SISTEM PREDIKSI PENGUNDURAN DIRI PESERTA DIDIK MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*

ARTIKEL SKRIPSI

Diajukan Untuk Memenuhi Sebagian Syarat Guna
Memperoleh Gelar Sarjana Komputer (S.Kom)
Pada Program Studi Teknik Informatika

OLEH :

PUPUT SHINTA DEWI
NPM: 12.1.03.02.0290

FAKULTAS TEKNIK
UNIVERSITAS NUSANTARA PERSATUAN GURU REPUBLIK INDONESIA
UN PGRI KEDIRI
2016

Skripsi oleh:

PUPUT SHINTA DEWI

NPM: 12.1.03.02.0290

Judul :

SISTEM PREDIKSI PENGUNDURAN DIRI PESERTA DIDIK MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER

Telah disetujui untuk diajukan Kepada
Panitia Ujian/Sidang Skripsi Program Studi Teknik Informatika
Fakultas Teknik UN PGRI Kediri

Tanggal : 11 Agustus 2016

Pembimbing I	Pembimbing II
Rini Indriati, S.Kom., M.Kom.	**Risky Aswi Ramadhani, M.Kom.**
NIDN. 0725057003	NIDN. 0708049001

Skripsi oleh:

PUPUT SHINTA DEWI

NPM: 12.1.03.02.0290

Judul :

SISTEM PREDIKSI PENGUNDURAN DIRI PESERTA DIDIK MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER

Telah dipertahankan di depan Panitia Ujian/Sidang Skripsi
Program Studi Teknik Informatika Fakultas Teknik UN PGRI Kediri
Pada tanggal : 11 Agustus 2016

Dan Dinyatakan telah Memenuhi Persyaratan

Panitia Penguji :

1. Ketua : Rini Indriati, S.Kom., M.Kom. _____
2. Penguji I : Suratman, SH., M.Pd. _____
3. Penguji II : Risky Aswi Ramadhani, M.Kom. _____

Mengetahui,
Dekan Fakultas Teknik

Dr. Suryo Widodo, M.Pd.
NIP. 19640202 199103 1 002

SISTEM PREDIKSI PENGUNDURAN DIRI PESERTA DIDIK MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*

Puput Shinta Dewi
12.1.03.02.0290
Teknik – Teknik Informatika
puput.shinta@yahoo.co.id
Rini Indriati, S.Kom., M.Kom. dan Risky Aswi Ramadhani, M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

AGNES *English Course* merupakan lembaga kursus bahasa Inggris yang setiap tahunnya membuka pendaftaran untuk menerima peserta didik baru. Namun dari keseluruhan peserta didik yang terdaftar, terdapat sejumlah peserta didik yang mengundurkan diri tanpa adanya konfirmasi kepada pihak lembaga kursus. Hal ini menyebabkan pihak lembaga kursus tidak mampu mempertahankan peserta didiknya. Keterbatasan pihak lembaga kursus dalam memprediksi pengunduran diri peserta didik dalam jumlah besar tidak mungkin dilakukan secara manual karena membutuhkan waktu yang cukup lama. Oleh karena itu diperlukan metode agar dapat mencapai tujuan dari penelitian ini untuk membangun aplikasi yang membantu pihak lembaga kursus dalam memprediksi pengunduran diri peserta didik. Suatu metode yang bisa memfasilitasinya adalah Metode *Naive Bayes Classifier*.

