

Penerapan *Naïve Bayes* untuk Memprediksi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademis

Gustientiedina¹, Muhammad Siddik², Yenny Desnelita³

^{1,2,3} Teknik Informatika, Sekolah Tinggi Ilmu Komputer Pelita Indonesia

¹gustientiedina@lecturer.pelitaindonesia.ac.id,

²siddik@lecturer.pelitaindonesia.ac.id, ³yenny.desnelita@lecturer.pelitaindonesia.ac.id

Abstrak— Penelitian ini bertujuan untuk menerapkan data mining menggunakan metode *Naïve Bayes* untuk mengukur kualitas pelayanan akademis pada beberapa perguruan tinggi di Pekanbaru. Target khusus yang ingin dicapai adalah untuk memperoleh informasi penting secara kuantitatif dalam pengukuran kualitas pelayanan akademis di perguruan tinggi sebagai bahan pertimbangan dalam pengambilan keputusan dan mengembangkan sistem yang dapat membantu masyarakat perguruan tinggi dalam mengambil kebijakan yang berhubungan dengan peningkatan kualitas pelayanan akademis. *Naïve Bayes* merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema atau aturan bayes dengan asumsi independensi yang kuat pada fitur. Indikator penilaian yang digunakan adalah *tangible* (bukti langsung), *reability* (keandalan), *responsiveness* (daya tanggap), *assurance* (jaminan), dan *empathy* (empati). Penerapan *Naïve Bayes* ini diharapkan mampu untuk memprediksi tingkat kepuasan mahasiswa terhadap kualitas pelayanan akademis pada perguruan tinggi. Berdasarkan hasil klasifikasi dengan menggunakan metode *Naïve Bayes* menunjukkan bahwa *Naïve Bayes* cocok digunakan untuk mengukur tingkat kepuasan mahasiswa terhadap pelayanan akademis pada perguruan tinggi dengan tingkat akurasi sebesar 96,71% dengan nilai *precision* sebesar 96,15% dan nilai *recallnya* sebesar 98,43%.

Kata kunci— *Naïve Bayes*, kepuasan mahasiswa, pelayanan akademis.

Abstract— This study to implement data mining using the *Naïve Bayes* method to measure the quality of academic services at several university in Pekanbaru. The specific target to be achieved is to obtain important information quantitatively in measuring the quality of academic services in tertiary institutions as a material for consideration in decision making and developing systems that can assist tertiary institutions in taking policies related to improving the quality of academic services. *Naïve Bayes* is a simple probabilistic based prediction technique based on the application of the Bayes theorem or rules with a strong assumption of independence on features. The assessment indicators used are tangible (direct evidence), reliability (reliability), responsiveness (responsiveness), assurance (assurance), and empathy (empathy). The application of *Naïve Bayes* is expected to be able to predict the level of student satisfaction with the quality of academic services at tertiary institutions. Based on the results of the classification using the *Naïve Bayes* method shows that *Naïve Bayes* is suitable for measuring the level of student satisfaction with academic services at higher education with an accuracy rate of 96.71% with a precision value of 96.15% and a recall value of 98.43%..

Keywords— *Naïve Bayes*, student satisfaction, academic service.

I. PENDAHULUAN

Pelayanan akademis adalah usaha yang dilakukan oleh perguruan tinggi untuk memberikan kemudahan pada mahasiswa dalam hal yang berkaitan dengan kegiatan akademis. Kualitas pelayanan akademis institusi atau perguruan tinggi sangat dipengaruhi oleh masukan bagi pelayanan akademis diantaranya adalah mahasiswa, dosen, fasilitas sarana dan prasarana. Ketiga faktor tersebut saling tergantung dan mempengaruhi satu sama lain dalam penyelenggara perguruan tinggi.

