



Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS)

Ade Nur Hidayatulloh^{1*}, Ilham², Muhammad Bima Win Pane³, Tubagus Muhammad Fahri Amriyahya⁴, Aries Saifudin⁵

^{1,2,3,4,5}Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

E-mail: ¹adenurhidayatulloh812@gmail.com, ²ilhamahli23@gmail.com, ³m.bimawinpane1409@gmail.com,
⁴tbahry@gmail.com, ⁵aries.saifudin@unpam.ac.id

(* : coressponding author)

Abstrak - Penelitian ini mengkaji prediksi tingkat kelulusan peserta uji sertifikasi Microsoft Office Specialist versi 2013 (Word dan Excel). Tujuannya adalah untuk memberikan informasi terkait tingkat kelulusan peserta dan menawarkan solusi alternatif dalam menentukan program sertifikasi yang tepat sebelum melakukan uji sertifikasi. Metode Naïve Bayes digunakan untuk klasifikasi kelulusan sertifikasi, memungkinkan peserta mengetahui status kelulusan atau ketidaklulusan mereka. Naïve Bayes merupakan metode klasifikasi yang mengandalkan probabilitas dan statistik untuk memprediksi peluang di masa depan berdasarkan pengalaman sebelumnya. Dalam penelitian ini, pengembangan sistem dilakukan dengan metode CRISP-DM untuk memastikan pengerjaan yang terstruktur. Pengujian dilakukan dengan metode blackbox untuk memeriksa setiap fungsi yang ada dalam aplikasi yang dibangun. Penelitian ini menghasilkan nilai probabilitas sebesar 0,001042 dengan tingkat akurasi mencapai 99%. Hasil ini menunjukkan bahwa metode Naïve Bayes dapat digunakan untuk memprediksi tingkat kelulusan peserta sertifikasi (Word dan Excel) dengan tingkat keakuratan yang tinggi, sehingga peserta dapat menentukan program sertifikasi yang tepat sebelum mengikuti uji sertifikasi.

Kata Kunci: Prediksi, Sertifikasi, Naïve Bayes, CRISP-DM, BlackBox

Abstract - *This study focuses on predicting pass rates for the Microsoft Office Specialist 2013 certification, specifically for Word and Excel. The aim is to provide insights into the pass rates and offer alternative solutions for selecting the appropriate certification program before taking the test. The Naive Bayes algorithm is employed to classify certification outcomes, enabling participants to understand their likelihood of passing or failing. Naive Bayes, a probabilistic classifier based on probability and statistics, predicts future outcomes based on prior data. In this research, the CRISP-DM methodology was utilized to ensure a structured system development process, and BlackBox testing was conducted to evaluate each function within the developed application. The findings indicate a probability value of 0.001042 with an accuracy of 99%. These results demonstrate that the Naive Bayes method can effectively predict certification pass rates (for Word and Excel) with high accuracy, aiding participants in selecting the most suitable certification program before taking the test.*

Keywords: Prediction, Certification, Naïve Bayes, CRISP-DM, BlackBox

1. PENDAHULUAN

Perkembangan teknologi saat ini sangat pesat, hampir semua bidang menggunakan teknologi berbasis komputer. Oleh karena itu, para profesional yang bekerja di bidang teknologi informasi harus meningkatkan profesionalisme mereka agar profesi tersebut tidak diambil alih oleh individu yang bukan berasal dari latar belakang teknologi informasi. Teknologi informasi terus berkembang baik secara revolusioner, seperti perkembangan perangkat keras, maupun secara evolusioner, seperti yang terjadi pada perangkat lunak.

Sertifikasi dalam bidang teknologi informasi merupakan hal yang sangat penting karena memberikan kredibilitas kepada pemegangnya. Sertifikasi ini menunjukkan bahwa para profesional di bidang teknologi informasi memiliki pengetahuan dan kompetensi yang dapat diandalkan. Selain itu, sertifikasi juga memberikan keunggulan kompetitif bagi perusahaan, khususnya dalam pasar global, karena kemampuan dan pengetahuan profesional teknologi informasi telah diuji dan terdokumentasi. Namun, yang lebih penting dari sertifikasi itu sendiri adalah kemampuan nyata dan pengetahuan dasar yang mendalam mengenai bidang tersebut. Ada banyak alasan untuk memperoleh sertifikasi IT.