Naive Bayes Classifier adalah pendekatan yang mengacu pada *Teorema Bayes*, dengan mengkombinasikan pengetahuan sebelumnya dengan pengetahuan baru. Metode ini digunakan untuk mengklasifikasikan peserta didik dalam kategori mengundurkan diri atau bertahan dengan memanfaatkan data peserta didik. Sehingga menghasilkan informasi prediksi pengunduran diri peserta didik melalui teknik *data mining*. Variabel penentu yang digunakan meliputi jarak tempuh, usia, pekerjaan, pendidikan terakhir, jenis kelamin, dan program kursus. Hasil pengujian tehadap 100 data testing yang sudah dilakukan pada sistem prediksi pengunduran diri dengan menggunakan metode *Naive bayes Classifier* menghasilkan 77 data peserta didik atau sebesar 77 % terklasifikasi bertahan dan 23 data atau sebesar 23 % terklasifikasi mengundurkan diri. Diharapkan berdasarkan hasil prediksi pengunduran diri peserta didik, pihak lembaga kursus dapat melakukan tindakan-tindakan yang diperlukan dalam mempertahankan peserta didik serta meminimalkan kemungkinan kerugian yang timbul akibat pengunduran diri peserta didik.

Kata Kunci : *Data Mining*, Klasifikasi, *Naive Bayes Classifier* (NBC), Prediksi.

I. LATAR BELAKANG

Active Guidance of Nucleus English Studies (AGNES) *English Course* merupakan lembaga kursus bahasa Inggris yang telah berdiri sejak tahun 1992 dan telah banyak membantu peserta didiknya dalam meningkatkan kemampuan untuk berbahasa Inggris. Peserta didik pada AGNES *English Course* dibimbing agar tidak hanya mahir dalam bahasa Inggris tulis tapi juga mahir dalam bahasa Inggris lisan. Keberhasilan peserta didik dalam berbahasa Inggris merupakan salah satu

aspek penting dalam evaluasi keberhasilan penyelenggaraan kegiatan belajar mengajar pada AGNES *English Course*. Namun, sering kali sejumlah peserta didik pada AGNES *English Course* mengundurkan diri tanpa melakukan konfirmasi kepada pihak lembaga kursus. Pihak lembaga kursus pun tidak mampu mempertahankan peserta didiknya karena peserta didik telah terlebih dahulu mengundurkan diri.

Tahun 2014 misalnya, terdapat 231 peserta didik yang terdaftar sebagai peserta didik. Namun dari keseluruhan peserta didik yang bertahan hingga akhir program kursus berjumlah 182 peserta didik atau sebesar 79% dari keseluruhan peserta didik. Sedangkan 49 peserta didik mengundurkan diri atau sebesar 21% dari keseluruhan peserta didik. Apabila kemungkinan pengunduran diri peserta didik dapat diketahui sejak dini, maka pihak lembaga kursus dapat melakukan tindakan-tindakan yang diperlukan untuk mempertahankan peserta didik tersebut.

Keterbasan kemampuan pihak lembaga kursus untuk mempertahankan dan memprediksikan pengunduran diri peserta didik dalam jumlah besar tidak mungkin dilakukan secara manual karena membutuhkan waktu yang cukup lama. Serta belum adanya suatu sistem yang membantu dalam memprediksi pengunduran diri peserta didik, merupakan beberapa faktor penyebab lembaga kursus ini kehilangan peserta didiknya. Sehingga pengunduran diri peserta didik akan berdampak pada manajemen lembaga kursus. Oleh karena itu diperlukan suatu metode yang dapat memprediksi pengunduran diri.

Suatu metode yang dapat memfasilitasinya adalah metode *Naive Bayes Classifier*. *Bayes* merupakan teknik prediksi berbasis *probabilistik* sederhana yang berdasarkan pada penerapan *Teorema Bayes* dengan asumsi independensi (ketidaktergantungan) yang kuat (*naif*) (Prasetyo, 2012). Pada penelitian ini, menggunakan teknik data mining dengan metode *Naive Bayes Classifier* untuk memprediksi pengunduran diri peserta didik. Sampel data yang digunakan adalah data peserta didik kelas reguler angkatan tahun 2013 sampai 2015 secara acak sebagai *data training*, dengan variabel yang meliputi jarak tempuh, usia, pekerjaan, pendidikan terakhir, jenis kelamin dan program kursus. Data peserta didik yang masih aktif dalam kegiatan kursus akan digunakan sebagai *data testing*. Peserta didik akan diklasifikasikan dalam kategori mengunduran diri atau bertahan.

