



## Prediksi Perguruan Tinggi Negeri dengan Menggunakan Metode Naive Bayes

Rahmania Aulia Ikimah Putri<sup>1</sup>, Tacbir Hendro Pudjiantoro<sup>2</sup>

<sup>1</sup>Teknik Informatika, Sains dan Informatika, Universitas Jenderal Achmad Yani

<sup>2</sup> Teknik Informatika, Sains dan Informatika, Universitas Jenderal Achmad Yani  
Rahmania2aulia@gmail.com

### Abstract

*Achievement is the learning outcomes obtained by someone where the ability to increase over time, achievement is not caused by the growth process but with the learning situation. The ability to master knowledge that has been tested for certainty and truth so that it can be measured in the form of grades or what is usually referred to as academic achievement. At least XYZ School students are accepted into state universities. By using the Naive Bayes method, the prediction results are accurate enough to facilitate the Counseling Guidance Section (BK) directing students to register for State Universities (PTN) through the National Higher Education Entrance Entrance Test (SNMPTN)*

*Keywords: Achievement; Prediction; Naïve-Bayes;*

### Abstrak

Prestasi yaitu hasil belajar yang didapatkan oleh seseorang dimana kemampuan yang dapat bertambah selama beberapa waktu, prestasi tidak disebabkan proses pertumbuhan tetapi dengan adanya situasi belajar. Kemampuan menguasai ilmu pengetahuan yang telah diuji kepastian dan kebenarannya sehingga bisa diukur baik berupa nilai maupun yang biasanya disebut dengan prestasi akademik. Sedikitnya siswa/i Sekolah XYZ yang diterima di perguruan tinggi negeri. Dengan menggunakan metode naive bayes mendapatkan hasil prediksi yang cukup akurat untuk mempermudah Bagian Bimbingan Koreseling (BK) mengarahkan siswa/i mendaftar Perguruan Tinggi Negeri (PTN) melalui jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN)

Kata kunci: Prestasi; Prediksi; Naive Bayes

### 1. Pendahuluan

Pendidikan merupakan bagian terpenting dalam aspek kehidupan sebagai bekal dalam rangka membentuk manusia yang cerdas dan berkualitas. Sesuai dengan undang-undang Sistem Pendidikan Nasional Republik Indonesia No. 2 tahun 2003 Bab II Pasal 3 yaitu pendidikan nasional bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertaqwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab [1]

Prestasi belajar adalah puncak hasil belajar yang dapat mencerminkan hasil keberhasilan siswa terhadap tujuan belajar yang telah ditetapkan. Hasil belajar tersebut berupa nilai, nilai merupakan perumusan terakhir yang dapat diberikan oleh guru mengenai kemajuan atau prestasi belajar yang diperoleh siswa selama masa. Hasil belajar dikatakan tuntas apabila nilai yang didapat telah memenuhi Kriteria Ketuntasan Minimum (KKM)

yang telah ditetapkan oleh masing-masing guru mata pelajaran [2]

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database yang besar. Data mining juga sering disebut *knowledge discovery in database (KDD)*, yang semua prosesnya adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dengan set data yang berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan [3].

Prediksi juga merupakan suatu usaha untuk meramalkan masa depan dengan memeriksa masa lalu. Terdiri dari mengestimasi dari besarnya masa depan dari beberapa variabel seperti penjualan atas dasar pengetahuan dari masa lalu dan sekarang. Prediksi digunakan untuk memprediksi kejadian dimasa yang akan datang [4]

Pendidikan di Sekolah XYZ mempunyai suatu permasalahan yaitu sulitnya siswa/i Sekolah XYZ untuk masuk ke Perguruan Tinggi Negeri (PTN). Faktor seperti keaktifan dan pekerjaan orang tua menjadi faktor penentu apakah siswa/i tersebut dapat melanjutkan ke PTN yang di pilih. Penelitian sebelumnya mengemukakan pengaruh keaktifan mahasiswa dalam organisasi dan motivasi belajar baik secara parsial maupun simultan terhadap prestasi belajar. Hasil penelitian menunjukkan bahwa: (1) Terdapat pengaruh positif dan signifikan keaktifan mahasiswa dalam organisasi terhadap prestasi belajar mahasiswa. (2) Terdapat pengaruh positif dan signifikan motivasi belajar terhadap prestasi belajar mahasiswa. (3) Terdapat pengaruh yang signifikan keaktifan mahasiswa dalam organisasi dan motivasi belajar secara simultan terhadap prestasi belajar mahasiswa