Laboratorium ITCC (Information Technology Certification Center) adalah salah satu dari beberapa laboratorium komputer di kampus Sekolah Tinggi Teknik PLN yang bergerak dalam



bidang sertifikasi dan telah bekerjasama dengan vendor berskala internasional seperti Microsoft. ITCC memiliki otorisasi sebagai Pusat Ujian Terotorisasi (Authorized Testing Center) dari Certiport untuk penyelenggaraan pelatihan dan sertifikasi internasional dalam program aplikasi seperti Microsoft Office Specialist (MOS), Microsoft Technology Associate (MTA), Microsoft Certified Educator (MCE), Adobe Certified Associate (ACA), dan Autodesk Certified User (ACU).

Laboratorium ITCC didirikan pada tahun 2016 dan sejak tahun 2017 telah menyelenggarakan uji sertifikasi untuk Microsoft Office Word 2013 dan Microsoft Office Excel 2013. Meski masih tergolong baru, laboratorium ini saat ini hanya melayani peserta sertifikasi dari kalangan internal STT-PLN. Namun, dalam rencana jangka panjang, Laboratorium ITCC berencana untuk membuka layanan sertifikasi bagi peserta dari luar STT-PLN.

Salah satu tantangan dalam pelaksanaan uji sertifikasi adalah ketidakmampuan sistem untuk memprediksi kelulusan peserta. Baik penyelenggara maupun peserta sering kali tidak memiliki pengetahuan yang memadai tentang kriteria yang diperlukan untuk memilih antara sertifikasi Microsoft Office Word 2013 dan Microsoft Office Excel 2013, serta untuk memahami faktor-faktor yang mempengaruhi tingkat kelulusan. Oleh karena itu, diperlukan informasi prediktif untuk meningkatkan peluang kelulusan dalam program sertifikasi yang tepat.

Untuk mengatasi masalah ini, penelitian ini dilakukan dengan menerapkan teknik data mining menggunakan metode klasifikasi dengan algoritma Naïve Bayes. Algoritma ini, yang didasarkan pada teorema probabilitas yang dikembangkan oleh Thomas Bayes, memprediksi peluang di masa depan berdasarkan data historis. Algoritma ini diasumsikan bekerja dengan menganggap bahwa setiap atribut independen satu sama lain.

Penggunaan algoritma Naïve Bayes diharapkan dapat memprediksi tingkat kelulusan di masa depan secara lebih akurat, sehingga membantu peserta dalam memilih program sertifikasi yang tepat. Sistem ini akan memberikan informasi mengenai kemungkinan kelulusan berdasarkan kriteria yang dimiliki peserta, sehingga proses pemilihan program menjadi lebih efektif dan efisien. Diharapkan, penelitian ini akan mengurangi tingkat ketidaklulusan dalam uji sertifikasi dan memberikan panduan yang lebih baik bagi peserta dalam memilih program sertifikasi yang sesuai.

Dengan latar belakang tersebut, penelitian ini berfokus pada prediksi tingkat kelulusan peserta dalam program sertifikasi Microsoft Office Word dan Excel menggunakan algoritma Naïve Bayes di laboratorium ITCC.