Adanya penelitian ini diharapkan dapat bertujuan untuk memprediksi pengunduran diri peserta didik dengan mengolah data peserta didik serta dapat

mengimplementasikan metode *Naive Bayes Classifier*. Sehingga memberikan informasi yang berguna untuk meningkatkan kualitas atau dijadikan suatu acuan bagi pihak lembaga kursus untuk dapat melakukan tindakan-tindakan yang diperlukan dalam mempertahankan peserta didik.

II. METODE

Salah satu penerapan *Teorema Bayes* dalam klasifikasi adalah *Naive Bayes*. *Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara konditional saling bebas jika diberikan nilai output. Atau, dengan kata lain, diberikan nilai *output*, probabilitas mengamati secara bersama adalah produk dari probabilitas individu atau $P(a_1, a_2, \dots, a_n | \upsilon_j) = \Pi_i P(a_i | \upsilon_j)$. Memasukkan persamaan ini ke persamaan 1 akan didapat pendekatan yang dipakai dalam klasifier *Naive Bayes* :

$$\upsilon_{NB} = \arg\max_{\upsilon_j \epsilon V} P(\upsilon_j) \prod P(a_i | \upsilon_j) \dots (1)$$

Dimana υ_{NB} adalah nilai hasil klasifikasi *Naive Bayes* dan $P(a_i | \upsilon_j)$ adalah rasio atara $\frac{n_c}{n}$, di mana n_c adalah jumlah *data training* untuk $\upsilon = \upsilon_j$ dan $a = a_i$ dan n adalah total kemungkinan output (Santoso, 2007).

Memprediksi pengunduran diri peserta didik dengan menggunakan metode *Naive Bayes Classifier* diperlukan variabel untuk melakukan perhitungan sehingga akan didapatkan hasil klasifikasi. Klasifikasi yang dimaksud adalah klasifikasi mengundurkan diri atau bertahan berdasarkan veriabel yang telah ditentukan. Variabel yang telah ditentukan terlebih dahulu dilakukan pengelompokan berdasarkan tipe data diskrit (kategori) dan data kontinu (numerik). Variabel penentu yang digunakan dalam mengklasifikasikan data peserta didik, yaitu $a_1 =$ jarak tempuh, $a_3 =$ pekerjaan, $a_4 =$ pendidikan terakhir, $a_5 =$ jenis kelamin, dan $a_6 =$ program kursus untuk variabel dengan tipe data diskrit (kategori), sementara $a_2 =$ usia adalah varabel dengan tipe data kontinu (numerik).

Adapun alur dari metode *Naive Bayes* dalam sistem prediksi pengunduran diri peserta didik adalah sebagai berikut :

a. Membaca *Data Training* Peserta Didik
b. Menghitung Jumlah, Mean, Standar Deviasi, dan Probabilitas

Menghitung jumlah, mean, standar deviasi dan probabilitas berdasarkan pada *data testing* peserta didik. Sebagai contoh *data testing* dengan kriteria bernama Adinda Putri Adhim, jarak tempuh antara rumah dan tempat kursus dekat (< 4 km), usia 11 tahun, pekerjaannya sebagai pelajar serta pendidikan terakhir pada saat mendaftar adalah TK, berjenis kelamin

perempuan dan program kursus yang diambil adalah EFC. Berdasarkan data diatas dapat diketahui klasifikasi data peserta didik apabila diberikan contoh input *data testing* (uji) seperti diatas yakni berupa nama = Adinda Putri Adhim, a_1 = dekat, a_2 = 11, a_3 = pelajar, a_4 = TK, a_5 = P dan a_6 = EFC menggunakan metode *Naive Bayes Classifier* dalam pengklasifikasianya. Data uji peserta didik tersebut dapat ditentukan masuk klasifikasi "Bertahan" atau "Mengundurkan Diri" melalui langkah berikut.

1) Mencari nilai Probabilitas $\mathrm{P}(v_j)$.

