

Penerapan Naïve Bayes pada Potensi Akademik Siswa SD Negeri 5 Singakerta

Ni Kadek Winda Patrianingsih ^{(1)*}, I Kadek Arya Sugianta ⁽²⁾

Informatika, Fakultas Bisnis, Teknologi, Sosial dan Humaniora, Universitas Bali Internasional, Bali

e-mail : {windapatrianingsih,aryasugianta}@iikmpbali.ac.id.

* Penulis korespondensi.

Artikel ini diajukan 28 Maret 2023, direvisi 12 April 2023, diterima 13 April 2023, dan dipublikasikan 26 Mei 2023.

Abstract

Student potential cannot only be measured based on the result of academic scores, and many things influence student academic determination. The purpose of this research is to prove that students' potential is influenced by many things, such as character, academic activity, socioeconomic status, and distance of residence. By using the naïve Bayes method and testing with the confusion matrix, it will give results for this research. The data is from V-grade students at SD Negeri 5 Singakerta, with 120 students assisted by the homeroom teacher. Based on the results of the tests that have been carried out using a data sample of 10 students and 1 data using the Naïve Bayes, it is obtained that students have academic potential, and the results with the confusion matrix are accuracy of 75%, precision of 81%, and recall of 89%. In this case, it can be concluded that the academic potential of students can not only be measured based on the results of the final grade, but many other factors have an effect, the application of the Naïve Bayes in students' academic potential is appropriate to use.

Keywords: *Student Potential, Student Academic, Data Mining, Naïve Bayes, Confusion Matrix*

Abstrak

Potensi siswa tidak hanya dapat diukur berdasarkan hasil nilai akademik, banyak hal yang berpengaruh dalam penentuan akademik siswa. Tujuan penelitian ini adalah untuk membuktikan bahwa potensi siswa dipengaruhi oleh banyak hal seperti, karakter, aktivitas akademik, status sosial ekonomi, dan jarak tempat tinggal. Dengan menggunakan algoritma Naïve Bayes dan pengujian dengan *confusion matrix*, akan memberikan hasil terhadap penelitian ini. Data yang digunakan yaitu data siswa kelas V di SD Negeri 5 Singakerta dengan perolehan data melalui peserta didik yang didampingi wali kelas sebanyak 120 siswa. Berdasarkan hasil perhitungan manual dengan Naïve Bayes memberikan keputusan terhadap data *testing* dengan hasil siswa tersebut memiliki potensi akademik. Pengujian dengan *confusion matrix* memberikan hasil yaitu *accuracy* sebesar 75%, *precision* sebesar 81%, dan *recall* sebesar 89%. Dalam hal ini dapat disimpulkan bahwa potensi akademik siswa tidak hanya dapat diukur berdasarkan hasil nilai akhir, namun banyak faktor lain yang berpengaruh, oleh karena itu penerapan algoritma Naïve Bayes dalam potensi akademik siswa layak digunakan.

Kata Kunci: *Potensi Siswa, Akademik Siswa, Data Mining, Naïve Bayes, Confusion Matrix*

1. PENDAHULUAN

Pendidikan memiliki peranan penting dalam kemajuan bangsa. Ilmu pengetahuan merupakan salah satu gambaran dari penggunaan dan pemanfaatan kemajuan teknologi komputer, yang terdiri dari berbagai bagian dari ilmu. Setiap individu memiliki atribut karakter yang berbeda-beda. Karakter yang disebut merupakan suatu ciri khas yang terdapat di dalam diri sendiri dalam keadaan yang berbeda (Martantoh & Yanih, 2022). Sekolah merupakan suatu jembatan bagi siswa di dalam menimba ilmu sebagai bekal siswa tersebut di masa depan. Faktor yang mempengaruhi pencapaian dari siswa adalah guru yang mampu memberikan pengajaran yang efektif dan berkualitas. Perkembangan proses pembelajaran merupakan suatu unsur dasar sebagai bentuk keberhasilan yang di mana setiap siswa memiliki tingkat kecepatan dalam



memahami materi dan nantinya akan menjadi perhatian khusus bagi sekolah untuk menentukan tingkat keberhasilan (Triandini et al., 2021). Perolehan apresiasi siswa akan ditampung dalam sebuah rapat yang di mana proses pembelajaran siswa telah dikumpulkan selama enam bulan atau satu semester. Hasil dari pembelajaran siswa tersebut akan diukur dan menghasilkan sebuah penilaian potensi dari masing-masing siswa (Susilo & Nur'aini, 2018).