2. METODOLOGI

2.1 Analisa Kebutuhan

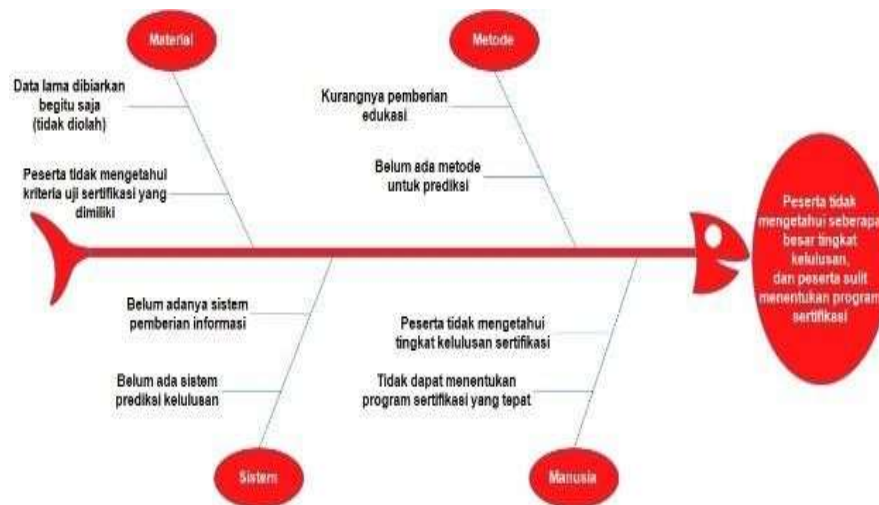
Untuk melakukan analisis kebutuhan, langkah pertama yang diperlukan adalah tahapan awal dari penelitian. Pada tahap ini, dilakukan evaluasi kebutuhan data menggunakan pendekatan CRISP-DM. Proses ini melibatkan serangkaian langkah yang meliputi:

a. *Business Understanding Phase* (Fase Pemahaman Bisnis)

Untuk tahap ini, akan dilakukan identifikasi tujuan proyek serta kebutuhan secara terperinci dalam lingkup bisnis atau unit penelitian secara menyeluruh. Studi ini bertujuan untuk mengembangkan sistem penentuan yang mampu memproyeksikan tingkat kelulusan peserta sertifikasi berdasarkan kriteria uji sertifikasi menggunakan metode Algoritma Naive Bayes. Untuk mencapai sasaran ini, dilakukan analisis awal terlebih dahulu.

1) Analisa Masalah

Untuk menganalisis permasalahan ini, dilakukan pendekatan menggunakan metode Fishbone guna mengidentifikasi berbagai faktor yang mempengaruhi terjadinya masalah. Metode Fishbone, atau lebih dikenal sebagai Diagram Tulang Ikan, digunakan untuk menggambarkan hubungan sebab-akibat dari suatu permasalahan. Berikut adalah diagram Fishbone yang telah disusun oleh peneliti untuk mengilustrasikan faktor-faktor yang relevan:



Gambar 1. Diagram Tulang Ikan

Diagram tulang ikan menggambarkan hubungan sebab-akibat dalam konteks suatu permasalahan. Setiap faktor penyebab diuraikan sebagai berikut:

- Manusia:** Kendala dalam menentukan program sertifikasi yang tepat sering kali timbul karena minimnya pemahaman peserta terhadap kriteria sertifikasi dan ketidakmampuan untuk melakukan prediksi yang akurat. Akibatnya, peserta sering kali memilih program yang tidak sesuai.
- Metode:** Laboratorium belum sepenuhnya menyediakan edukasi yang memadai mengenai pemilihan program sertifikasi yang sesuai. Hal ini disebabkan oleh kurangnya metode yang dapat mendukung proses prediksi yang efektif.
- Sistem:** Hingga saat ini, belum ada sistem yang dapat memprediksi dengan tepat tingkat kelulusan dalam uji sertifikasi, sehingga informasi terkait sertifikasi seringkali kurang tersedia.
- Material:** Data dari uji sertifikasi sebelumnya sering kali tidak dimanfaatkan secara optimal, menyebabkan kurangnya pemahaman peserta terhadap kriteria evaluasi yang digunakan.