Langkah ini digunakan untuk mencari nilai probabilitas class terhadap *data training*, class yang dimaksud adalah kategori klasifikasi peserta didik yang sudah masuk kedalam *data training*. Dataset yang didapat berjumlah 171 data. Data peserta didik dengan kategori bertahan berjumlah 134 data dan peserta didik dengan ketegori mengundurkan diri berjumlah 37 data.

Nilai $\mathrm{P}(v_j)$ yang didapat adalah :

$$\mathrm{P}(v_j = \text{bertahan}) = \frac{134}{171} = 0{,}7836$$

$$\mathrm{P}(v_j = \text{mengundurkan diri}) = \frac{37}{171} = 0{,}2164$$

2) Mencari nilai Probabilitas $\mathrm{P}(a_i \mid v_j)$ dari data diskrit (kategori) menggunakan *m-estimate.*

Langkah ini digunakan untuk mencari nilai probabilitas dari *data testing* terhadap *data training* untuk masing - masing variabel yang ada. Pada penelitian ini dalam mengestimasi probabilitas menggunakan rumus, yang disebut dengan *m-estimate.* Dimana,

$$\mathrm{P}(a_i \mid v_j) = \frac{n_c + mp}{n + m} \quad \text{.....(2)}$$

n adalah jumlah data training untuk class bertahan atau mengundurkan diri ($\upsilon = \upsilon_j$), n_c adalah jumlah data training untuk class bertahan atau mengundurkan diri ($\upsilon = \upsilon_j$) berdasarkan variabel ($a = a_i$), p adalah *prior estimate* untuk $\mathrm{P}(a, i \backslash v_j)$ atau merupakan besar kemungkinan suatu atribut untuk dipilih dalam suatu variabel, dan m adalah ukuran sampel ekuivalen. Nilai m bisa diberi nilai sembarang, misalnya m = 6 tetapi konsisten untuk semua variabel.

Nilai $\mathrm{P}(a_i \mid v_j)$ menggunakan *m-estimate* didapat:

$$\mathrm{P}(a_1 = \text{dekat} \mid v_j = \text{bertahan}) = \frac{101 + 6 * 0.33}{134 + 6} = 0{,}7357$$

$$\mathrm{P}(a_1 = \text{dekat} \mid v_j = \text{mengundurkan diri}) = \frac{15 + 6 * 0.33}{37 + 6} = 0{,}3953$$

$$\mathrm{P}(a_3 = \text{pelajar} \mid v_j = \text{bertahan}) = \frac{117 + 6 * 0.25}{134 + 6} = 0{,}8464$$

$$\mathrm{P}(a_3 = \text{pelajar} \mid v_j = \text{mengundurkan diri}) = \frac{35 + 6 * 0.25}{37 + 6} = 0{,}8488$$

$$\mathrm{P}(a_4 = \text{TK} \mid v_j = \text{bertahan}) = \frac{34 + 6 * 0.2}{134 + 6} = 0{,}2514$$

$$\mathrm{P}(a_4 = \text{TK} \mid v_j = \text{mengundurkan diri}) = \frac{4 + 6 * 0.2}{37 + 6} = 0{,}1209$$

$P(a_5 = P \mid v_j = \text{bertahan}) = \frac{77 + 6 * 0.5}{134 + 6} = 0{,}5714$

$P(a_5 = P \mid v_j = \text{mengundurkan diri}) = \frac{21 + 6 * 0.5}{37 + 6} = 0{,}5581$

$P(a_6 = EFC \mid v_j = \text{bertahan}) = \frac{30 + 6 * 0.25}{134 + 6} = 0{,}2250$

$P(a_6 = EFC \mid v_j = \text{mengundurkan diri}) = \frac{5 + 6 * 0.25}{37 + 6} = 0{,}1512$

3) Mencari nilai Probabilitas $P(a_i \mid v_j)$ dari data kontinu (numerik) menggunakan *Distribusi Gaussian.*

Langkah ini digunakan untuk mencari nilai probabilitas dari *data testing* terhadap *data training* untuk variabel usia. *Distribusi Gaussia* biasanya dipilih untuk merepresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas $P(a_i \mid v_j)$. Dikarakteristikkan dengan dua parameter yakni mean (μ) dan standart deviasi (σ).