SD Negeri 5 Singkerta merupakan sekolah dasar yang berada di Banjar Tunon, Kecamatan Ubud, Kabupaten Gianyar. Terdapat 120 siswa kelas lima pada semester ini. Perkembangan dunia digital memiliki berbagai manfaat dalam pemenuhan kebutuhan manusia yang salah satunya yaitu perkembangan *data science*. *Data science* memiliki fungsi signifikan dalam pengolahan data (Saleh, 2015b). Sistem penilaian dan pencatatan pada SD Negeri 5 Singkerta masih bersifat manual sehingga para wali kelas mempunyai kendala dalam proses penentuan potensi akademik siswa. Potensi siswa di SD Negeri 5 Singkerta masih dinilai dari hasil nilai akhir masing-masing siswa saja, sedangkan masih ada aspek penilaian yang dapat mendukung dari potensi siswa tersebut. Sistem penilaian tersebut tentu saja belum cukup untuk kemajuan potensi akademik siswa di SD Negeri 5 Singkerta. Salah satu alternatif yang dapat dilakukan untuk membantu pihak wali kelas dalam proses penilaian akhir yaitu dengan memanfaatkan perkembangan teknologi yang ada, khususnya *data mining*.

Dalam penelitian ini teknik *data mining* yang akan diterapkan yaitu algoritma Naïve Bayes. Pemilihan algoritma ini berdasarkan suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang didefinisikan. Naïve Bayes adalah algoritma yang sederhana dan tiap-tiap atribut bersifat bebas, yang memungkinkan tiap atribut dapat berkontribusi terhadap hasil akhir (Marpaung et al., 2021). Penggunaan *data mining* khususnya Naïve Bayes diharapkan dapat memberikan informasi yang dapat digunakan untuk menentukan potensi akademik siswa di SD Negeri 5 Singkerta, sehingga kemampuan siswa dapat dilihat tidak hanya dari nilai akhir saja, namun dari segi ekonomi maupun jarak tempat tinggal dapat mempengaruhi.

Klasifikasi algoritma Naïve Bayes dalam *data mining* untuk menentukan konsentrasi siswa menghasilkan bahwa metode Naïve Bayes berhasil mengklasifikasikan 109 data siswa dari 120 data yang diuji dengan persentase keakuratan 90,8333% (Saleh, 2015a). Selanjutnya, dalam penelitian (Putri, 2020) yaitu penerapan model Naïve Bayes untuk memprediksi potensi pendaftaran siswa di SMK Taman Siswa Teluk Betung berbasis *web* memperoleh hasil akhir sebuah sistem informasi yang memiliki kemampuan untuk memprediksi potensi banyaknya siswa yang mendaftar dan tidak mendaftar dengan menggunakan aplikasi yang menghasilkan perhitungan akurasi sebesar 86%. Implementasi algoritma naïve bayes untuk klasifikasi penerima beasiswa (studi kasus universitas hamzanwadi) memberikan kesimpulan bahwa algoritma naïve bayes sangat baik digunakan dalam seleksi kelulusan pemberkasan beasiswa bidikmisi (Nurhidayati et al., 2023).

Dalam penelitian sebelumnya yaitu penentuan topik tugas akhir dengan metode Naïve Bayes, membuktikan bahwa metode ini menunjukkan hasil yang sangat layak berdasarkan kuisioner yang disebar ke dua responden yaitu dosen dan ahli sistem oleh (Ghaniy & Sihotang, 2019). Prediksi potensi mahasiswa berprestasi dengan metode Naïve Bayes di Stikom Dinamika Bangsa Jambi memberikan hasil persentase akurasi sebesar 87% oleh (Resti et al., 2021).

Selanjutnya, dalam penerapan metode Naïve Bayes Classifier untuk mengklasifikasi tingkat prestasi akademik santri Pondok Pesantren Mahasiswa (PPM) Baitul Jannah Malang (Y. M. Firdaus, 2019), penelitian tersebut dibuatkan sebuah sistem untuk mengklasifikasi tingkat prestasi akademik agar memudahkan para guru dalam proses penentuan prestasi akademik dan aplikasi diuji dengan menghasilkan keakuratan dengan nilai 76%.

Kelebihan dari penerapan Naïve Bayes yang dijelaskan dalam penelitian sebelumnya yaitu Naïve Bayes mampu memberikan hasil keakuratan dalam melakukan prediksi prestasi siswa, sehingga penelitian ini diambil berdasarkan penelitian sebelumnya bahwa penggunaan algoritma Naïve Bayes akan mampu menangani kumpulan data siswa yang terpilih yaitu siswa kelas lima di SD



Negeri 5 Singakerta untuk memberikan hasil yang akurat dan tepat. Selain itu faktor pendukung dari penggunaan naïve bayes dinilai bahwa berpotensi baik dalam melakukan klasifikasi suatu data dan memberikan hasil akurasi dan efesiensi yang optimal (Kawani, 2019).

Mengacu pada penjelasan latar belakang di atas, maka pentingnya dilakukan pengukuran hasil potensi akademik siswa di SD Negeri 5 Singakerta, yang kedepannya akan membantu pihak sekolah dalam peningkatan mutu pendidikan serta memberikan kriteria-kriteria yang tidak biasa dijumpai dalam melakukan penentuan potensi dengan hasil secara akurat dan tepat.

