0,06
Kunjungan ke Sekolah	0,18	0,03
Total	1,00	1,00

**Perhitungan Naives Bayes Manual**

Berikut adalah hasil perhitungan *Naives Bayes* manual.

1. Hasil perhitungan (*Example Set*) sampel 7 data.

**Tabel 10. Hasil Perhitungan (*Example Set*)**

Row No.	Status	Prediction (Status)	Confidence (Registrasi)	Confidence (Tidak Registrasi)
1	Registrasi	Registrasi	0,007	0,000
2	Registrasi	Registrasi	0,009	0,000
3	Registrasi	Registrasi	0,003	0,000
4	Registrasi	Registrasi	0,027	0,000
5	Registrasi	Registrasi	0,014	0,000
6	Tidak Registrasi	Tidak Registrasi	0,004	0,008
7	Registrasi	Registrasi	0,006	0,000

2. Hasil Perhitungan *Accuracy*

Secara umum *precision*, *recall*, dan *accuracy* dapat dirumuskan sebagai berikut:

- a. Untuk kelas Registrasi

$$Precision = \frac{41}{41 + 3} = \frac{41}{44} = 0,9318 = 93,18\%$$

$$Recall = \frac{41}{41 + 1} = \frac{41}{42} = 0,9762 = 97,62\%$$

- b. Untuk kelas Tidak Registrasi

$$Precision = \frac{8}{8 + 1} = \frac{8}{9} = 0,8889 = 88,89\%$$

$$Recall = \frac{8}{8 + 3} = \frac{8}{11} = 0,7273 = 72,73\%$$

$$Accuracy = \frac{8 + 41}{8 + 3 + 1 + 41} = \frac{49}{53} = 0,9245 = 92,45\%$$

$$Error = \frac{3+1}{8+3+1+41} = \frac{4}{53} = 0,0755 = 7,55\%$$

### 3. Presentase Registrasi dan Tidak Registrasi

Presentase Registrasi dan Tidak Registrasi yang didapat dari hasil uji 263 data dengan Naives Bayes.

**Tabel 11. Presentase**

	Registrasi	Tidak Registrasi
<i>Presentase</i>	83,6%	16,4%

## PENUTUP

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa dari proses klasifikasi berdasarkan alamat, asal sekolah, gelombang, informasi PPDB dan pendapatan orang tua didapat hasil untuk strategi promosi yaitu: alamat wilayah rancaekek dengan jumlah pendaftar 71%, gelombang 1 menjadi peminat paling banyak dengan jumlah 54% dan informasi PPDB yang berpengaruh sebagian besar yaitu dari teman 30% dan kunjungan ke sekolah 18%. Berdasarkan hasil analisa pada penelitian yang dilakukan, didapatkan hasil prediksi dalam menentukan calon siswa registrasi atau tidak registrasi dengan cepat dan akurat, dari pengujian yang dilakukan dengan membandingkan data training dengan data testing menggunakan *tools RapidMiner* didapat dengan tingkat akurasi sebesar 92,45%.

Adapun saran bagi Yayasan Skye, diharapkan dalam menentukan strategi promosi dengan melihat atribut informasi PPDB yaitu dari teman, dengan memberikan *fee* rekomendasi kepada siswa yang mampu merekomendasikan orang lain untuk daftar ke Yayasan Skye dan kunjungan ke sekolah dalam peranannya dapat memperhatikan dari segi pemilihan materi yang akan disampaikan seperti lebih menonjolkan keunggulan atau pembeda dari sekolah lain, memperhatikan cara bersosialisasi agar informasi dapat diterima dengan baik, lalu pemilihan personal dari segi penampilan dan usia dapat diperhatikan agar orang lain dapat lebih tertarik untuk mendaftar ke Yayasan Skye. Lalu, menambahkan variabel alasan memilih Skye dalam form PPDB, diharapkan masukan-masukan tersebut dapat meningkatkan kualitas pendidikan di Yayasan Skye Media Cipta Teknologi Rancaekek.

## DAFTAR PUSTAKA

- [1] A. H. Mirza, "Application of Naive Bayes Classifier Algorithm in Determining New Student Admission Promotion Strategies," *J. Inf. Syst. Informatics*, vol. 1, no. 1, pp. 14-28, 2019, doi: 10.33557/journalisi.v1i1.2.
- [2] Dwiarti, R. Utami, D. Rahma Utami, A. Kurniawan, and U. Enri, "Naive Bayes Klasifikasi untuk Rekomendasi Strategi Promosi Tindakan pada Aesthetic

- 
- Dental Clinic Karawang," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 601-610, 2021.
- [3] A. H. Sani, A. Setiawan, and ..., "Penerapan Metode Naive Bayes Dalam Rekomendasi Strategi Penerimaan Peserta Didik Baru," *J. Comput. ...*, vol. 4, no. 1, pp. 245-251, 2022, doi: 10.47065/josyc.v4i1.2438.
- [4] Diana, *Metode dan Aplikasi Sistem Pendukung Keputusan*. Yogyakarta: CV BUDI UTAMA, 2018.
- [5] Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi pada BPR Pantura," *J. Penelit. Tek. Inform.*, pp. 128-146, 2019, [Online]. Available: <https://repository.nusamandiri.ac.id/index.php/repo/viewitem/13890>.
- [6] L. Mardiana, D. Kusnandar, and N. Satyahadewi, "Analisis Diskriminan Dengan K Fold Cross Validation Untuk Klasifikasi Kualitas Air Di Kota Pontianak," *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 11, no. 1, pp. 97-102, 2022.
- [7] A. R. Alfarisi, H. Tjandrasa, and I. Arieshanti, "Perbandingan Performa antara Imputasi Metode Konvensional dan Imputasi dengan Algoritma," *J. Tek. Pomits*, vol. 2, no. 1, pp. 1-4, 2013.
- [8] Dwi Untari, K. Hastuti, E. Y. Hidayat, Dwi Untari, N. Limão, and N. Y. L. Gaol, "Data Mining untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5," *J. Penelit. Ilmu Komput.*, vol. 2, no. November, pp. 31-48, 2018.
- [9] S. Dewi and Oktaviawati, "Penerapan Algoritma C4.5 untuk Pehamanan Siswa SMK Pada Pelajaran Kompetensi Keahlian," *Intern. (Information Syst. Journal)*, vol. 5, no. 2, pp. 116-125, 2022.
- [10] M. F. Nugraha and S. B. Rahayu, "Penerapan Naïve Bayes dalam Mengklasifikasi Calon Penerima Bantuan Pangan Non Tunai di Desa Nanjung Mekar," ... *Syst. Journal*), vol. 5, no. 2, pp. 137-146, 2022, [Online]. Available: <https://jurnal.masoemiversity.ac.id/index.php/internal/article/view/634><https://jurnal.masoemiversity.ac.id/index.php/internal/article/download/634/423>.
-