Dimana,

$$P(a_i \mid v_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp^{-\frac{(x_i - \mu_{ij})^2}{2\,\sigma_{ij}^2}} \quad .....(3)$$

π adalah nilai Pi yang bernilai 3.14, σ_{ij} adalah standar deviasi *data training* untuk class bertahan atau mengundurkan diri ($v = v_j$) berdasarkan variabel ($a = a_i$), x adalah nilai dari variabel pada inputan tertentu dan exp bernilai 2.718282. Berikut adalah nilai σ_{ij} dan μ_{ij} :

$$\mu_{ij} = \frac{(x_1 + x_2 + \ldots + x_n)}{n} \quad .....(4)$$

$$\mu_{iBertahan} = \frac{(17 + 15 + \ldots + x_n)}{134} = 15.6642$$

$$\mu_{iMengundurkan\ D} = \frac{(17 + 19 + \ldots + x_n)}{37} = 16.5405$$

$$\sigma_{ij} = \sqrt{\frac{\Sigma(x_i - \mu)}{n - 1}} \quad .....(5)$$

$$\sigma_{iBertahan} = \sqrt{\frac{(17 - 15.6642)^2 + (15 - 15.6642)^2 + \ldots + (x_n - 15.6642)^2}{134 - 1}} = 5.2787$$

$$\sigma_{iMengundurkan\ d} = \sqrt{\frac{(17 - 16.5405)^2 + (19 - 16.5405)^2 + \ldots + (x_n - 16.5405)}{37 - 1}} = 3.0188$$

Nilai $P(a_i \mid v_j)$ menggunakan *Distribusi Gaussian* didapat :

$$P(a_2 = 11 \mid v_j = \text{bertahan}) = \frac{1}{\sqrt{2\pi}\sigma_{iBertahan}} \exp^{-\frac{(x_i - \mu_{iBertahan})^2}{2\,\sigma_{iBertahan}^2}} = \frac{1}{\sqrt{2 * 3.14} * 5.2787} 2.718282^{-\frac{(11 - 15.6}{2 * 5.27}} = 0.0512$$

$$P(a_2 = 11 \mid v_j = \text{mengundurkan}) = \frac{1}{\sqrt{2\pi}\sigma_{iMengundurkan\ Diri}} \exp^{-\frac{(x_i - \mu_{ij}}{2\,\sigma_{iMengundurk}^2}} = \frac{1}{\sqrt{2 * 3.14} * 3.0188} 2.718282^{-\frac{(11 - 16.5}{2 * 3.01}} = 0.0245$$

4) Menentukan nilai v_{NB} adalah nilai hasil klasifikasi pada data uji mengunakan pendekatan yang dipakai dalam klasifier *Naive Bayes*.

$$v_{NB} = \arg\max_{v_j \in V} P(v_j) \prod P(a_i \mid v_j)$$

$$= \arg\max_{v_j \in [\text{bertahan, mengundurkan diri}]} P(v_j = \text{bertahan})$$

$$P(a_1 \mid v_j) P(a_2 \mid v_j) P(a_3 \mid v_j) P(a_4 \mid v_j) P(a_5 \mid v_j) P(a_6$$

$$= 0{,}7836 * 0{,}7357 * 0.0512 * 0{,}8464 * 0{,}2514 * 0{,}5714 * 0{,}2250$$

$$= 0.0008071$$

$$v_{NB} = \arg\max_{v_j \in V} P(v_j) \prod P(a_i \mid v_j)$$

$$= \arg\max_{v_j \in [\text{bertahan}, \text{mengundurkan diri}]} P(v_j = \text{mengundurkan diri})$$

$$P(a_1 \mid v_j) P(a_2 \mid v_j) P(a_3 \mid v_j) P(a_4 \mid v_j) P(a_5 \mid v_j) P(a_{}$$

$$= 0{,}2164 * 0.3953 * 0.0245 * 0{,}8488 * 0{,}1209 * 0{,}5581 * 0{,}1512$$

$$= 0.0000182$$

5) Normalisasi nilai probabilitas.