2. METODE PENELITIAN

2.1 Naïve Bayes

Kategorisasi probabilitas berdasarkan Teorema Bayes disebut Naïve Bayes. Untuk membuat perhitungan yang diperlukan lebih mudah, Naïve Bayes memperhitungkan dampak dari nilai atribut lainnya. Rumus probabilitas Bayesian yang luas dapat dilihat pada Pers. (1) (Sarasvananda et al., 2022).

$$\frac{P(C_i|X)=P(X|C_i)P(C_i)}{P(X)} \quad (1)$$

Berdasarkan rumus probabilitas Naïve Bayes dijelaskan bahwa $P(C_i|X)$ merupakan probabilitas C_i jika diberi bukti X , selanjutnya $P(C_i)$ merupakan probabilitas C_i tanpa memandang bukti, $P(X|C_i)$ yaitu probabilitas X terjadi akan mempengaruhi C_i , dan $P(X)$ adalah probabilitas X tanpa memandang bukti apapun.

2.2 Confusion Matrix

Confusion matrix merupakan suatu cara untuk mengukur performa dari kinerja klasifikasi yang di mana mendapatkan keluaran dua kelas atau lebih. *Confusion matrix* digambarkan berupa tabel dengan empat kombinasi berbeda dari nilai aktual dan nilai prediksi, dari hasil tersebut dimasukkan ke dalam tabel yaitu *true positif*, *true negatif*, *false positif* dan *false negatif*. Dengan menggunakan perhitungan ini akan menghasilkan seberapa tingkat *accuracy*, *precision*, dan *recall*. Adapun persamaan dalam melakukan pengujian dengan *confusion matrix* terdapat pada Tabel 1 (A. F. Firdaus et al., 2021).

Tabel 1 *Confusion Matrix*

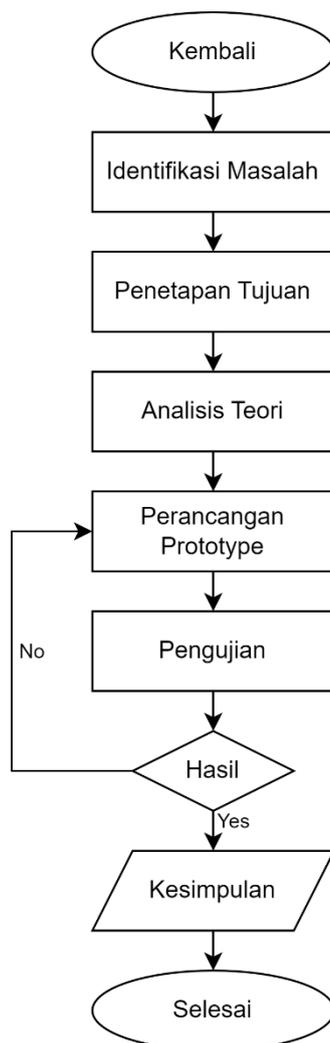
| Nilai Aktual | Nilai Prediksi | |
|--------------|---------------------|---------------------|
| | Positif | Negatif |
| Positif | True Positive (TP) | False Negative (FN) |
| Negatif | False Positive (FP) | True Negative (TN) |

Berdasarkan Tabel 1 dijelaskan bahwa *true positif* merupakan data positif yang diprediksi benar, *true negatif* merupakan data negatif yang diprediksi benar, *false positif* merupakan data negatif namun diprediksi sebagai data positif, dan nilai prediksi *false negatif* merupakan data positif namun diprediksi sebagai data negatif.

2.3 Alur Penelitian

Dalam penelitian ini data yang ditambah adalah data siswa kelas V di SD Negeri 5 Singakerta. Penelitian ini difokuskan untuk mengetahui keakuratan dari penerapan metode Naïve Bayes dalam menentukan potensi akademik siswa di SD Negeri 5 Singakerta. Adapun rancangan alur penelitian digambarkan pada Gambar 1.





Gambar 1 Alur Penelitian

Berdasarkan rancangan alur pada Gambar 1, proses penelitian ini akan diawali dengan identifikasi masalah. Dalam penelitian ini, ditemukan permasalahan yang terjadi di SD Negeri 5 Singakerta yaitu dalam penentuan potensi akademik siswa, selanjutnya dilakukan penetapan tujuan sehingga mendapatkan hasil sesuai dengan perumusan masalah. Proses selanjutnya yaitu analisis teori, kemudian perancangan *prototype*, dan dilakukan pengujian yang menggunakan *confusion matrix*. Pengujian berhasil dan memperoleh hasil yang akurat, jika hasil yang diperoleh akurat dan baik dari proses pengujian bisa dilanjutkan ke tahap kesimpulan, jika tidak maka pengujian harus diulang kembali sampai menemukan hasil yang baik dan akurat, dan proses penelitian bisa selesai.