## 2. Metode Penelitian

Metode ini didasarkan pada *classifier Naive bayes* dan tujuannya adalah untuk mengetahui apa yang siswa dapat peroleh di hasil akhir semester mereka. Mereka dapat memperoleh manfaat dari prediksi hasil siswa dalam beberapa cara. Guru dan siswa mengambil langkah-langkah penting untuk mengembangkan hasil siswa yang hasil prediksi tidak terpenuhi dan satu set pelatihan data siswa diambil untuk membangun model naïve bayes dan kemudian diterapkan pada data uji untuk menemukan hasil akhir siswa. semester.[5]

Memprediksi adalah salah satu teknik penambangan data yang digunakan dalam berbagai disiplin ilmu untuk menemukan dan menganalisis data historis dan membuat prediksi dengan menggunakan teknik yang berbeda seperti statistik, penambangan data, dan pembelajaran mesin. Metode statistik seperti regresi dapat secara matematis mewakili interaksi antara variabel yang berbeda di bawah beberapa pertimbangan. Kemudian, teknik *Artificial Intelligence* (AI) mendapatkan popularitas dan menggantikan model statistik di sebagian besar aplikasi karena mereka dikenal efisien dan kurang memakan waktu dalam pemodelan sistem yang kompleks dibandingkan dengan model matematika seperti regresi [6]

Kesalahan dapat diperkecil menggunakan prediksi dimana proses ini memperkirakan secara sistematis mengenai sesuatu yang akan terjadi dimasa depan berdasarkan masa lalu. Jawaban yang dihasilkan melalui prediksi tidak harus pasti terjadi melainkan berusaha untuk mencari hasil yang akan mungkin terjadi.[7]

Menurut teorema Bayes, probabilitas masalah klasifikasi dengan memberikan karakteristik sampel X dan sampel milik kategori Y adalah  $P(y|x) = P(x|y)P(x)$ . Dalam hal ini, X adalah vektor eigen, dan mengatur dimensi X ke M. Menurut rumus probabilitas penuh dan asumsi Naif (kondisi karakteristik independen),

rumus (1) dapat diekspresikan  $P(y = c_k|x) = \prod_{i=1}^n P(x_i | y = c_k) \sum_k P(y = c_i) P(x_i | y = c_k)$ . Dalam hal ini, kita hanya dapat memperkirakan probabilitas bersyarat dari masing-masing karakteristik X di setiap kelas. Probabilitas sebelumnya dari kelas Y dapat dihitung dengan set pelatihan. Demikian pula, dengan statistik pada set pelatihan, vektor probabilitas bersyarat yang sesuai dengan setiap jenis karakteristik independen bersyarat dapat diperoleh [8]

$$P(X|Y) = \frac{P(X|Y)P(Y)}{P(X)}$$

Keterangan:

X : Data dengan class yang belum diketahui

Y : Hipotesis data X merupakan suatu class spesifik

P(Y|X) : Probanilitas hipotesis Y berdasarkan kondisi X

P(Y) : Probabilitas hipotesis Y

P(X|Y) : Probabilitas X berdasarkan kondisi hipotesis Y

P(X) : Probalitas X

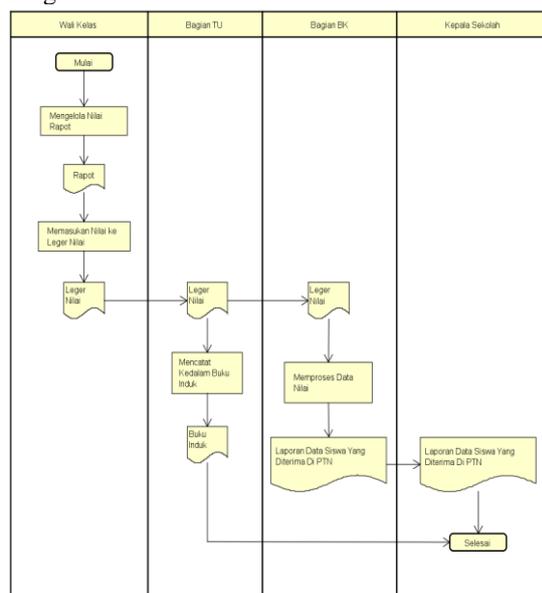
Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan [9]

Kelemahan harus mengasumsi bahwa antar fitur tidak terkait (independent) Dalam realita, keterkaitan itu ada Keterkaitan tersebut tidak dapat dimodelkan oleh Naïve Bayesian Classifier [10]