Data mining merupakan suatu metode untuk menemukan pengetahuan dalam suatu tumpukan data yang cukup besar. *Data mining* adalah proses menggali dan menganalisa sejumlah data yang sangat besar untuk memperoleh sesuatu yang benar, baru dan bermanfaat dan akhirnya dapat ditemukan suatu corak atau pola dalam data tersebut [1]. *Data mining* merupakan proses dengan menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang besar [2]. *Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu : Deskripsi, Estimasi, Prediksi, Klasifikasi, *Clustering*, dan Asosiasi [3].

Naïve Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan

teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [4]. *Naïve Bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada database dengan *bigdata* [5].

Terdapat beberapa penelitian yang telah dilakukan sebelumnya yang berkaitan dengan metode *naïve bayes*, seperti Alfa Saleh dalam penelitiannya menerapkan metode klasifikasi *naïve bayes* dalam memprediksi besarnya penggunaan listrik rumah tangga. Dari 60 data penggunaan listrik rumah tangga yang diuji dengan metode *naïve bayes*, maka diperoleh hasil persentase 78,3333% untuk keakuratan prediksi, di mana dari 60 data penggunaan listrik rumah tangga yang diuji terdapat 47 data penggunaan listrik rumah tangga yang berhasil diklasifikasikan dengan benar [6]. Hasil penelitian Haditsah Annur tentang klasifikasi masyarakat miskin menggunakan metode *naïve bayes*, diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori *Good*. Sementara nilai *Precision* sebesar 92% dan *Recall* sebesar 86% [7]. Penelitian Fuzy Yustika Manik dan Kana Saputra Saragih tentang klasifikasi belimbing menggunakan *naïve bayes* berdasarkan fitur warna RGB, didapatkan hasil klasifikasi menunjukkan akurasi menggunakan ekstraksi ciri RGB sebesar 80% [8].

Berdasarkan uraian diatas maka peneliti merasa tertarik untuk melakukan penelitian tentang prediksi tingkat kepuasan mahasiswa terhadap sistem pelayanan akademis pada perguruan tinggi menggunakan metode *naïve bayes*,

diharapkan dapat meningkatkan akurasi penilaian terhadap pelayanan akademis sehingga Perguruan Tinggi mempunyai suatu standar dalam memberikan penilaian terhadap pelayanan akademis.

II. METODOLOGI PENELITIAN

A. Pengumpulan Data

Data primer diperoleh dengan melakukan beberapa teknik pengumpulan data yaitu dengan menyebarkan kuesioner kepada mahasiswa semester genap tahun ajaran 2018/2019 STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru, pengamatan langsung, dan wawancara, yang digunakan sebagai acuan penelitian dalam menganalisa metode *naïve bayes* untuk memprediksi kepuasan mahasiswa terhadap pelayanan akademis di STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru.

Sedangkan data sekunder merupakan data relevan yang melengkapi data primer sehingga dapat memaksimalkan hasil penelitian, diperoleh dari berbagai sumber yang menunjang penjelasan penelitian, yaitu dengan menggunakan studi kepustakaan.

B. Pengolahan Data Awal

Data survey dalam bentuk kuesioner yang dibagikan kepada mahasiswa semester genap tahun ajaran 2018/2019 STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru yang terdiri dari beberapa aspek atau faktor penilaian, kemudian dipersempit menjadi 5 (lima) indikator sebagai atribut yaitu *tangible* (bukti langsung), *reability* (keandalan), *responsiveness* (daya tanggap), *assurance* (jaminan), *empathy* (empati).

C. Model yang Diusulkan

Metode yang diusulkan untuk memprediksi kepuasan mahasiswa terhadap pelayanan akademis di STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru adalah dengan menggunakan metode *naïve bayes*. Persamaan dari teorema Bayes adalah [9]:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (1)$$

Keterangan :

- B : Data dengan kelas yang belum diketahui.
- A : Hipotesa data B merupakan suatu kelas spesifik.
- P(A) : Probabilitas hipotesa A (*prior probability* / probabilitas awal).
- P(B) : Probabilitas B.
- P(B|A) : Probabilitas hipotesa B berdasarkan kondisi A.
- P(A|B) : Probabilitas hipotesa A berdasarkan kondisi B (*Posterior probability* / probabilitas akhir).