2.4 Pengumpulan Data

Data yang digunakan bersumber dari SD Negeri 5 Singakerta, yang di mana pengisian data tersebut diisi oleh peserta didik yang didampingi oleh wali kelas, kemudian data digunakan sebagai penunjang di data rapot siswa. Data yang diambil untuk pemilihan atribut adalah data siswa kelas V SD Negeri 5 Singakerta yang berjumlah 120 siswa dengan rincian sebagai pada Gambar 2.



| Kode | Atribut | | | | Kode | Atribut | | | | Kode | Atribut | | | |
|------|---------|-------------|--------|--------|------|---------|-------------|--------|--------|------|---------|-------------|--------|--------|
| | A1 | A2 | A3 | A4 | | A1 | A2 | A3 | A4 | | A1 | A2 | A3 | A4 |
| V01 | A | Aktif | Sedang | Tinggi | V21 | A | Tidak Aktif | Sedang | Rendah | V41 | B | Aktif | Dekat | Tinggi |
| V02 | A | Aktif | Sedang | Sedang | V22 | B | Tidak Aktif | Sedang | Sedang | V42 | B | Aktif | Dekat | Sedang |
| V03 | A | Sedang | Sedang | Rendah | V23 | B | Tidak Aktif | Dekat | Tinggi | V43 | B | Aktif | Sedang | Sedang |
| V04 | B | Aktif | Sedang | Sedang | V24 | B | Aktif | Dekat | Sedang | V44 | B | Aktif | Sedang | Tinggi |
| V05 | B | Aktif | Dekat | Tinggi | V25 | A | Aktif | Sedang | Sedang | V45 | B | Aktif | Dekat | Sedang |
| V06 | C | Aktif | Dekat | Sedang | V26 | C | Tidak Aktif | Dekat | Tinggi | V46 | B | Aktif | Dekat | Tinggi |
| V07 | C | Tidak Aktif | Sedang | Sedang | V27 | A | Aktif | Dekat | Sedang | V47 | B | Sedang | Sedang | Tinggi |
| V08 | A | Tidak Aktif | Dekat | Tinggi | V28 | A | Tidak Aktif | Jauh | Tinggi | V48 | B | Aktif | Dekat | Sedang |
| V09 | B | Sedang | Dekat | Sedang | V29 | B | Tidak Aktif | Dekat | Tinggi | V49 | B | Aktif | Sedang | Tinggi |
| V10 | C | Aktif | Jauh | Tinggi | V30 | B | Aktif | Sedang | Sedang | V50 | B | Aktif | Sedang | Sedang |
| V11 | B | Sedang | Dekat | Tinggi | V31 | B | Aktif | Sedang | Sedang | V51 | B | Aktif | Dekat | Tinggi |
| V12 | C | Tidak Aktif | Dekat | Tinggi | V32 | B | Tidak Aktif | Sedang | Tinggi | V52 | B | Aktif | Aktif | Tinggi |
| V13 | C | Tidak Aktif | Dekat | Sedang | V33 | C | Tidak Aktif | Dekat | Tinggi | V53 | B | Aktif | Dekat | Sedang |
| V14 | C | Aktif | Sedang | Sedang | V34 | B | Aktif | Dekat | Sedang | V54 | B | Aktif | Dekat | Rendah |
| V15 | B | Aktif | Dekat | Tinggi | V35 | B | Aktif | Jauh | Tinggi | V55 | B | Tidak Aktif | Jauh | Sedang |
| V16 | C | Tidak Aktif | Dekat | Sedang | V36 | B | Aktif | Dekat | Tinggi | V56 | B | Aktif | Dekat | Tinggi |
| V17 | B | Tidak Aktif | Jauh | Tinggi | V37 | B | Sedang | Sedang | Tinggi | V57 | B | Aktif | Sedang | Sedang |
| V18 | C | Tidak Aktif | Dekat | Tinggi | V38 | B | Aktif | Sedang | Sedang | V58 | A | Aktif | Dekat | Tinggi |
| V19 | A | Tidak Aktif | Sedang | Tinggi | V39 | B | Aktif | Sedang | Rendah | V59 | C | Tidak Aktif | Dekat | Tinggi |
| V20 | B | Aktif | Sedang | Sedang | V40 | B | Aktif | Sedang | Sedang | V60 | B | Aktif | Sedang | Sedang |