Dengan normalisasi, jumlah probabilitas akan sama dengan 1, menghitung probabilitas konditional untuk pilihan bertahan dan mengundurkan diri jika diberikan nilai-nilai atribut. Untuk contoh ini, probabilitasnya adalah

$$\text{Bertahan} = \frac{0.0008071}{0.0008071 + 0.0000182} = 0.9780$$

$$\text{Mengundurkan Diri} = \frac{0.0000182}{0.0008071 + 0.0000182} = 0.0220$$

Mengkasifikasikan apakah peserta didik termasuk dalam kategori bertahan atau mengundurkan diri, dapat dengan melihat pada nilai akhir probabilitas yang hampir mendekati nilai 1 atau bernilai sama dengan 1. Dari hasil perhitungan normalisasi diketahui bahwa hasil akhir yang diperoleh untuk nilai akhir probabilitas bertahan = 0.9780, sedangkan untuk nilai akhir probabilitas mengundurkan diri = 0.0220. Sehingga peserta didik yang bernama Adinda Putri Adhim, jarak tempuh antara rumah dan tempat kursus dekat (< 4 km), usia 11 tahun, pekerjaannya sebagai pelajar serta pendidikan terakhir pada saat mendaftar adalah TK, berjenis kelamin perempuan dan program kursus yang diambil adalah EFC diprediksi sebagai peserta didik yang berkemungkinan bertahan hingga program kursus berakhir.

III. HASIL DAN KESIMPULAN

A. Hasil Implementasi Sistem

Hasil implementasi sistem dari penelitian ini sebagai berikut :

a. Form Input Data Testing

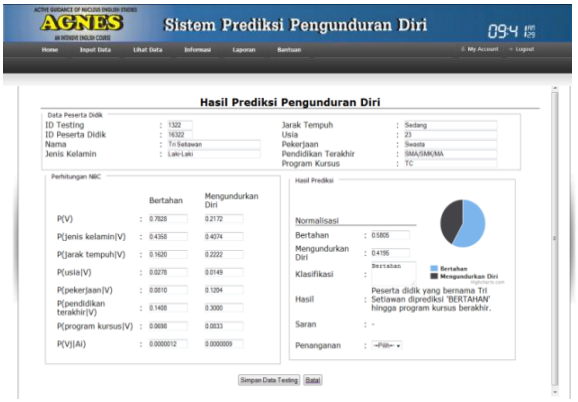
Form ini berfungsi untuk menginputkan *data testing* (uji) peserta didik. Form ini berisikan data id, nama, jenis kelamin, jarak tempuh, usia, pekerjaan, pendidikan terakhir dan program kursus. Tampilan *input* untuk form *input* data peserta didik dapat seperti Gambar 1.

Gambar 1. Form *Input Data Testing*

b. Hasil Prediksi

Halaman hasil prediksi ini bertujuan untuk menampilkan informasi hasil prediksi pengunduran diri dari *data testing*

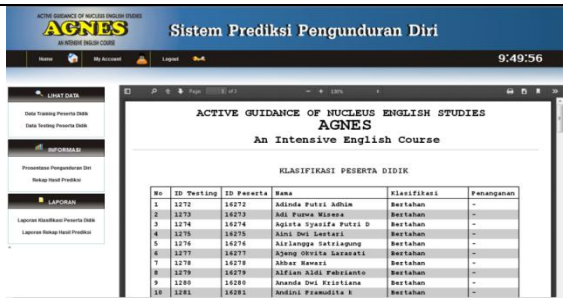
(uji) peserta didik yang telah diinputkan. Peserta didik akan di klasifikasikan ke dalam klasifikasi bertahan atau mengundurkan diri mengunakan metode *Naive Bayes Classifier*. Informasi yang ditampilkan pada halaman ini terdiri dari *data testing* (uji) peserta didik, nilai probabilitas, nilai normalisasi, grafik prosentase, hasil klasifikasi dan saran penanganan. Pada halaman hasil prediksi peserta didik terdapat fasilitas untuk simpan data dan batal untuk mengakhiri apabila data tidak diinginkan untuk disimpan. Gambar 2 berikut adalah tampilan dari halaman hasil prediksi.



