

IDENTIFIKASI SISWA BERMASALAH MENGGUNAKAN METODE NAÏVE BAYES DI SMK MUHAMMADIYAH 3 BANJARMASIN

Reza Fauzan ⁽¹⁾, Joni Riadi ⁽²⁾,
Frista Rizky Rinandi ⁽³⁾
reza.fauzan@poliban.ac.id ⁽¹⁾
joni_riadi@poliban.ac.id ⁽²⁾
fristabjmkalsel@gmail.com ⁽³⁾

(1), (2), (3) Teknik Informatika Politeknik Negeri Banjarmasin

Abstrak

Perilaku siswa pada saat ini begitu sangat mengkhawatirkan. Ada banyak perilaku yang bersifat destruktif, seperti tawuran, geng-gengan, minuman keras, perilaku seks bebas, pencurian, pemerkosaan, perampokan, dan banyak lagi. Kenakalan pada siswa sekolah merupakan gejala patologis sosial pada siswa yang disebabkan oleh satu bentuk pengabaian sosial. Beberapa faktor seperti keluarga, sekolah, dan teman sepermainan dianggap menjadi faktor penyebab perilaku siswa bermasalah. Untuk dapat mengidentifikasi siswa bermasalah yang terjadi dimasa sekarang, penulis mengumpulkan data secara langsung dilapangan. Dengan menggunakan metode *Naïve Bayes*, sistem pendukung keputusan siswa bermasalah dan tidak bermasalah di SMK Muhammadiyah 3 Banjar-masin telah dapat di selesaikan. Berdasarkan pengujian yang telah di lakukan, nilai presisi yang ditunjukkan sebe-sar 92,5%. Nilai recall dihasilkan sebesar 51,3%.

Kata Kunci: *Naïve Bayes*, Sistem Identifikasi, Sekolah

1. PENDAHULUAN

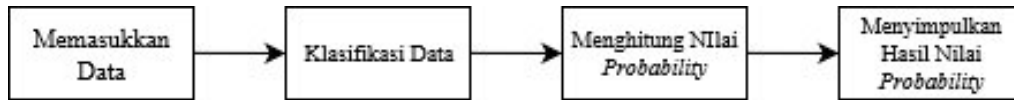
Perilaku siswa pada saat ini begitu sangat mengkhawatirkan. Ada banyak perilaku yang bersifat destruktif, seperti tawuran, geng-gengan, minuman keras, perilaku seks bebas, pencurian, pemerkosaan, perampokan, dan banyak lagi. Hal ini tentu saja memunculkan keprihatinan tersendiri dan sekaligus menjadi tamparan keras bagi praktik pendidikan di Indonesia. Perilaku siswa tersebut menjadi ukuran apakah praktik pendidikan di Indonesia sudah berhasil ataukah tidak [1].

Pada Jurnal Ilmiah Ilmu Komputer di Fakultas Ilmu Komputer Universitas AL Asyariah Mandar dan Fakultas Sains dan Teknologi Universitas Islam Negeri Makasar, *Naïve Bayes* di implementasikan untuk memprediksi kelulusan untuk memasuki perguruan tinggi. Dalam pembahasan ini terdapat 2 dataset yang digunakan, yaitu yang pertama data penerimaan mahasiswa baru yang memiliki 15 atribut dan 363 baris yang terdiri nama, tempat, tanggal lahir, asal sekolah, tahun lulus, nama orang tua, nomor ujian, nilai tes tertulis, nilai tes

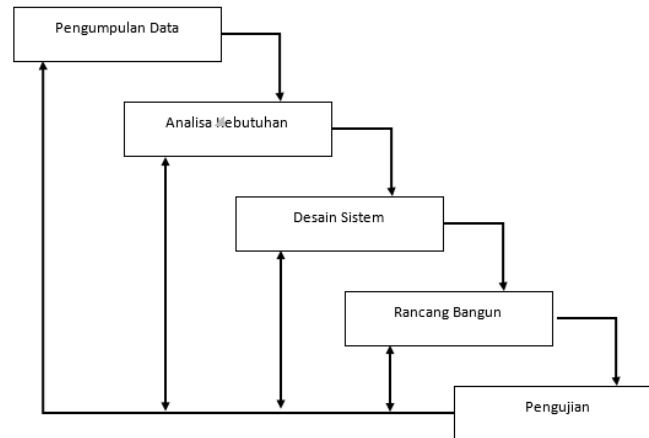
wawancara, nilai rata-rata, ket.Lulus, lulus pilihan dan sebagainya. Kedua, data hasil ujian masuk perguruan tinggi yang terdiri dari 5 atribut dan 362 baris yang terdiri dari nomor ujian, nilai tes tulisan, nilai tes wawancara, rata-rata dan keterangan lulus. [2]

Pada hasil evaluasinya menunjukkan data yang diklasifikasikan secara benar (*correct classified instances*) sesuai dengan pengelompokan pilihan lulus pilihan pertama, pilihan kedua dan tidak lulus oleh algoritma sebanyak 93.6288% atau sebanyak 338 data dan data yang di klasifikasikan namun tidak sesuai dengan class yang di prediksi (*incorrect classified instances*) yang seharusnya merupakan kelompok pilihan kedua atau tidak lulus tetapi dimasukkan kedalam kelompok pilihan pertama yaitu sebanyak 6.3712 % atau sebanyak 23 data.