| Kode | Atribut | | | | Kode | Atribut | | | | Kode | Atribut | | | |
|------|---------|-------------|--------|--------|------|---------|-------------|--------|--------|------|---------|-------------|--------|--------|
| | A1 | A2 | A3 | A4 | | A1 | A2 | A3 | A4 | | A1 | A2 | A3 | A4 |
| V61 | B | Aktif | Dekat | Rendah | V81 | B | Aktif | Sedang | Tinggi | V101 | A | Aktif | Dekat | Sedang |
| V62 | B | Aktif | Dekat | Sedang | V82 | B | Aktif | Dekat | Sedang | V102 | C | Tidak Aktif | Dekat | Tinggi |
| V63 | B | Aktif | Jauh | Tinggi | V83 | B | Aktif | Dekat | Tinggi | V103 | A | Aktif | Dekat | Sedang |
| V64 | B | Aktif | Dekat | Sedang | V84 | B | Aktif | Sedang | Tinggi | V104 | B | Aktif | Dekat | Tinggi |
| V65 | C | Tidak Aktif | Sedang | Sedang | V85 | B | Aktif | Dekat | Sedang | V105 | B | Aktif | Dekat | Sedang |
| V66 | B | Aktif | Sedang | Tinggi | V86 | B | Aktif | Sedang | Tinggi | V106 | A | Aktif | Sedang | Sedang |
| V67 | B | Aktif | Sedang | Sedang | V87 | B | Aktif | Sedang | Sedang | V107 | A | Aktif | Dekat | Sedang |
| V68 | B | Aktif | Sedang | Rendah | V88 | B | Tidak Aktif | Sedang | Rendah | V108 | A | Aktif | Jauh | Tinggi |
| V69 | B | Aktif | Sedang | Sedang | V89 | B | Aktif | Sedang | Sedang | V109 | B | Aktif | Dekat | Tinggi |
| V70 | C | Tidak Aktif | Dekat | Tinggi | V90 | B | Aktif | Dekat | Tinggi | V110 | A | Aktif | Sedang | Sedang |
| V71 | B | Aktif | Dekat | Sedang | V91 | B | Aktif | Dekat | Sedang | V111 | A | Aktif | Dekat | Tinggi |
| V72 | C | Tidak Aktif | Sedang | Sedang | V92 | A | Tidak Aktif | Sedang | Sedang | V112 | A | Aktif | Dekat | Sedang |
| V73 | B | Aktif | Dekat | Tinggi | V93 | B | Aktif | Dekat | Tinggi | V113 | B | Aktif | Sedang | Sedang |
| V74 | C | Aktif | Dekat | Sedang | V94 | B | Aktif | Dekat | Sedang | V114 | B | Aktif | Sedang | Tinggi |
| V75 | B | Aktif | Jauh | Tinggi | V95 | B | Aktif | Jauh | Tinggi | V115 | B | Aktif | Dekat | Sedang |
| V76 | B | Tidak Aktif | Dekat | Tinggi | V96 | B | Aktif | Dekat | Tinggi | V116 | B | Aktif | Dekat | Tinggi |
| V77 | B | Aktif | Sedang | Sedang | V97 | B | Aktif | Sedang | Sedang | V117 | B | Aktif | Sedang | Tinggi |
| V78 | B | Aktif | Dekat | Tinggi | V98 | B | Tidak Aktif | Dekat | Sedang | V118 | B | Aktif | Dekat | Sedang |
| V79 | A | Aktif | Dekat | Sedang | V99 | B | Aktif | Sedang | Sedang | V119 | A | Aktif | Sedang | Tinggi |
| V80 | B | Aktif | Sedang | Sedang | V100 | B | Tidak Aktif | Sedang | Tinggi | V120 | A | Aktif | Sedang | Sedang |

Gambar 2 Atribut Data Siswa

2.5 Variabel Penelitian

Atribut penilaian mencakup antara lain yaitu karakter, aktivitas akademik, tempat tinggal, dan status sosial ekonomi. Uraian atribut dalam variabel penelitian ini adalah:

- 1) Atribut Karakter
 Karakter siswa yang diukur dalam kategori sangat baik (A), baik (B), cukup (C), kurang baik (D), dan buruk (E).
- 2) Atribut Aktivitas Akademik
 Aktivitas siswa dalam bidang akademik terdiri dari aktif dan tidak aktif.
- 3) Atribut Tempat Tinggal
 Jarak dari rumah ke sekolah diukur berdasarkan kategori dekat, sedang, dan jauh.
- 4) Atribut Sosial Ekonomi
 Atribut status sosial ekonomi diukur berdasarkan kategori tinggi, sedang, dan rendah.

3. HASIL DAN PEMBAHASAN

Metode yang akan digunakan untuk memprediksi potensi akademik siswa untuk mempermudah penilaian oleh wali kelas sehingga dapat berjalan dengan lancar dan metodis serta mencapai tujuan yang telah ditetapkan. Langkah-Langkah berikut termasuk dalam tahapan penelitian ini yaitu:

3.1 Analisis Data

Berdasarkan data yang telah terkumpulkan hasil dari implementasi Naïve Bayes dalam menentukan potensi akademik siswa akan dilakukan perhitungan secara manual dari data



training dan data *testing* yang telah ditentukan. Selanjutnya pengujian dengan *confusion matrix* akan dilakukan menggunakan keseluruhan data yang diberikan dari pihak sekolah. Hasil prediksi dalam Naïve Bayes akan membantu pihak sekolah tentunya para wali kelas dalam melakukan pengelolaan nilai dan memberikan keputusan terhadap siswa yang memiliki potensi akademik dan tidak memiliki potensi akademik berdasarkan kriteria yang sudah ditetapkan dalam penelitian ini.