Gambar 2. Halaman Hasil Prediksi

c. Laporan Klasifikasi

Halaman ini bertujuan untuk menampilkan laporan hasil klasifikasi data testing peserta didik. Laporan disajikan dalam bentuk tabel, file dapat diunduh dalam format .pdf. Gambar 3 berikut adalah tampilan dari laporan klasifikasi peserta didik.



Gambar 3. Laporan Klasifikasi

B. Kesimpulan

Kesimpulan yang dapat diambil berdasarkan hasil penelitian dan pembahasan yang telah dilakukan bahwa dari data peserta didik dapat digunakan untuk melakukan prediksi pengunduran diri peserta didik berdasarkan variabel yang telah ditentukan. Serta, metode *Naive Bayes Classifier* dapat diimplentasikan kedalam sistem prediksi pengunduran diri peserta didik untuk memberikan klasifikasi peserta didik dalam kategori bertahan atau mengundurkan diri. Dengan metode ini pihak lembaga kursus dapat memecahkan permasalahan pengunduran diri berdasarkan hasil prediksi pengunduran diri dari data peserta didik yang dimasukkan kedalam sistem.

Hasil pengujian tehadap 100 data testing yang sudah dilakukan pada sistem prediksi pengunduran diri dengan menggunakan metode *Naive bayes Classifier* menghasilkan 77 data peserta didik atau sebesar 77 % terklasifikasi bertahan dan 23 data atau sebesar 23 %

terklasifikasi mengundurkan diri. Diharapkan berdasarkan hasil prediksi pengunduran diri peserta didik, pihak lembaga kursus dapat melakukan tindakan-tindakan yang diperlukan dalam mempertahankan peserta didik serta meminimalkan kemungkinan kerugian yang timbul akibat pengunduran diri peserta didik.

IV. DAFTAR PUSTAKA

Kusrini & Sri. 2007. *Penggunaan Pohon Keputusan untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa di STMIK AMIKOM Yogyakarta*. Makalah disajikan dalam Seminar Nasional Teknologi, STMIK AMIKOM Yogyakarta, Yogyakarta, 24 November.

Ndaumanu, R.I., Kusrini. & Arief, M.R. 2014. Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa dengan Metode K-Nearest Neighbor. *Jatisi*, 1 (1): 1-15.

Pakereng, I. & Wahyono, T. 2004. *Sistem Basis Data: Konsep dan Pendekatan Praktikum*. Yogyakarta: Graha Ilmu.

Prasetyo, E. 2012. *Data Mining: Konsep Dan Aplikasi Menggunakan MATLAB*. Yogyakarta: Andi.

Ridwan, M., Suyono, H. & Sarosa, M. 2013. Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Jurnal EECCIS*, 7 (1): 59-64.

Santoso, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.

Simarmata, J. & Paryudi, I. 2006. *Basis Data*. Yogyakarta: Andi.

Sutabri, T. 2004. *Pemograman Terstruktur*. Yogyakarta: Andi.

Tacbir, Faiza & Age. 2011. *Penerapan Data Mining Untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa Baru*. Makalah disajikan dalam Konferensi Nasional Sistem dan Informatika, Universitas Jenderal Achmad Yani Cimahi, Bali, 12 November.

Wasiati, H. & Wijayanti, D. 2014. Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes (Studi Kasus: Di P.T. Karyatama Mitra Sejati Yogyakarta). *Indonesian Journal on Networking and Security*, 3 (2): 45-51.