Naïve Bayes dapat melakukan pengklasifikasian dengan metode probabilitas dan statistik, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. Nilai presentase keakuratan menunjukkan keefektifan dataset penerimaan Mahasiswa Baru yang diterapkan ke dalam metode *Naïve Bayes*



Gambar 1. Rancangan Umum Sistem



Gambar 2. Metode Waterfall

Classification. Implementasi *Naïve Bayes* menggunakan aplikasi WEKA dapat menelusuri karakteristik atribut dari dataset dengan luaran Pilihan Lulus. Pengelompokan Pilihan Lulus dilakukan berdasarkan atribut terpilih yaitu Prodi, Pilihan Pertama Kedua dan Nilai Rata [3].

2. METODE PENELITIAN

Sistem bekerja ketika pengguna memasukkan data. Kemudian data yang telah dimasukan di klasifikasi sesuai kategori yang telah ditentukan. Setelah mendapatkan hasil mengenai nilai *probability*. Kemudian nilai *probability* dihitung. Lalu hasil perhitungan dari nilai *probability* di simpulkan. Ilustrasi ada pada Gambar 1.

Pada Gambar 2 adalah metode pengembangan perangkat lunak menggunakan metode *waterfall* dengan tahap awal mengumpulkan data secara menyeluruh di lapangan. Kemudian data tersebut analisa dan diseleksi sesuai kriteria yang di perlukan untuk penelitian. Kemudian data yang sudah terkumpul di rancang dengan menggunakan bahasa pemrograman *java* dengan metode *Naïve Bayes*. Setelah itu hasil pada sistem diujikan.

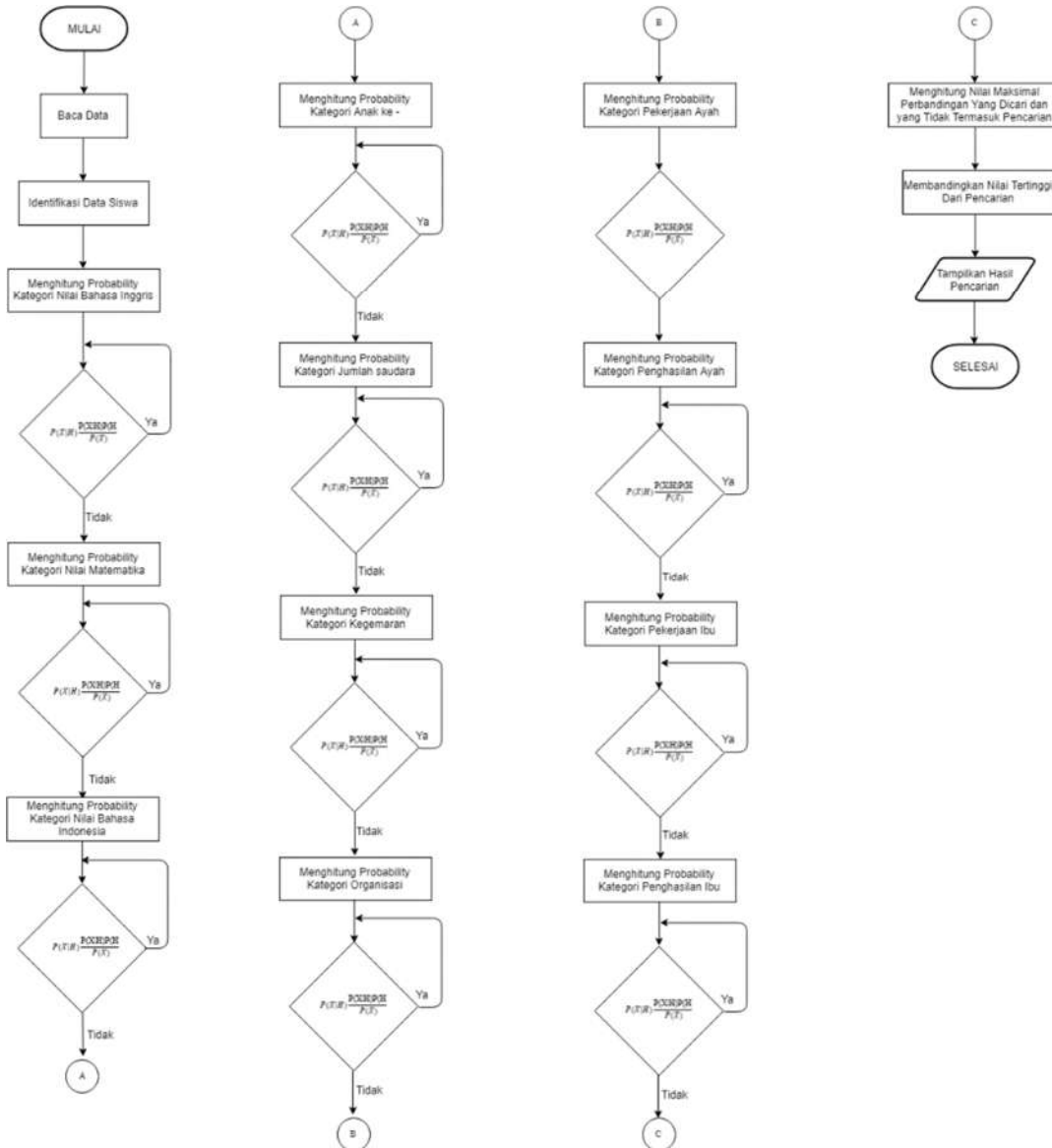
Sistem membaca data terlebih dahulu kemudian data tersebut diidentifikasi. Menghitung sesuai klasifikasi yang ada seperti Nilai Bahasa Inggris, Nilai Matematika, Nilai Bahasa Indonesia, Anak