Kaitan pada *naïve bayes* dan klasifikasi, korelasi hipotesis dan bukti klasifikasi bahwa hipotesis dalam teorema adalah label kelas yang menjadi tujuan dari pemetaan. Dalam *naïve bayes classifier* maksud independensi yang kuat pada atribut adalah sebuah atribut pada sebuah data tidak berkaitan dengan ada atau tidaknya atribut lain dalam data yang sama. Independensi (tidak ketergantungan), dengan kata lain *naïve bayes classifier* model yang digunakan adalah model atribut independensi [10]. Konsep probabilitas merupakan salah satu bentuk model statistik. Salah satu metode yang menggunakan konsep

probabilitas adalah *naïve bayes classification* (NBC). Pada metode ini, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain. Apabila diberikan k atribut yang saling bebas (*independence*), nilai probabilitas dapat diberikan sebagai berikut. Klasifikasi *naïve bayes classifier* yang memiliki nilai atribut yang terpisah atau independen ini dapat dinyatakan dengan rumus sebagai berikut:

$$P(X_1, \dots, X_k|C) = P(X_1|C) \times P(X_2|C) \times \dots \times P(X_k|C) \quad (2)$$

Keterangan :

- $P(X_1, \dots, X_k|C)$ = Probabilitas kelas C menghasilkan intence X
- $P(X_1|C)$ = Probabilitas kelas C menghasilkan nilai yang diamati untuk atribut ke-1 (X_1)
- $P(X_2|C)$ = Probabilitas kelas C menghasilkan nilai yang diamati untuk atribut ke-1 (X_2)

Untuk menghitung nilai akhir kelas menggunakan rumus:

$$C_M = \text{argmax}_C P(X|C) \quad (3)$$

Keterangan :

- C_{MAP} = Hipotesa nilai tertinggi
- $\text{argmax}_{C \in C}$ = Nilai rata-rata dari setiap kelas

D. Pengujian Model

Model yang telah didapatkan dan dikembangkan dalam penelitian ini akan diterapkan pada data kepuasan mahasiswa terhadap pelayanan akademis di STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru melalui suatu simulasi menggunakan Rapidminer 5.3. Dataset terdiri dari atribut *tangible* (bukti langsung), *reability*. (keandalan), *responsiveness* (daya tanggap), *assurance* (jaminan), *empathy* (empati) dan keterangan sebagai label.

E. Evaluasi dan Validasi Hasil

Evaluasi dilakukan dengan menganalisa hasil klasifikasi. *Confusion Matrix* adalah sebuah metode yang digunakan untuk menganalisa seberapa baik sebuah model klasifikasi yang dibangun dalam mengenali data dengan kelas-kelas yang berbeda. Dengan performa dari suatu model klasifikasi dapat diukur dengan tingkat akurasi. Akurasi yang dihasilkan dari sebuah klasifikasi yang memberikan hasil latih dengan bentuk persentase dari kelompok data latih yang telah diklasifikasikan dengan benar, dengan perhitungan [1]:

TABEL I
CONFUSION MATRIX

	Predicted Class		
	Yes	No	
Actual Class	Yes	TP	FN
	No	FP	TN

Keterangan :

Actual class (yes): kelas sebenarnya yang bernilai *yes*.
Actual class (no): kelas sebenarnya yang bernilai *no*.
Predicted class (yes): prediksi kelas yang bernilai *yes*.
Predicted class (no): prediksi kelas yang bernilai *no*.

TP (*True Positive*) : jumlah data yang diklasifikasikan dengan benar (*Actual class (yes)* = *Predicted class (yes)*).

FN (*False Negative*): jumlah data yang diklasifikasikan dengan salah (*Actual class (yes)* = *Predicted class (no)*).

FP (*False Positive*): jumlah data yang diklasifikasikan dengan salah (*Actual class (no)* = *Predicted class (yes)*).

