

KLASIFIKASI PENERIMAAN CALON SISWA DI SMPN 1 SITUBONDO DENGAN METODE NAÏVE BAYES

¹ *Alfan (111065xxxxxxxxxxxxx)*

² *Daryanto, S.Kom, M.Kom,* ³ *Heny Wahyu, S.Kom*

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

ABSTRAK

Madrasah Aliyah Negeri (MAN) merupakan jenjang pendidikan menengah yang mengutamakan penyiapan siswa untuk melanjutkan pendidikan ke jenjang yang lebih tinggi dengan cara pengkhususan. Perwujudan pengkhususan tersebut berupa diselenggarakannya penjurusan yang dimulai di kelas X (sepuluh), yakni penjurusan Matematika dan Ilmu Pengetahuan Alam, Ilmu-Ilmu Sosial, Bahasa, dan Agama. Dengan demikian, penjurusan merupakan upaya strategis dalam memberikan fasilitas kepada siswa untuk menyalurkan bakat, minat, dan kemampuan yang dimilikinya untuk dikembangkan secara optimal. Oleh karena itu, sekolah harus mampu menempatkan siswa ke dalam jurusan yang tepat. Untuk menghindari kemungkinan terjadinya ketidak tepatan penjurusan siswa dalam pilihan jurusan tertentu, perlu adanya upaya pengukuran dan penilaian keefektifan perencanaan penjurusan, keefektifan pelaksanaan penjurusan, keberhasilan siswa setelah penjurusan, serta kendala-kendala yang dihadapi dalam penjurusan. Kemudian algoritma *k-means* adalah metode klasterisasi yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana, mudah diimplementasikan, serta memiliki kemampuan untuk mengklaster data yang besar, mampu menangani data banyak.

Kata kunci : *Sistem Analisa Penentuan Penjurusan Siswa MAN 1 Jember Menggunakan Algoritma K-Means Clustering.*

1. Pendahuluan

Lembaga pendidikan seperti pendidikan SMP/MTS sering kali membutuhkan suatu bentuk keputusan dalam menentukan penerimaan siswa baru yang akan mendaftar di SMP/MTS tersebut. Keputusan yang diambil dalam menentukan calon siswa yang akan mendaftar harus sesuai dengan kemampuan calon siswa tersebut yaitu nilai matematika, bahasa indonesia, bahasa inggris dan nilai ujian nasional. Pembuatan keputusan harus benar-benar mempertimbangkan pilihan yang sesuai untuk penentuan calon siswa, sehingga dibutuhkan sistem yang dapat mengklasifikasikan pola pemilihan calon siswa baru dengan mempertimbangkan : Danem dan Piagam-piagam penghargaan, dengan menggunakan metode *Naïve Bayes*. Kemampuan teknologi informasi untuk mengumpulkan dan menyimpan berbagai tipe data jauh meninggalkan kemampuan untuk menganalisis, meringkas, dan mengekstrak pengetahuan dari data. Metode

Naïve Bayes merupakan *classifier*, yaitu metode yang dapat mengklasifikasikan data. Dari permasalahan diatas penulis menyelesaikan permasalahan yang ada di SMPN 1 SITUBONDO, yaitu dengan mengklasifikasikan siswa yang mendaftarkan di sekolah SMPN 1 SITUBONDO berdasarkan nilai kriteria-kriteria yang telah ditetapkan sekolah. Setelah selesai melakukan perhitungan rata-rata dari tiap kriteria tersebut barulah kemudian dapat diketahui apakah siswa tersebut dalam klasifikasi diterima atau tidak diterima.

Dari latar belakang permasalahan diatas, penulis ingin mengangkat sebuah judul “Klasifikasi Penerimaan Calon Siswa Di SMPN 1 Situbondo Dengan Metode Naive Bayes”.

2. Dasar Teori

2.1 Pengertian Klasifikasi

Klasifikasi adalah proses untuk menemukan model yang menjelaskan atau membedakan

konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui (Mulyanto, 2009). Di dalam klasifikasi diberikan sejumlah record yang dinamakan data training set, yang terdiri dari beberapa atribut, atribut dapat berupa kontinyu ataupun kategoris, salah satu atribut menunjukkan kelas untuk record.