3.2 Kriteria dan Sub Kriteria

Analisis potensi akademik siswa dengan menggunakan Naïve Bayes membutuhkan kriteria dan subkriteria dalam proses perhitungannya. Subkriteria yang ada juga perlu didefinisikan rentang nilai bobotnya. Kriteria dan subkriteria dapat dilihat pada Tabel 2 (Yendarman, 2016).

Tabel 2 Kriteria dan Sub Kriteria

| Kriteria | Subkriteria | Nilai |
|-----------------------|-----------------|-------|
| Karakter | A (Sangat Baik) | 5 |
| | B (Baik) | 4 |
| | C (Cukup) | 3 |
| | D (Kurang Baik) | 2 |
| | E (Buruk) | 1 |
| Aktivitas Akademik | Aktif | 5 |
| | Sedang | 3 |
| | Tidak Aktif | 1 |
| Tempat Tinggal | Dekat | 5 |
| | Sedang | 3 |
| | Jauh | 1 |
| Status Sosial Ekonomi | Tinggi | 5 |
| | Sedang | 3 |
| | Rendah | 1 |

3.3 Hasil Perhitungan Naïve Bayes

Berdasarkan hasil pengumpulan data yang telah dilakukan, jumlah populasi data dalam penelitian ini yaitu 120 orang. Pada proses perhitungannya data akan dibagi menjadi 2, yaitu data *training* yang menggunakan data dengan kode V01 sampai V10 dan data *testing* menggunakan data dengan kode V11. Adapun simulasi disajikan seperti pada Tabel 3.

Tabel 3 Data Training

| Kode | ATRIBUT | | | | Potensi |
|------|---------|-------------|--------|--------|---------|
| | A1 | A2 | A3 | A4 | |
| V01 | A | Aktif | Sedang | Tinggi | Ya |
| V02 | A | Aktif | Sedang | Sedang | Ya |
| V03 | A | Sedang | Sedang | Rendah | Ya |
| V04 | B | Aktif | Sedang | Sedang | Ya |
| V05 | B | Aktif | Dekat | Tinggi | Ya |
| V06 | C | Aktif | Dekat | Sedang | Tidak |
| V07 | C | Tidak Aktif | Sedang | Sedang | Tidak |
| V08 | A | Tidak Aktif | Dekat | Tinggi | Tidak |
| V09 | B | Sedang | Dekat | Sedang | Ya |
| V10 | C | Aktif | Jauh | Tinggi | Ya |

Tabel 3 merupakan tabel data *training* dari algoritma Naïve Bayes. Selanjutnya diperlukan data *testing* untuk dapat menghasilkan keputusan terkait data siswa, data *testing* tersebut seperti pada Tabel 4.



Tabel 4 Data Testing

| Kode | ATRIBUT | | | | Potensi |
|------|---------|--------|-------|--------|---------|
| | A1 | A2 | A3 | A4 | ? |
| V11 | B | Sedang | Dekat | Tinggi | |

Berikut merupakan tahapan penyelesaian data uji menggunakan perhitungan Naïve Bayes.

1) Tahap 1, menghitung jumlah kelas/label

- $P(Y = Ya) = 7/10$
 Jumlah data "Ya" pada kolom "Potensi" dibagi jumlah data.
- $P(Y = Tidak) = 3/10$
 Jumlah data "Tidak" pada kolom "Potensi" dibagi jumlah data.

2) Tahap 2, menghitung jumlah kasus yang sama dengan kelas yang sama

- $P(Karakter = A|Y = Ya) = 3/7$
- $P(Karakter = A|Y = Tidak) = 1/3$
- $P(Karakter = B|Y = Ya) = 3/7$
- $P(Karakter = B|Y = Tidak) = 0/3$
- $P(Karakter = C|Y = Ya) = 1/7$
- $P(Karakter = C|Y = Tidak) = 2/3$

- $P(AktivitasAkademik = Aktif|Y = Ya) = 5/7$
- $P(AktivitasAkademik = Aktif|Y = Tidak) = 1/3$
- $P(AktivitasAkademik = Sedang|Y = Ya) = 2/7$
- $P(AktivitasAkademik = Sedang|Y = Tidak) = 0/3$
- $P(AktivitasAkademik = TidakAktif|Y = Ya) = 0/7$
- $P(AktivitasAkademik = TidakAktif|Y = Tidak) = 2/3$

- $P(TempatTinggal = Dekat|Y = Ya) = 2/7$
- $P(TempatTinggal = Dekat|Y = Tidak) = 2/3$
- $P(TempatTinggal = Sedang|Y = Ya) = 4/7$
- $P(TempatTinggal = Sedang|Y = Tidak) = 1/3$
- $P(TempatTinggal = Jauh|Y = Ya) = 1/7$
- $P(TempatTinggal = Jauh|Y = Tidak) = 0/3$

- $P(StatusSosialEkonomi = Rendah|Y = Ya) = 1/7$
- $P(StatusSosialEkonomi = Rendah|Y = Tidak) = 0/3$
- $P(StatusSosialEkonomi = Sedang|Y = Ya) = 3/7$
- $P(StatusSosialEkonomi = Sedang|Y = Tidak) = 2/3$
- $P(StatusSosialEkonomi = Tinggi|Y = Ya) = 3/7$
- $P(StatusSosialEkonomi = Tinggi|Y = TIDAK) = 1/3$

3) Tahap 3, kalikan semua hasil variabel Ya dan Tidak, sesuai dengan sampel data *testing*.

| | | | | |
|-----|---|--------|-------|--------|
| V11 | B | Sedang | Dekat | Tinggi |
|-----|---|--------|-------|--------|

Selanjutnya, dilakukan proses perhitungan terhadap variabel-variabel tersebut, untuk mendapatkan nilai dari sampel data *testing*.