Berdasarkan analisis di atas, penulis akan mengembangkan sebuah aplikasi sistem yang dapat meningkatkan efisiensi dalam pengolahan data, sehingga dapat membantu dalam prediksi tingkat kelulusan peserta sertifikasi dengan lebih baik dan efektif.

b. *Data Understanding Phase* (Fase Pemahaman Data)

Untuk tahap ini, dilakukan pengumpulan informasi. Dalam penelitian ini, data yang digunakan merupakan hasil dari uji sertifikasi di laboratorium ITCC pada periode 2017 – 2018 dengan jumlah total 1024 dataset. Berikut ini adalah detail dari data yang berhasil dikumpulkan:

Tabel 1. Data Hasil Uji Sertifikasi

No	Program	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Kriteria 5	Keterangan	Nilai
1	MOS-01	100%	100%	100%	100%	100%	Lulus	1000
2	MOS-01	93%	94%	98%	92%	85%	Lulus	924
3	MOS-01	96%	93%	97%	96%	68%	Lulus	900
4	MOS-01	93%	97%	99%	95%	61%	Lulus	890
5	MOS-01	93%	90%	94%	83%	91%	Lulus	902



..
..
..
..
1022	MOS-01	58%	58%	57%	61%	84%	Tidak	637
1023	MOS-01	57%	63%	55%	61%	71%	Tidak	614
1024	MOS-01	56%	58%	61%	64%	61%	Tidak	600

c. *Data Preparation Phase* (Fase Pengolahan Data)

Pada tahap pengelolaan data, dilakukan seleksi dan transformasi data. Data yang telah digabungkan akan diolah pada proses pengelolaan data ini.

1) *Data Selection* (Seleksi Data)

Untuk mengevaluasi data yang akan digunakan, seleksi data diperlukan karena tidak semua data relevan untuk analisis yang dilakukan. Hanya data yang tepat yang diambil untuk mendukung kevalidan analisis tersebut. Peneliti memilih dengan cermat data yang relevan yang sesuai dengan kebutuhan studi ini sebagai berikut:

Tabel 2. Data Selection

No	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Kriteria 5	Keterangan	Nilai
1	100%	100%	100%	100%	100%	Lulus	1000
2	93%	94%	98%	92%	85%	Lulus	924
3	96%	93%	97%	96%	68%	Lulus	900
4	93%	97%	99%	95%	61%	Lulus	890
5	93%	90%	94%	83%	91%	Lulus	902
..
..
..
..
1022	58%	58%	57%	61%	84%	Tidak	637
1023	57%	63%	55%	61%	71%	Tidak	614
1024	56%	58%	61%	64%	61%	Tidak	600

2) *Data Transformation* (Pengubahan Data)

Data yang sudah dipilih akan diproses lebih lanjut dalam tahap transformasi data. Tahap ini melibatkan penggabungan data dan pengklasifikasian ke dalam berbagai kategori. Berikut adalah tabel yang menunjukkan hasil pengklasifikasian data:

a) Kriteria1

Variabel ini berisi range nilai yang telah di klasifikasikan menjadi 5 Kategori, yaitu A, B, C, D dan E. Berikut kriteria1 yang telah di klasifikasi sebagai berikut:



Tabel 3. Klasifikasi Kriteria

Kriteria1	Kategori
90%-100%	A
80%-89%	B
65%-79%	C
55%-64%	D
<55%	E

b) Kriteria2

Variabel ini berisi range nilai yang telah di klasifikasikan menjadi 5 Kategori,yaitu A, B, C, D dan E. Berikut kriteria2 yang telah di klasifikasi sebagai berikut:

Tabel 4. Klasifikasi Kriteria 2

Kriteria2	Kategori
90%-100%	A
80%-89%	B
65%-79%	C
55%-64%	D
<55%	E

c) Kriteria3

Variabel ini berisi range nilai yang telah di klasifikasikan menjadi 5 Kategori,yaitu A, B, C, D dan E. Berikut kriteria3 yang telah di klasifikasi sebagai berikut:

Tabel 5. Klasifikasi Kriteria 3

Kriteria3	Kategori
90%-100%	A
80%-89%	B
65%-79%	C
55%-64%	D
<55%	E

d) Kriteria4

Variabel ini berisi range nilai yang telah di klasifikasikan menjadi 5 Kategori,yaitu A, B, C, D dan E. Berikut kriteria4 yang telah di klasifikasi sebagai berikut:

Tabel 6. Klasifikasi Kriteria 4

Kriteria4	Kategori
90%-100%	A
80%-89%	B



65%-79%	C
55%-64%	D
<55%	E

e) Kriteria5

Variabel ini berisi range nilai yang telah di klasifikasikan menjadi 5 Kategori,yaitu A, B, C, D dan E. Berikut kriteria5 yang telah di klasifikasi sebagai berikut:

Tabel 7. Klasifikasi Kriteria 5

Kriteria5	Kategori
90%-100%	A
80%-89%	B
65%-79%	C
55%-64%	D
<55%	E

f) Tingkat Kelulusan

Variabel tingkat keberhasilan merupakan kategori atau informasi yang digunakan untuk menentukan hasil suatu keputusan. Pada tahap pengelompokan data, klasifikasi telah ditentukan secara pasti untuk mencegah kesalahan dalam perhitungan yang dilakukan oleh program. Data keputusan tingkat keberhasilan memiliki dua nilai yaitu “LULUS” dan “TIDAK”.

Berikut ini merupakan data yang telah di klasifikasi berdasarkan kategori masing-masing atribut:

Tabel 8. Tabel Data *Transformation*

No	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Kriteria 5	Keterangan
1	A	A	A	A	A	Lulus
2	A	A	A	A	B	Lulus
3	A	A	A	A	C	Lulus
4	A	A	A	A	D	Lulus
5	A	A	A	B	A	Lulus
..
..
..
..
1022	D	D	D	D	D	Tidak
1023	D	D	D	D	D	Tidak
1024	D	D	D	D	D	Tidak



d. *Modelling Phase* (Fase Permodelan)

Untuk tahap ini, data yang sudah diklasifikasikan akan diproses lebih lanjut melalui berbagai model. Proses pemodelan ini bertujuan untuk memaksimalkan pencapaian hasil yang diinginkan.

1. *Algoritma Naïve Bayes*

Setelah menyelesaikan semua tahapan pengolahan data, akan dihasilkan data latihan. Data latihan ini akan digunakan untuk melakukan komputasi menggunakan metode Naive Bayes. Berikut ini adalah gambaran proses komputasi dengan menggunakan metode Naive Bayes yang diilustrasikan dalam bentuk flowchart, yaitu:



Gambar 2. Flowchart Naïve Bayes

Berikut ini adalah tahapan yang dijalankan, di mana dilakukan penghitungan menggunakan Metode Naive Bayes dengan menggunakan data latih dan data uji sebagai data evaluasi:

a) *Baca Data Training*

Berikut adalah data latihan yang digunakan untuk melakukan perhitungan komputasi.

Tabel 9. Tabel *Training*

No	Kriteria 1	Kriteria 2	Kriteria 3	Kriteria 4	Kriteria 5	Keterangan
1	A	A	A	A	A	Lulus
2	A	A	A	A	B	Lulus
3	A	A	A	A	C	Lulus
4	A	A	A	A	D	Lulus
5	A	A	A	B	A	Lulus



..
..
..
..
1022	D	D	D	D	D	Tidak
1023	D	D	D	D	D	Tidak
1024	D	D	D	D	D	Tidak

Dari data training diatas, akan di uji menggunakan data testing sebagai berikut, yaitu:

Tabel 10. Tabel *Testing*

Nama	Kriteria1	Kriteria2	Kriteria3	Kriteria4	Kriteria5	Tingkat Kelulusan
Bowval Valentino	90	88	78	85	75	?

Data testing diatas akan di klasifikasi berdasarkan data preparation (datatransformation), yaitu:

Tabel 11. Tabel *Testing Transformation*

Nama	Kriteria1	Kriteria2	Kriteria3	Kriteria4	Kriteria5	Tingkat Kelulusan
Bowval Valentino	A	B	C	B	C	?