TN (*True Negative*): jumlah data yang diklasifikasikan dengan benar (*Actual class (no)* = *Predicted class (no)*).

TP dan TN menunjukkan bahwa model klasifikasi melakukan klasifikasi dengan benar, FP dan FN menunjukkan bahwa model klasifikasi melakukan klasifikasi dengan salah. Sehingga untuk menghitung akurasi dari model yang dibangun dengan persamaan berikut:

$$Akurasi = \frac{TP+TN}{TP+FP+TN+} \times 100\% \quad (4)$$

Precision didefinisikan sebagai rasio item relevan yang dipilih terhadap semua item yang terpilih. *Precision* merupakan probabilitas bahwa sebuah item yang dipilih adalah relevan. Dapat diartikan sebagai kecocokan antara permintaan informasi dengan jawaban terhadap permintaan itu. *Precision* dapat dihitung menggunakan persamaan berikut ini:

$$Precision = \frac{TI}{TP+} \times 100\% \quad (5)$$

Recall didefinisikan sebagai rasio dari item relevan yang dipilih terhadap total jumlah item relevan yang tersedia. *Recall* merupakan probabilitas bahwa suatu item yang relevan akan dipilih. *Recall* dapat dihitung dengan jumlah rekomendasi yang relevan yang dipilih oleh *user* dibagi dengan jumlah semua rekomendasi yang relevan baik dipilih maupun rekomendasi yang tidak terpilih. *Recall* dapat dihitung menggunakan persamaan berikut ini:

$$Recall = \frac{T}{TP+} \times 100\% \quad (6)$$

Sistem rekomendasi akan dianggap baik jika nilai *precision* dan *recallnya* tinggi.

Evaluasi pengukuran Rapidminer yaitu untuk menemukan nilai *akurasi*, nilai *precision*, dan nilai *recall* dari metode *naïve bayes*.

Validasi hasil penelitian dilakukan dengan mengambil sampel secara acak sebanyak 213 dari data kuesioner kepuasan mahasiswa terhadap pelayanan akademis di STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru, kemudian data diuji dengan metode *naïve bayes*.

III. HASIL DAN PEMBAHASAN

F. Pengolahan Data

Gambar 1. Data Awal

Data yang dikumpulkan adalah data kuesioner yang dibagikan kepada mahasiswa semester genap hun ajaran 2018/2019 STIKOM Pelita Indonesia Pekanbaru dan STIE Pelita Indonesia Pekanbaru.

Agar penelitian ini mendapatkan hasil yang diharapkan maka penulis mengelompokan data sampel per parameter menjadi dua interval nilai seperti terlihat pada Tabel II.

TABEL II
KELOMPOK NILAI KUISONER PER PARAMETRE

No	Interval Nilai	Kelompok
1	1 - 4	Rendah
2	5 - 7	Tinggi

Mengklasifikasikan hasil menjadi dua interval nilai seperti terlihat pada Tabel III.

TABEL III
KLASIFIKASI HASIL NILAI KUISONER

No	Interval Nilai	Klasifikasi
1	1 – 4	Tidak Puas
2	5 – 7	Puas

Data awal di olah dengan Microsoft Excel sehingga menghasilkan data set seperti terlihat pada gambar 2.

M	Tangibles	Reliability	Responsiveness	Assurance	Empathy	Hasil
1	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
2	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
3	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
4	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
5	Rendah	Tinggi	Rendah	Tinggi	Tinggi	Puas
6	Rendah	Rendah	Rendah	Rendah	Tinggi	Tidak Puas
7	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
8	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
9	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
10	Tinggi	Tinggi	Rendah	Tinggi	Rendah	Tidak Puas
11	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
12	Tinggi	Tinggi	Rendah	Tinggi	Rendah	Puas
13	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
14	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
15	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
16	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
17	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
18	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
19	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
20	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas

Gambar 2. Data Set

G. Proses Data Mining pada RapidMiner

Selanjutnya data akan diproses menggunakan *RapidMiner* dan untuk pengujiannya menggunakan *Split Validation* yang terdapat pada *RapidMiner*. Gambar 3 adalah data yang akan di-import ke *RapidMiner*.