2.2 Klasifikasi Naïve Bayes

Klasifikasi *Naïve Bayes* adalah klasifikasi statistic yang dapat memprediksi kelas suatu anggota probabilitas. Algoritma ini memanfaatkan teori probabilitas yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas dimasa depan berdasarkan pengalaman dimasa sebelumnya. Dua kelompok peneliti, satu oleh Pantel dan Lin dan yang lain oleh Microsoft Research memperkenalkan metode statistic *bayesianzi*. Tetapi yang membuat *Naïve Bayesian* ini terkenal adalah pendekatan yang dilakukan oleh Paul Graham.

Naïve Bayes berasumsi bahwa efek dari suatu pada kelas yang diberikan adalah independen terhadap nilai atribut yang lainnya. Asumsi ini biasa disebut dengan *class conditional independence*. Itu dibuat untuk menyederhanakan komputasi yang terkait dan dalam hal ini disebut sebagai “*naive*”. (Perdana, 2009 : 7)

Naïve Bayes Classifier (NBC) merupakan sebuah pengklasifikasian probabilitas sederhana yang mengaplikasikan Teorema *Bayes* dengan asumsi ketidak tergantungan (*independent*) yang tinggi. Keuntungan penggunaan NBC adalah metode ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.

Naïve Bayes klasifikasi adalah metode yang berdasarkan probabilitas dan Teorema *Bayesian* dengan asumsi bahwa setiap variabel bersifat bebas (*independence*) dan mengamsusikan bahwa keberadaan sebuah fitur (*variable*) tidak ada kaitanya dengan beradaan fitur (*variable*) yang lain. *Naïve Bayes* adalah model penyederhanaan dari metode *bayes*. *Naïve bayes* inilah yang digunakan di dalam *macine learning* sebagai metode untuk mendapatkan hipotesis untuk suatu keputusan. (Basuki, 2006)

Salah satu penerapan Teorema Bayes adalah dalam klasifikasi Naïve Bayes sebagai berikut :

$$P(B_r | A) = \frac{P(B_r \cap A)}{P(A | B_r)} = \frac{P(B_r) * P(A | B_r)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r) * P(A | B_r)}{\sum_{i=1}^k P(B_i)}$$

Rumus diatas merupakan rumus untuk data diskrit, sedangkan untuk data kontinue rumus yang dipakai adalah :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

σ = standar deviasi

μ = nilai rata-rata kemunculan

x = nilai rata-rata

π = phi, nilainya 3,14

e = bilangan natural, nilainya 2,712...

Contoh perhitungan *Naïve Bayes Classifier* :

3. Tahapan Penelitian

3.1 Tahapan Penelitian

Dalam pengerjakan Tugas Akhir ini diperlukan langkah-langkah Kegiatan Penelitian untuk mendapatkan hasil yang maksimal. Untuk itu penulis merencanakan suatu langkah-langkah yang dapat memaksimalkan dalam pengerjaan Tugas Akhir ini. Langkah-langkah itu adalah sebagai berikut :

a. Study Literatur

Untuk Melakukan pencarian literatur, referensi, dan tutorial tentang *Naïve Bayes* dan semua materi yang berhubungan dengan masalah yang akan dibahas dengan mengambil/mendownload paper di internet.

b. Pengumpulan Dataset

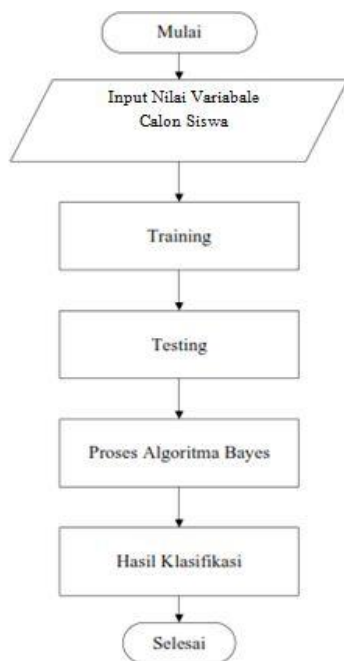
Penelitian ini akan menggunakan dataset siswa yang mendaftarkan di sekolah SMPN 3 SITUBONDO.