b) Hitung $P(C1)$ pada setiap Class

Untuk menghitung Probabilitas class, yaitu:

$$P(C1) = \frac{\text{Classtingkatkelulusan} = \text{LULUS}}{\text{Jumlahdatauji}}$$

$$P(C2) = \frac{\text{Classtingkatkelulusan} = \text{TIDAK}}{\text{Jumlahdatauji}}$$

$$1) \quad P(\text{LULUS}) = \frac{916}{1024} = 0.894531$$

$$2) \quad P(\text{TIDAK}) = \frac{108}{1024} = 0.105469$$

c) Hitung probabilitas pada setiap kriteria

$$P(X/H).P(H)$$

$$P(X/H) = \frac{P(X/H).P(H)}{P(X)}$$

$$P(X)$$

1) Kriteria = A

$$i. \quad P(\text{Kriteria1} = A \mid \text{Tingkat kelulusan} = \text{YA}) = \frac{249}{916} = 0.271834$$



ii. $P(\text{Kriteria1} = A \mid \text{Tingkat kelulusan} = \text{TIDAK}) = \frac{7}{108} = 0.064815$

2) Kriteria2 = B

i. $P(\text{Kriteria2} = B \mid \text{Tingkat kelulusan} = \text{YA}) = \frac{243}{916} = 0.265284$

ii. $P(\text{Kriteria2} = B \mid \text{Tingkat kelulusan} = \text{TIDAK}) = \frac{13}{108} = 0.12037037$

3) Kriteria3 = C

i. $P(\text{Kriteria3} = C \mid \text{Tingkat kelulusan} = \text{YA}) = \frac{227}{916} = 0.247817$

ii. $P(\text{Kriteria3} = C \mid \text{Tingkat kelulusan} = \text{TIDAK}) = \frac{29}{108} = 0.268519$

4) Kriteria4 = B

i. $P(\text{Kriteria4} = B \mid \text{Tingkat kelulusan} = \text{YA}) = \frac{241}{916} = 0.265284$

ii. $P(\text{Kriteria4} = B \mid \text{Tingkat kelulusan} = \text{TIDAK}) = \frac{13}{108} = 0.120370$

5) Kriteria5 = C

i. $P(\text{Kriteria5} = C \mid \text{Tingkat kelulusan} = \text{YA}) = \frac{222}{916} = 0.245633$

ii. $P(\text{Kriteria5} = C \mid \text{Tingkat kelulusan} = \text{TIDAK}) = \frac{56}{108} = 0.287037$

d) Perkalian nilai probabilitas pada setiap class

1) Probabilitas LULUS

$$\frac{\frac{\frac{\frac{\frac{\frac{916}{1024} \times 249}{916} \times 243}{916} \times 227}{916} \times 241}{916} \times 222}{916}$$

$$0.894531 * 0.271834 * 0.265284 * 0.247817 * 0.265284 * 0.245633 = 0.001042$$

2) Probabilitas TIDAK

$$\frac{\frac{\frac{\frac{\frac{\frac{108}{1024} \times 7}{108} \times 13}{108} \times 29}{108} \times 13}{108} \times 56}{108}$$

$$0.105469 * 0.064815 * 0.120370 * 0.268519 * 0.120370 * 0.287037 = 0.000008$$

e) Hitung Persentase Keakurasian Kelulusan



- 1) Akurasi Persentase LULUS

$$\left(\frac{nLULUS}{nTIDAK + nLULUS} \right) * 100\%$$
$$\left(\frac{0.001042}{0.000008 + 0.001042} \right) * 100\%$$
$$99.24\%$$

- 2) Akurasi Persentase TIDAK

$$\left(\frac{nTIDAK}{nLULUS + nTIDAK} \right) * 100\%$$
$$\left(\frac{0.000008}{0.001042 + 0.000008} \right) * 100\%$$
$$0.76\%$$

- f) Membandingkan Nilai Akurasi

Dari hasil tersebut menyatakan $P(X|C1) > P(X|C2)$ atau $P(X|C2) < P(X|C1)$

- g) Hasil

Berdasarkan informasi yang diberikan, dari 1024 data pelatihan yang telah diklasifikasikan dan diuji di laboratorium, Metode Naive Bayes berhasil mengestimasi tingkat kelulusan peserta sertifikasi dengan akurasi mencapai 99.24%.