Ya

$$R = P(Karakter = B|Y = Ya) \times P(AktivitasAkademik = TidakAktif|Y = Ya) \times P(TempatTinggal = Dekat|Y = Ya) \times P(StatusSosialEkonomi = Sedang|Y = YA) \times P(Y = YA)$$

$$= 3/7 \times 2/7 \times 2/7 \times 3/7$$



= 0,01499375

Tidak

$$R = P(\text{Karakter} = B|Y = Ya) * P(\text{Aktivitas Akademik} = \text{Tidak Aktif}|Y = Ya) * P(\text{Tempat Tinggal} = \text{Dekat}|Y = Ya) * P(\text{Status Sosial Ekonomi} = \text{Sedang}|Y = Ya) * P(Y = Ya)$$

$$= 3/7 \times 2/7 \times 2/7 \times 3/7$$

$$= 0$$

4) Tahap 4, bandingkan hasil pada kelas Ya dan Tidak. Karena hasil (P|Ya) lebih besar dibanding (P|Tidak), maka keputusan yang dihasilkan dari data uji bahwa siswa dengan kode V11 **memiliki potensi akademik.**

3.4 Hasil Pengujian

Berdasarkan hasil perhitungan yang telah didapatkan, selanjutnya dilakukan pengujian dengan menggunakan *confusion matrix*. Pengujian ini dilakukan oleh hasil yang diberikan oleh pihak sekolah sesuai dengan penilaian per semester oleh masing-masing wali kelas dengan jumlah data sebanyak 120 siswa. Pengujian dengan *confusion matrix* ini akan menghasilkan tiga nilai yaitu *accuracy*, *precision*, dan *recall*. Tabel 5 menunjukkan hasil prediksi dari algoritma Naïve Bayes.

Tabel 5 Hasil Prediksi Naïve Bayes

| Aktual | Prediksi | |
|------------------|------------|------------------|
| | Berpotensi | Tidak Berpotensi |
| Berpotensi | 85 | 10 |
| Tidak Berpotensi | 20 | 5 |

$$Accuracy(\%) = \frac{85 + 5}{120} = 75\%$$

Kedekatan nilai prediksi dengan nilai asli berada pada tingkat 75 %.

$$Precision(\%) = \frac{85}{85 + 20} = 81\%$$

Prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif berada pada tingkat 81 %.

$$Recall = \frac{85}{85 + 10} = 89\%$$

Ukuran ketepatan dari suatu kejadian yang diinginkan berada pada tingkat 89%.

Hasil dari pengujian *confusion matrix* dengan metode Naïve Bayes yaitu memiliki nilai *accuracy* sebesar 75 %, nilai *precision* sebesar 81 %, dan nilai *recall* sebesar 89%.

4. KESIMPULAN

Sistem penilaian di SD Negeri 5 Singakerta masih bersifat subjektif dengan menggunakan nilai akhir sebagai acuan dalam menentukan potensi akademik siswa. Potensi siswa tidak dapat diukur hanya dengan satu faktor, melainkan banyak faktor pendukung yang dapat berpengaruh dalam penentuan potensi akademik siswa. Penelitian ini menjelaskan bahwa pengukuran terhadap potensi akademik siswa dapat berpengaruh dengan status sosial ekonomi, karakter, aktivitas akademik, dan juga jarak tempat tinggal. Perkembangan teknologi yang semakin maju, dapat membantu pihak sekolah dalam melakukan penyelesaian permasalahan ini, yaitu dengan teknik *data mining* dengan menggunakan algoritma Naïve Bayes. Naïve Bayes merupakan suatu algoritma yang berfungsi untuk memprediksi suatu kelas dengan melakukan pengelompokan



secara efektif. Adapun data dalam penelitian ini bersumber dari data SD Negeri 5 Singakerta. Data tersebut terkumpul sebanyak 120 data siswa, yang di mana data dengan jumlah 120 tersebut sudah diberikan oleh pihak sekolah yang digunakan untuk melakukan pengujian dengan *confusion matrix*. Pengujian ini dilakukan dengan mendapatkan hasil yaitu *accuracy* sebesar 75%, *precision* sebesar 81%, dan *recall* sebesar 89%. Hal ini membuktikan bahwa penilaian potensi akademik siswa dapat dilakukan dengan algoritma Naïve Bayes dengan menghasilkan data akurasi yang cukup baik. Hal lain yang perlu dikembangkan dalam penelitian selanjutnya yaitu dapat dilakukan implementasi dengan perangkat lunak berbasis *web*, dengan menambahkan algoritma lain dalam perhitungannya sehingga menghasilkan tingkat akurasi yang lebih baik.