Data diperoleh langsung dari TU SMPN 3 SITUBONDO.

c. Implementasi dan Rancangan Sistem

Sistem ini akan di implementasikan dengan menggunakan bahasa *Visual Basic.Net* adalah sebuah alat untuk mengembangkan dan membangun aplikasi yang bergerak dengan sistem .NET Framework, dan menggunakan bahasa BASIC.

dijelaskan bahwa flowchart metode Teorema Bayes ini menjelaskan alur dari aplikasi klasifikasi dengan menggunakan metode *Naïve Bayes*. Pertama input nilai variabel calon penerima siswa, setelah itu data di training, setelah data ditraining data diproses testing kemudian diproses dengan algoritma *Naive bayes*, setelah melakukan proses training dan testing data akan terklasifikasi.



Gambar Flowchart Metode Naïve Bayes

d. Pengujian Data Training dan Penentuan Kelas

Setelah mengumpulkan data akan dilakukan pembuatan data training dan penentuan kelas “calon siswa” dan “tidak calon siswa” dari data siswa di SMPN 3 SITUBONDO, yaitu membuat data training dimana terdapat 3 variabel yang terdiri dari :

- Danem
- Piagam – Piagam Penghargaan

Tabel 3.1 Kriteria Penentuan Penerima Beasiswa

Nama Kriteria	Nilai Tidak Diterima	Nilai Diterima
Danem	0 - 50	60 - 100
Piagam Penghargaan	0 - 50	60 - 100

Untuk kriteria di SMPN 3 SITUBONDO dalam menenukan standart prasyarat calon siswa untuk masuk di sekolah

SMPN 3 SITUBONDO Danem dan Piagam Penghargaan adalah 60-100 dikarenakan pihak sekolah telah menentukan batasan untuk siswa yang masuk di sekolah tersebut dengan nilai total siswa untuk danem dan piagam penghargaan. Kemudian yang diterima dengan kriteria danem nilainya mulai 60 – 100, sedangkan tidak diterima 0 – 50. Nilai tersebut diambil dari total danem siswa. Kemudian Untuk kriteria Piagam penghargaan nilainya mulai 60 – 100, sedangkan tidak diterima 0 – 50. Nilai tersebut diambil dari total piagam penghargaan siswa.

3.2 Perhitungan Naïve Bayes

Berikut ini adalah analisa terhadap kriteria-kriteria untuk penentuan penerima bantuan beasiswa.

Data training yang terdiri dari 327 calon siswa akan di klasifikasikan oleh sistem untuk selanjutnya dijadikan acuan dalam perhitungan naïve bayes. Kemudian kita cari berapa banyak kemunculan :

1. Nilai rata-rata siswa yang diterima mendapatkan calon siswa
2. Nilai rata-rata siswa yang diterima masuk ke daftar tidak diterima
3. Nilai rata-rata siswa yang tidak diterima masuk ke daftar diterima

Setelah kita menuliskan nilai rata-rata sesuai kriteria diatas maka kita cari standar deviasi dan nilai means(rata-rata)nya, yaitu dengan cara:

1. Rumus Mean :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

2. Rumus Standart Deviasi :

$$S_d = \sqrt{\frac{\sum_{i=1}^{i=n} (x - \bar{x})^2}{n-1}}$$

dimana:

x = nilai tiap data

\bar{x} = nilai rata-rata

n = jumlah banyaknya data

S_d = Standar deviasi

Setelah itu penghitungan *Naïve bayes* akan dilanjutkan dimana kali ini dicontohkan pada siswa yang bernama Fatriatul lailatul.