- e. *Evaluation Phase* (Fase Evaluasi)

Fase evaluasi merupakan langkah untuk menentukan apakah model yang telah dibuat telah mencapai tujuan awalnya atau belum. Tujuan utama dari pengembangan model ini adalah untuk mencapai tingkat akurasi yang tinggi, sebagai bukti keberhasilan dari penelitian yang dilakukan. Jika hasil yang diperoleh belum sesuai dengan harapan, maka dilakukan analisis lebih mendalam guna menghasilkan rekomendasi perbaikan atau penyesuaian yang diperlukan untuk menentukan tingkat kelulusan peserta sertifikasi. Dengan demikian, hal ini akan membantu laboratorium dan calon peserta sertifikasi dalam menentukan tingkat kelulusan serta program sertifikasi yang sesuai dengan sistem prediksi yang telah tersedia.

- f. *Deployment Phase* (Fase Penyebaran)

Proses penyebaran adalah fase implementasi untuk mengembangkan sebuah aplikasi yang berisi rangkuman pengetahuan yang sudah diperoleh agar sistem ini dapat dimanfaatkan oleh pengguna lainnya. Dalam konteks ini, proses penyebaran data mining dapat dilakukan terhadap semua calon peserta uji sertifikasi di Laboratorium ITCC. Namun, dalam penelitian ini, tahap implementasi atau penyebaran tidak dilaksanakan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Rancangan Aplikasi

Aplikasi untuk menentukan kelulusan peserta dalam program sertifikasi Word dan Excel dapat diakses oleh dua jenis pengguna: peserta sebagai pengguna reguler dan pengelola laboratorium ITCC sebagai administrator. Setiap pengguna memiliki hak akses yang berbeda terhadap aplikasi tersebut. Algoritma Naive Bayes digunakan untuk mengevaluasi prestasi peserta dalam sertifikasi Word dan Excel dengan membandingkan kedua program tersebut.



	atribut1	atribut2	atribut3	atribut4	atribut5	atribut6	Kategori
1	100	100	100	100	100	100	LULUS
2	100	100	100	100	100	100	LULUS
3	100	100	100	100	100	100	LULUS
4	100	100	100	100	100	100	LULUS
5	100	100	100	100	100	100	LULUS
6	100	100	100	100	100	100	LULUS
7	100	100	100	100	100	100	LULUS
8	100	100	100	100	100	100	LULUS
9	100	100	100	100	100	100	LULUS
10	100	100	100	100	100	100	LULUS

Gambar 3. Hasil Aplikasi Training

	atribut1	atribut2	atribut3	atribut4	atribut5	atribut6	Kategori
1	100	100	100	100	100	100	LULUS
2	100	100	100	100	100	100	LULUS
3	100	100	100	100	100	100	LULUS
4	100	100	100	100	100	100	LULUS
5	100	100	100	100	100	100	LULUS
6	100	100	100	100	100	100	LULUS
7	100	100	100	100	100	100	LULUS
8	100	100	100	100	100	100	LULUS
9	100	100	100	100	100	100	LULUS
10	100	100	100	100	100	100	LULUS

Gambar 4. Hasil Aplikasi Testing

3.2 Pembahasan

Aplikasi ini dirancang untuk memprediksi tingkat kelulusan peserta sertifikasi, membantu dalam penentuan program sertifikasi bagi calon peserta. Prediksi dilakukan dengan menggunakan data training sebanyak 1024 entri yang sebelumnya telah melalui tahapan seleksi dan transformasi data. Data training ini kemudian digunakan untuk menghitung menggunakan metode Naive Bayes pada data testing yang akan dievaluasi. Tujuan pengujian ini adalah untuk menilai tingkat akurasi prediksi yang telah dilakukan.

Sebelumnya, belum ada aplikasi yang mendukung prediksi tingkat kecocokan jenis tanaman sayuran karena pengolahan data masih dilakukan secara manual, sehingga menghambat penerapan metode pengolahan data tertentu.

Berdasarkan penelitian menggunakan pendekatan Naive Bayes, probabilitas kelas LULUS diperoleh sebesar 0.89453125, sedangkan untuk kelas TIDAK sebesar 0.10546875, dengan tingkat akurasi mencapai 99%. Hasil tersebut menunjukkan bahwa peserta sertifikasi yang memiliki kriteria1, kriteria2, kriteria3, kriteria4, dan kriteria5 sebagai data testing memiliki tingkat kelulusan yang lebih tinggi dibandingkan dengan yang tidak memenuhi persyaratan untuk sertifikasi.