DAFTAR PUSTAKA

- Firdaus, A. F., Saedudin, R., & Andeswari, R. (2021). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa. *EProceedings of Engineering*, 8(5), 9274–9279. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15825/15538>
- Firdaus, Y. M. (2019). Penerapan Metode Naïve Bayes Classifier untuk Mengklasifikasi Tingkat Prestasi Akademik Santri Pondok Pesantren Mahasiswa (PPM) Baitul Jannah Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(1), 327–336. <https://doi.org/10.36040/JATI.V3i1.1398>
- Ghaniy, R., & Sihotang, K. (2019). Penerapan Metode Naïve Bayes Classifier Untuk Penentuan Topik Tugas Akhir. *Teknois: Jurnal Ilmiah Teknologi Informasi Dan Sains*, 9(1), 63–72. <https://doi.org/10.36350/jbs.v9i1.7>
- Kawani, G. P. (2019). Implementasi Naive Bayes. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(2), 73–81. <https://doi.org/10.20895/inista.v1i2.73>
- Marpaung, S., -, S., & -, I. (2021). Penerapan Metode Naïve Bayes Dalam Memprediksi Prestasi Siswa Di SMA Negeri 1 Panombeian Panei. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 4(2), 8–13. <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v4i2.1522>
- Martantoh, E., & Yanih, N. (2022). Implementasi Metode Naïve Bayes Untuk Klasifikasi Karakteristik Kepribadian Siswa Di Sekolah MTS Darussa'adah Menggunakan Php Mysql. *Jurnal Teknologi Sistem Informasi*, 3(2), 166–175. <https://doi.org/10.35957/jtsi.v3i2.2896>
- Nurhidayati, N. I., Yahya, Y., Fathurrahman, F., Samsu, L. M., & Amnia, W. (2023). Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi). *Infotek: Jurnal Informatika Dan Teknologi*, 6(1), 177–188. <https://doi.org/10.29408/jit.v6i1.7529>
- Putri, S. eka Y. (2020). Penerapan Model Naive Bayes untuk Memprediksi Potensi Pendaftaran Siswa di SMK Taman Siswa Teluk Betung Berbasis Web. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 1(1), 93–99. <https://doi.org/10.33365/jatika.v1i1.228>
- Resti, R., Abidin, D. Z., & Rasywir, E. (2021). Penerapan Data Mining Klasifikasi Untuk Memprediksi Potensi Mahasiswa Berprestasi Di Stikom Dinamika Bangsa Jambi Dengan Metode Naive Bayes. *Jurnal Ilmiah Mahasiswa Teknik Informatika*, 3(2), 97–114. <https://ejournal.stikom-db.ac.id/index.php/jimti/article/view/1047>
- Saleh, A. (2015a). Klasifikasi Metode Naive Bayes dalam Data Mining untuk Menentukan Konsentrasi Siswa (Studi Kasus di MAS PAB 2 Medan). *KeTIK - Konferensi Nasional Pengembangan Teknologi Informasi Dan Komunikasi*, 200–208.
- Saleh, A. (2015b). Penerapan Data Mining dalam Menentukan Jurusan Siswa. *Seminar Nasional Informatika 2015*, 1(1), 351–355. <https://e-journal.potensi-utama.ac.id/ojs/index.php/SNI/article/view/291>
- Sarasvananda, I. B. G., Selivan, D., Radhitya, M. L., & Putra, I. N. T. A. (2022). Analisis Sentimen Pada Pembelajaran Daring Di Indonesia Melalui Twitter Menggunakan Naïve Bayes Classifier. *SINTECH (Science and Information Technology) Journal*, 5(2), 227–233. <https://doi.org/10.31598/sintechjournal.v5i2.1241>
- Susilo, G., & Nur'aini, T. A. (2018). Pengaruh Tes Potensi Akademik Siswa terhadap Prestasi



- Belajar Matematika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, 4(1), 21. <https://doi.org/10.30998/jkpm.v4i1.2975>
- Triandini, M., Defit, S., & Nurcahyo, G. W. (2021). Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia. *Jurnal Informasi Dan Teknologi*, 3, 167–173. <https://doi.org/10.37034/jidt.v3i3.120>
- Yendarman, Y. (2016). Peningkatan Kemampuan Menetapkan Kkm dengan Diskusi Kelompok Kecil Bagi Guru Matematika SMKN 1, SMKN 4, Smk PGRI 2 Kota Jambi. *Jurnal Ilmiah Universitas Batanghari Jambi*, 16(3), 122–129. <https://doi.org/10.33087/JIUBJ.V16I3.25>