## 3. Hasil dan Pembahasan

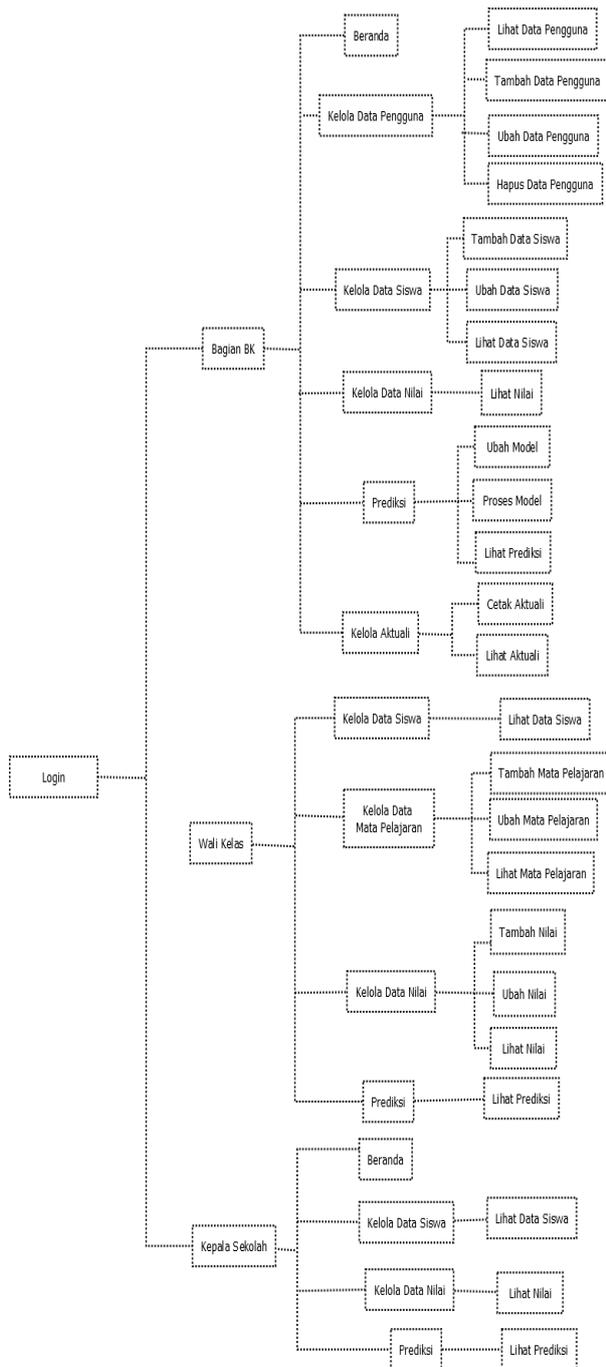
### 1. Hasil Kebutuhan Fungsional Sistem

Berikut ini adalah model kebutuhan fungsionalitas sistem yang digambarkan menggunakan Flowmap Diagram



Gambar 1 Flowmap Diagram





Gambar 6 Peta Aplikasi

## 2. Perhitungan Metode

Berikut hasil dari prediksi PTN dengan menggunakan metode naive bayes. Dengan data 50 yang sudah diperhitungkan dengan menggunakan metode naive bayes.

Dataset PTN memiliki 13 kelas masalah, sehingga dapat diketahui probabilitas (P) masing-masing kelas dengan cara membagi nilai frekuensi/jumlah data untuk masing-masing kelas masalah, dengan jumlah total nilai frekuensi/jumlah data pada kelas tersebut, seperti yang ditunjukkan dalam perhitungan berikut ini :

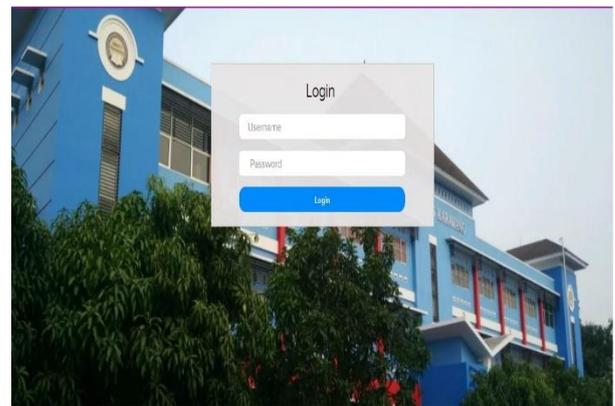
Tabel 1 Class Probabilitas

PTN	IPB	4/50	0,08
PTN	UGM	4/50	0,08
PTN	PNJ	2/50	0,04
PTN	UNPAD	4/50	0,08
PTN	UNSIKA	14/50	0,28
PTN	UNJ	1/50	0,02
PTN	UPI	4/50	0,08
PTN	UIN JAKARTA	1/50	0,02
PTN	UIN BANDUNG	2/50	0,04
PTN	UNS	2/50	0,04
PTN	SWASTA	10/50	0,2
PTN	UNSOED	1/50	0,02
PTN	AKA	1/50	0,02