M	Tangibles	Reliability	Responsive	Assurance	Empathy	Hasil
text	binomi...	binomi...	binomi...	binomi...	binomi...	binomi...
id	attribute	attribute	attribute	attribute	attribute	label
1	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
2	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
3	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
4	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas
5	Rendah	Tinggi	Rendah	Tinggi	Tinggi	Puas
6	Rendah	Rendah	Rendah	Rendah	Tinggi	Tidak Puas
7	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
8	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
9	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
10	Tinggi	Tinggi	Rendah	Tinggi	Rendah	Tidak Puas
11	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Puas
12	Tinggi	Tinggi	Rendah	Tinggi	Rendah	Puas
13	Tinggi	Tinggi	Rendah	Tinggi	Tinggi	Puas

Gambar 3. Data yang di-Import ke *RapidMiner*

Setelah melakukan perhitungan dengan metode *naïve bayes* pada *RapidMiner*, maka dihasilkan *probabilitas* seperti Gambar 4.

Attribute	Parameter	Puas	Tidak Puas
Tangibles	value=Tingg	0.953	0.128
Tangibles	value=Rend:	0.047	0.872
Tangibles	value=unknc	0.000	0.000
Reliability	value=Tingg	0.968	0.198
Reliability	value=Rend:	0.032	0.802
Reliability	value=unknc	0.000	0.000
Responsive	value=Tingg	0.638	0.070
Responsive	value=Rend:	0.362	0.930
Responsive	value=unknc	0.000	0.000
Assurance	value=Tingg	0.961	0.209
Assurance	value=Rend:	0.039	0.791
Assurance	value=unknc	0.000	0.000
Empathy	value=Tingg	0.898	0.174
Empathy	value=Rend:	0.102	0.826
Empathy	value=unknc	0.000	0.000

Gambar 4. *Probabilitas*

H. Perhitungan Manual mencari *Probabilitas*

- P(Tangibles=Rendah|Kelas=Puas) = 6/127 = 0.0472
- P(Tangibles=Rendah|Kelas=Tidak Puas) = 75/86 = 0.8721
- P(Tangibles=Tinggi|Kelas=Puas) = 121/127 = 0.9528
- P(Tangibles=Tinggi|Kelas=Tidak Puas) = 11/86 = 0.1279
- P(Reliability=Rendah|Kelas=Puas) = 4/127 = 0.0315
- P(Reliability=Rendah|Kelas=Tidak Puas) = 69/86 = 0.8023
- P(Reliability=Tinggi|Kelas=Puas) = 123/127 = 0.9685
- P(Reliability=Tinggi|Kelas=Tidak Puas) = 17/86 = 0.1977
- P(Responsiveness=Rendah|Kelas=Puas) = 46/127 = 0.3622
- P(Responsiveness=Rendah|Kelas=Tidak Puas) = 80/86 = 0.9302
- P(Responsiveness=Tinggi|Kelas=Puas) = 81/127 = 0.6378
- P(Responsiveness=Tinggi|Kelas=Tidak Puas) = 6/86 = 0.0698
- P(Assurance=Rendah|Kelas=Puas) = 5/127 = 0.0394
- P(Assurance=Rendah|Kelas=Tidak Puas) = 68/86 = 0.7907
- P(Assurance=Tinggi|Kelas=Puas) = 122/127 = 0.9606
- P(Assurance=Tinggi|Kelas=Tidak Puas) = 18/86 = 0.2093
- P(Empathy=Rendah|Kelas=Puas) = 13/127 = 0.1024
- P(Empathy=Rendah|Kelas=Tidak Puas) = 71/86 = 0.8256
- P(Empathy=Tinggi|Kelas=Puas) = 114/127 = 0.8976
- P(Empathy=Tinggi|Kelas=Tidak Puas) = 15/86 = 0.1744

I. Pengujian

Dari hasil pengolahan data menggunakan metode *naïve bayes* pada *RapidMiner*, maka dihasilkan *Confusion Matrix* seperti Gambar 5.

	true Puas	true Tidak Puas
pred. Puas	125	5
pred. Tidak Puas	2	81

Gambar 5. *Confusion Matrix*

Dari *Confusion Matrix* diperoleh jumlah *true positive* (TP) sebanyak 125 record, *false positive* (FP) sebanyak 2 record, jumlah *true negative* (TN) sebanyak 81 record, dan *false negative* (FN) sebanyak 5 record.