Rumus Naïve Bayes Continue :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

dimana :

σ = standar deviasi

μ = nilai rata-rata kemunculan rata-rata Diterima/Tidak diterima,

contohnya nilai rata-rata Diterima yang masuk di daftar Diterima, nilai rata - rata Tidak diterima yang masuk di daftar Diterima.

x = nilai rata-rata Diterima / Tidak diterima

π = phi, nilainya 3,14

e = bilangan natural, nilainya 2,712...

$$f(K1|Danem) = \frac{1}{10.3692048\sqrt{2\pi}} e^{-\frac{-1(62-76)^2}{2(10.3692048)^2}} = 0.014476239$$

$$f(K2|P. Penghargaan) = \frac{1}{11.12762695\sqrt{2\pi}} e^{-\frac{-1(76-76)^2}{2(11.12762695)^2}} = 0.038438801$$

$$f(K1| Danem) = \frac{1}{9.443114624\sqrt{2\pi}} e^{-\frac{-1(62-75)^2}{2(9.443114624)^2}} = 0.01638198$$

$$f(K2| P. Penghargaan) = \frac{1}{10.11997001\sqrt{2\pi}} e^{-\frac{-1(76-76)^2}{2(10.11997001)^2}} = 0.03943128$$

Sehingga :

$$\text{Likelihood Diterima} = K1 \text{ Danem} \times K2 \text{ Piagam Penghargaan} \\ = 0.014476239 \times 0.038438801$$

$$= 0.000556449$$

$$\text{Likelihood Tidak Diterima} = K1 \text{ Danem} \times K2 \text{ Piagam Penghargaan} \\ = 0.01638198 \times 0.03943128 \\ = 0.000645963$$

Nilai probabilitas dapat dihitung dengan melakukan normalisasi terhadap *likelihood* tersebut sehingga jumlah nilai yang diperoleh = 1

$$\text{Probabilitas Diterima} = \frac{\text{Like Diterima}}{(\text{Like Diterima} + \text{Like Tidak Diterima})} \\ = \frac{0.000556449}{(0.000556449 + 0.000645963)} \\ = 0.462777601$$

$$\text{Probabilitas Tidak Diterima} = \frac{\text{Like Tidak Diterima}}{(\text{Like Diterima} + \text{Like Tidak Diterima})} \\ = \frac{0.000645963}{(0.000556449 + 0.000645963)} \\ = 0.537222399$$

Klasifikasi = Tidak Diterima karena nilai probabilitas Tidak Diterima lebih besar dari pada nilai probabilitas diterima.

4. Pengujian dan Analisa

Agar mempermudah proses perhitungan dan penerimaan calon siswa, maka dibuatlah program penerimaan calon siswa tersebut dengan bahasa pemrograman VB.Net

• Form Input Data Training

Pada form input data training dapat menambahkan atau menginputkan hasil penerimaan calon siswa di SMPN 3 Situbondo.

• Form Testing

Pada form testing akan mendata siswa di SMPN 3 Situbondo mulai dari diterima dan tidak diterima

• Form Perhitungan Naive Bayes

Pada form perhitungan algoritma naive bayes akan menghitung calon penerima

beasiswa di SMPN 3 Situbondo sesuai dengan criteria yang ada disekolah tersebut.

5. Kesimpulan

Dari pembuatan calon siswa baru Dengan Metode Naïve Bayes didapat beberapa kesimpulan antara lain :

Dari 100 data training terdapat 4 calon siswa baru yang terklasifikasi tidak sama setelah dilakukan perhitungan naïve bayes. Didapatkan hasil 50 data diterima dan 46 data tidak diterima. Hal tersebut dikarenakan terdapat perbedaan hasil mean dan standart deviasi antara perhitungan manual dan perhitungan sistem.

DAFTAR PUSTAKA

- Mulyanto, Agus, 2009. "Sistem Informasi Konsep & Aplikasi". Cetakan I, Yogyakarta: Pustaka Pelajar, hal. 204-206.
- Rezdy Anugrah Perdana. "*Aplikasi Rekomendasi Iklan Jejaring Sosial Facebook Menggunakan Algoritma Naïve Bayes*" Progam Studi Ilmu Komputer Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam Universitas Pendidikan Indonesia, 2009.
- Basuki, Ahmad, 2006. "Metode Byes". Kuliah PENS-ITS.
- Tan *et al.* 2006. *Introduction to Data Mining*. USA: Addison Wesley.