Hasil dari perhitungan naive bayes tersebut maka dihasilkan data PTN untuk tahun selanjutnya di Sekolah XYZ

## 3. Implementasi Antar Muka

Tampilan sistem halaman awal dari sistem prediksi PTN untuk masuk kedalam sistem. Antar muka halaman login Sistem



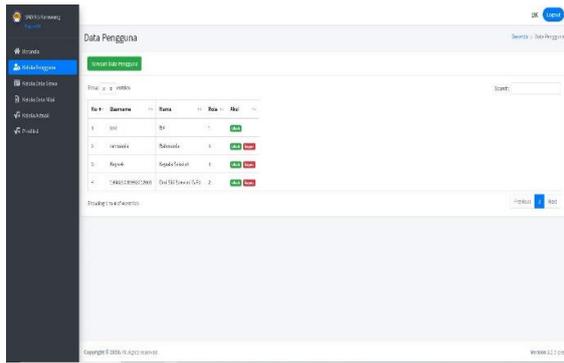
Gambar 7 Login

Antarmuka halaman awal untuk Bagian BK pada sistem prediksi



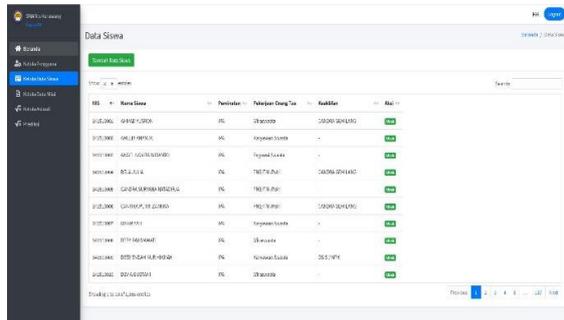
Gambar 8 Beranda

Antarmuka lihat pengguna untuk melihat pengguna pada system



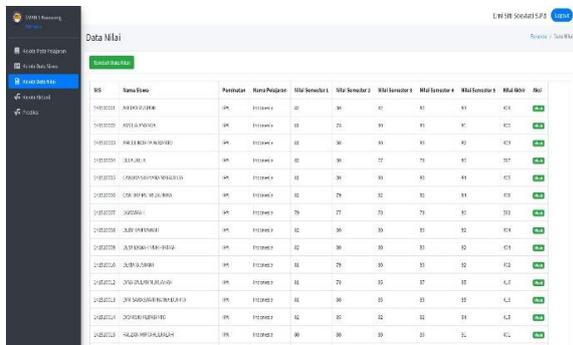
Gambar 9 Kelola Data Pengguna

Antarmuka lihat pengguna untuk melihat pengguna pada sistem



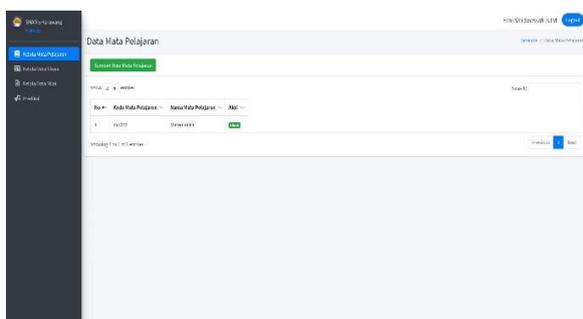
Gambar 10 Kelola Siswa

Antarmuka lihat nilai untuk melihat nilai baru pada sistem



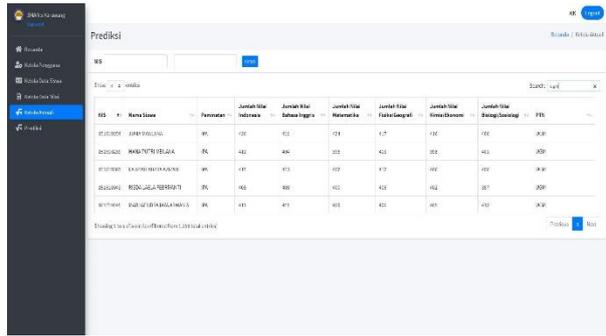
Gambar 11 Kelola Nilai

Antarmuka lihat mata pelajaran untuk melihat nilai baru pada sistem



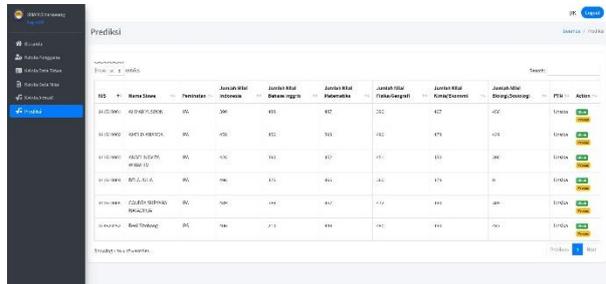
Gambar 12 Kelola Data Mata Pelajaran

Antarmuka lihat aktuali untuk melihat nilai baru pada sistem



Gambar 13 Kelola Aktuali

Antarmuka lihat prediksi untuk melihat nilai baru pada sistem



Gambar 14 Kelola Prediksi

#### 4. Kesimpulan

Berdasarkan pada analisis dan perancangan, penelitian ini dilakukan untuk menentukan prediksi yang dapat mendaftar di PTN sesuai dengan perhitungan dengan menggunakan Naive Bayes.

#### Ucapan Terimakasih

Penulis mengucapkan terimakasih kepada Perguruan Tinggi Universitas Jenderal Achmad Yani yang membantu dalam proses penelitian, serta Sekolah XYZ yang telah memfasilitasi penulis untuk melakukan penelitian

#### Daftar Rujukan

- [1] C. P. Intani and I. Ifdil, "Hubungan kontrol diri dengan prestasi belajar siswa," *IICET*, vol. 4, no. 4, pp. 65–70, 2018.
- [2] A. Yulianto, "Pengaruh Keaktifan Siswa Berorganisasi Terhadap Peningkatan Soft Skills dan Prestasi Belajar Siswa SMK Muhammadiyah Prambanan," *E-Jurnal Pendidikan Teknik Mesin*, vol. 3, no. 1, pp. 329–336, 2015.