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \\
 &= \frac{125 + 81}{125 + 2 + 81 + 5} \times 100\% \\
 &= \frac{206}{213} \times 100\% \\
 &= 96,71\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FN} \times 100\% \\
 &= \frac{125}{125 + 5} \times 100\% \\
 &= \frac{125}{130} \times 100\% \\
 &= 96,15\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP + FP} \times 100\% \\
 &= \frac{125}{125 + 2} \times 100\% \\
 &= \frac{125}{127} \times 100\% \\
 &= 98,43\%
 \end{aligned}$$

IV. KESIMPULAN

Dari hasil pengujian akhir yang dilakukan dari data sampel kuesioner dengan atribut atau indikator *tangible*, *reability*, *responsiveness*, *assurance*, dan *empathy* menggunakan metode *naïve bayes* didapatkan tingkat akurasi sebesar 96,71% dengan nilai *precision* sebesar 96,15% dan nilai *recallnya* sebesar 98,43%.

Berdasarkan hasil pengujian, maka metode *naïve bayes* bisa direkomendasikan untuk memprediksi tingkat kepuasan mahasiswa terhadap pelayanan akademis pada perguruan tinggi, karena nilai *precision* dan *recallnya* tinggi.

Penelitian ini bisa dikembangkan dengan menggabungkan atau membandingkan dengan algoritma klasifikasi lain seperti

algoritma *C4.5*, *K-Nearest Neighbor*, *Random Forest*, atau *Support Vector Machine* untuk memperoleh hasil prediksi yang lebih baik

UCAPAN TERIMAKASIH

Ucapan terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat Direktorat Jendral Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia telah mendanai penelitian ini dan Sekolah Tinggi Ilmu Komputer Pelita Indonesia.

REFERENSI

- [1] J. Han, J., Kamber, M., Pei, *Data Mining Concept And Techniques*, III ed. Amerika: Library of Congress Cataloging, 2012.
- [2] T. L. Kusriani, Emha, *Algoritma Data Mining*. Yogyakarta: Andi Offset, 2009.
- [3] Santosa, B, *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu, 2007.
- [4] Patil, T. R., Sherekar, M. S., "Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification," *International Journal of Computer Science and Applications*, Vol. 6, No. 2, Hal. 256-261, 2013.
- [5] Wati, R., "Penerapan Algoritma Genetika Untuk Seleksi Fitur Pada Analisis Sentimen Review Jasa Maskapai Penerbangan Menggunakan Naïve Bayes," *Jurnal Evolusi*, Vol. 4, No. 1, Hal. 25-31, ISSN: 2338-8161, 2016.
- [6] Saleh, A., "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Citec Journal*, Vol. 2, No. 3, Hal. 207-217, ISSN: 2354-5771, 2015.
- [7] Annur, H., "Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes," *ILKOM Jurnal Ilmiah*, Vol. 10, No. 2, Hal. 160-165, ISSN: 2548-7779, 2018.
- [8] Manik, F. Y., saragih, K. S., "Klasifikasi Belimbing Menggunakan Naïve Bayes Berdasarkan Fitur Warna RGB," *IJCCS*, Vol. 11, No. 1, Hal. 99-108, ISSN: 1978-1520, 2017.
- [9] Septiani, W. D., "Komperasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis," *Jurnal Pilar Nusa Mandiri*, Vol. 13, No. 1, Hal. 76-84, 2017.
- [10] Prasetio, E., *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi, 2012.