4. KESIMPULAN

Berdasarkan analisis dan diskusi yang telah dilakukan, dapat ditarik beberapa simpulan, antara lain:

1. Penerapan algoritma naïve bayes dalam aplikasi prediksi dapat menyajikan informasi mengenai prediksi status kelulusan peserta sertifikasi, yang berguna bagi peserta dan pengelola laboratorium ITCC untuk mengevaluasi tingkat kelulusan dalam sertifikasi tersebut.



2. Aplikasi ini dapat memberikan rekomendasi keputusan dalam memilih program sertifikasi yang tepat dengan membandingkan hasil uji prediksi antara aplikasi word dan excel.

REFERENSI

- Larose D, T., 2005, *Discovering knowledge in data: an introduction to data mining*, Jhon Wiley & Sons Inc.
- Syarli, & Muin, A. A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer*, 2(1), 1–5.
- Budiman, I., Prahasto, T., & Christyono, Y. (2012). Data Clustering Menggunakan Metodologi Crisp-Dm Untuk Pengenalan Pola Proporsi Pelaksanaan Tridharma, 2012(Snati), 15–16.
- A.S, Rosa dan Shalauddin, M. (2014). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek*. Bandung: Informatika Bandung.
- Siregar, R. R. A., Sinaga, F. A., & Arianto, R. (2017). Aplikasi Penentuan Dosen Penguji Skripsi Menggunakan Metode TF-IDF dan Vector Space Model. *Computatio: Journal of Computer Science and Information Systems*, 1(2), 171-186.
- T. Nisa, R. Siregar, And W. Suliyanti, “Estimasi Daya Beban Listrik Pada Gardu Induk Cengkareng Dengan Menggunakan Metode Time Series Model Dekomposisi”, *Teknologia*, Vol. 1, No. 2, Apr. 2019.
- [1] Purwanto, S. D., & Santoso, I. B. (2017). Sistem Identifikasi Boraks pada Bakso Daging Sapi Berbasis Android Menggunakan Algoritma Naive Bayes Classifier, 9(1), 33–37.
- Nurrohmat, M. A., & Nugroho, Y. S. (2015). Aplikasi Pemrediksi Masa Studi dan Predikat Kelulusan Mahasiswa Informatika Universitas Muhammadiyah Surakarta Menggunakan Metode Naïve Bayes. *Khazanah Informatika*, I(1), 29–34. <https://doi.org/10.23917/khif.v1i1.1179>
- Siregar, R., Siregar, Z., & Arianto, R. (2019). Klasifikasi Sentiment Analysis Pada Komentar Peserta Diklat Menggunakan Metode K-Nearest Neighbor. *KILAT*, 8(1). <https://doi.org/10.33322/kilat.v8i1.421>
- Siregar Rr, Putri Dr. Metode Support Vector Machine Pada Klasifikasi Audit Energi: Studi Kasus Gedung STT-PLN Jakarta. *Jurnal Informatika Dan Komputasi*. 2017 Mar 2;8(2):98- 104.
- Sangadji, I., & Arvio, Y. (2018, March). Dynamic Segmentation Of Behavior Patterns Based On Quantity Value Movement Using Fuzzy Subtractive Clustering Method. In *Journal of Physics: Conference Series* (Vol. 974, No. 1, p. 012009). IOP Publishing.
- R. Jurnal, “Implementasi Logika Fuzzy Untuk Sistem Otomatisasi Pengaturan Pengisian Batere Pembangkit Listrik Tenaga Surya”, *Energi*, Vol. 9, No. 2, Pp. 111-119, Nov. 2018.
- A. Prianty, R. Siregar, and R. Arianto, “Penanganan Gangguan Listrik Rumah Tangga Menggunakan Algoritma Greedy Untuk Penentuan Jarak Optimal”, *Teknologia*, vol. 2, no. 1, Aug. 2019.