ke-, Jumlah saudara, kegemaran, organisasi, Pekerjaan Ayah, Penghasilan Ayah, Pekerjaan Ibu, Penghasilan Ibu [4] [5]. Kemudian Sistem menghitung nilai maksimal perbandingan yang dicari maupun yang tidak dicari. Sistem menampilkan hasil dari pencarian.

Untuk masalah klasifikasi, yang dihitung adalah $P(H|X)$, yaitu peluang bahwa hipotesa benar (*valid*) untuk data sample X yang diamati, dimana :

- X adalah Data sampel dengan kelas (label) yang tidak diketahui.
- H adalah Merupakan hipotesa bahwa X adalah data dengan klasifikasi (label) C .
- $P(H)$ adalah peluang dari hipotesa H .
- $P(X)$ adalah peluang data sampel yang diamati.
- $P(X|H)$ adalah peluang data sampel X , bila diasumsikan bahwa hipotesa benar (*valid*).

Dalam pengujian sistem pakar diperlukan presisi dan *recall* untuk menjadi bahan perbandingan apakah sistem yang rancang efektif atau tidak efektif. Berdasarkan pengertiannya menyatakan bahwa presisi adalah perbandingan jumlah dokumen yang relevan terhadap *query* dengan jumlah dokumen yang terambil sesuai dari hasil pencarian, sedangkan *recall* adalah perbandingan jumlah dokumen relevan yang terambil dengan *query* yang diberikan



Gambar 3. Flowchart Sistem

dengan total kumpulan dokumen yang relevan dengan query.

Menurut teori Lancaster menjelaskan dalam pencarian presisi dan *recall* dengan rumus seperti ini :

$$Presisi = \frac{i}{j} \times 100\% \quad (1)$$

$$Recall = \frac{i}{k} \times 100\% \quad (2)$$

Dimana presisi merupakan jumlah dokumen yang relevan (i) dibagi dengan jumlah dokumen yang terpanggil dalam pencarian (j) dan dikalikan dengan

persentase penuh, sedangkan *recall* merupakan jumlah dokumen yang terpanggil (i) dibagi jumlah dokumen yang relevan dalam database (k) dan dikalikan dengan persentase penuh.

Jadi, efektifitas sistem temu kembali informasi dinilai berdasarkan teori Lancaster yaitu relevan dan tidak relevan [3]. Untuk *flowchart* sistem dapat dilihat pada Gambar 3.

3. HASIL DAN PEMBAHASAN

Hasil sistem di uji kebenarannya dengan penilaian guru di sekolah SMK Muhammadiyah 3 Banjarmasin. Guru menilai hasil pada sistem apakah relevan dengan keadaan siswa di sekolah tersebut yang sebenarnya. Hasil sistem dan penilaian guru di tampilkan seperti pada Tabel 1.