- [3] T. H. Pudjiantoro, F. Renaldi, and A. Teogunadi, "Penerapan Data Mining untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa Baru," *Konferensi Nasional Sistem dan Informatika*, vol. 1, no. February, pp. 51–60, 2011.
- [4] I. Ali and L. Sularto, "Optimasi Parameter Artificial Neural Network Menggunakan Algoritma Kemungkinan Prediksi Kelulusan Mahasiswa," *Jurnal ICT*, vol. 18, no. 1, pp. 54–59, 2019.
- [5] B. K. Francis and S. S. Babu, "Predicting Academic Performance of Students Using a Hybrid Data Mining Approach," *Journal of Medical System*, vol. 1, no. 1, p. 15, 2019.

- [6] S. Al-Janabi and M. A. Mahdi, "Evaluation prediction techniques to achievement an optimal biomedical analysis Samaher Al-Janabi \* and Muhammed Abaid Mahdi," *Int. J. Grid and Utility Computing*, vol. 10, no. 5, pp. 512–527, 2019.
- [7] A. Saleh, "Penerapan Data Mining Dengan Metode Klasifikasi Naive Bayes untuk Memprediksi Kelulusan Mahasiswa dalam Mengikuti English Proficiency Test," *Citec Journal*, vol. 1, no. 1, p. 6, 2017.
- [8] X. Zhang, R. Xue, B. Liu, W. Lu, and Y. Zhang, "Grade Prediction of Student Academic Performance with Multiple Classification Models," *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, vol. 1, no. 1, pp. 1086–1090, 2018.
- [9] A. Saleh, "Penerapan Data Mining dengan Metode Klasifikasi Naive Bayes untuk Memprediksi Kelulusan Mahasiswa dalam Mengikuti English Proficiency," *ResearchGate*, vol. 1, no. February 2015, p. 7, 2016.
- [10] A. A. Muin and Syarli, "Metode Naive Bayes Untuk Prediksi Kelulusan ( Studi Kasus : Data Mahasiswa Baru Perguruan Tinggi )," *Jurnal Ilmiah Ilmu Komputer*, vol. 2, no. 1, pp. 22–26, 2016.
- [11] National electronic Library for Health, 2003. *Can walking make you slimmer and healthier?* (Hitting the headlines article) [Online] (Updated 16 Jan 2005) Tersedia di : <http://www.nhs.uk.hth.walking>. [Accessed 10 April 2005]