Tabel 1. Hasil Penilaian

ID	Penilaian Sistem	Penilaian Guru
8	BERMASALAH	TIDAK RELEVAN
9	BEBAS MASALAH	RELEVAN
10	BEBAS MASALAH	RELEVAN
11	BEBAS MASALAH	RELEVAN
12	BEBAS MASALAH	RELEVAN
13	BEBAS MASALAH	RELEVAN
14	BEBAS MASALAH	RELEVAN
15	BEBAS MASALAH	RELEVAN
16	BEBAS MASALAH	RELEVAN
17	BEBAS MASALAH	RELEVAN
18	BEBAS MASALAH	RELEVAN
19	BEBAS MASALAH	RELEVAN
20	BEBAS MASALAH	RELEVAN
21	BERMASALAH	RELEVAN
22	BEBAS MASALAH	RELEVAN
23	BEBAS MASALAH	RELEVAN
24	BERMASALAH	RELEVAN
25	BEBAS MASALAH	RELEVAN
26	BEBAS MASALAH	RELEVAN
27	BEBAS MASALAH	RELEVAN
28	BEBAS MASALAH	RELEVAN
29	BEBAS MASALAH	RELEVAN
30	BEBAS MASALAH	RELEVAN
31	BEBAS MASALAH	RELEVAN
32	BEBAS MASALAH	RELEVAN
33	BEBAS MASALAH	RELEVAN
34	BEBAS MASALAH	RELEVAN
35	BEBAS MASALAH	RELEVAN
36	BEBAS MASALAH	RELEVAN
37	BEBAS MASALAH	RELEVAN
38	BEBAS MASALAH	TIDAK RELEVAN
39	BEBAS MASALAH	RELEVAN
40	BEBAS MASALAH	RELEVAN
41	BEBAS MASALAH	RELEVAN
42	BEBAS MASALAH	RELEVAN
43	BEBAS MASALAH	RELEVAN
44	BEBAS MASALAH	RELEVAN
45	BEBAS MASALAH	RELEVAN
46	BEBAS MASALAH	RELEVAN
47	BEBAS MASALAH	TIDAK RELEVAN

Oleh karena itu dapat dihitung hasil uji presisi dan *recall* sesuai dengan perhitungan sebagai berikut :

$$\text{Presisi (\%)} = \frac{37}{40} \times 100\% = 92,5\%$$

$$\text{Recall (\%)} = \frac{37}{72} \times 100\% = 51,3\%$$

Berdasarkan hasil uji yang dilakukan bahwa nilai presisi sebesar 92,5% dan nilai *recall* sebesar 51,3%. Hal ini menunjukkan bahwa nilai *recall* lebih rendah daripada nilai presisi, namun tingkat keefektifan dari sistem identifikasi siswa bermasalah tersebut sudah dikatakan efektif, karena efektifitas sistem dikatakan efektif jika hasil dari presisi dan *recall* tersebut diatas 50%.

4. KESIMPULAN

Sistem pendukung keputusan siswa bermasalah dan tidak bermasalah di SMK Muhammadiyah 3 Banjarmasin. Telah dapat di selesaikan. Berdasarkan pengujian yang telah di lakukan, nilai presisi yang ditunjukan sebesar 92,5% . Nilai *recall* dihasilkan sebesar 51,3%.

5. DAFTAR PUSTAKA

- [1] Nikmah Rahmawati (2016) 'KENAKALAN REMAJA DAN KEDISIPLINAN: Perspektif Psikologi dan Islam', 11(April), pp. 267–288.
- [2] Syarli and Muin, A. A. (2016) 'Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus : Data Mahasiswa Baru Perguruan Tinggi)', Jurnal Ilmiah Ilmu Komputer, 2(1), pp. 1–5.
- [3] Saleh, A. (2015) 'Implementasi Metode Klasifikasi Naive Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga', Citec Journal, 2(3), pp. 207–217. doi: doi.org/10.24076/citec.2015v2i3.49
- [4] Cristedi Permana Barus (2013) 'Sosial Ekonomi Keluarga dan Hubungannya dengan Kenakalan Remaja di Desa Lantasan Baru Kecamatan Patumbak Kabupaten Deli Serdang', Welfare StatE, 2(1). Available at: <https://jurnal.usu.ac.id/index.php/ws/article/view/2138/1172>.
- [5] N., Nindya P., M. R. (2012) 'Hubungan antara Kekerasan Emosional Pada Anak terhadap Kecenderungan Kenakalan Remaja', Jurnal Psikologi Klinis dan Kesehatan Mental, 1(02), pp. 1–9. Available at: [http://journal.unair.ac.id/filerPDF/110810221_ringkasan\(1\)_FIX_PKM.pdf](http://journal.unair.ac.id/filerPDF/110810221_ringkasan(1)_FIX_PKM.pdf).

- [6] Afrizal Subhan, A. (2017) 'Rancang Bangun Aplikasi Pembelajaran Dasar Pemrograman Berbasis Mobile Phone', *Jurnal Teknik Informatika Politeknik Sekayu (TIPS)*, VI(1), pp. 4–19