

**KLASIFIKASI DOKUMEN KARYA AKHIR MAHASISWA
MENGUNAKAN NAÏVE BAYES CLASSIFIER (NBC)
BERDASARKAN ABSTRAK KARYA AKHIR DI JURUSAN
TEKNIK ELEKTRO UNIVERSITAS NEGERI JAKARTA**



**NUR INDAH PRATIWI
5235107412**

**Skripsi ini Ditulis Untuk Memenuhi Sebagian Persyaratan dalam
Memperoleh Gelar Sarjana**

**PROGRAM STUDI PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA
2014**

**KLASIFIKASI DOKUMEN KARYA AKHIR MAHASISWA
MENGUNAKAN NAÏVE BAYES CLASSIFIER (NBC) BERDASARKAN
ABSTRAK KARYA AKHIR DI JURUSAN TEKNIK ELEKTRO
UNIVERSITAS NEGERI JAKARTA**

NUR INDAH PRATIWI

ABSTRAK

Dokumen karya akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta setiap tahunnya bertambah, pengklasifikasian dokumen menjadi hal yang sangat penting untuk mengorganisasikan dokumen sehingga dapat memudahkan pencarian. Pengembangan Sistem klasifikasi dokumen bertujuan untuk mengembangkan sebuah sistem yang dapat mengklasifikasikan dokumen karya akhir mahasiswa berdasarkan abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier* (NBC). Sehingga, dapat memudahkan pengklasifikasian dokumen karya akhir di Jurusan Teknik Elektro. Dalam penelitian ini menggunakan metode eksperimen dan menggunakan 100 dokumen abstrak, 90 dokumen sebagai *data train* dan 10 dokumen sebagai *data test*. Data diambil dari skripsi mahasiswa Jurusan Teknik Elektro Universitas Negeri Jakarta dari 14 Maret 2014 sampai dengan 27 Maret 2014. Setelah melakukan proses pengembangan perangkat lunak, dihasilkan sebuah sistem klasifikasi yang bernama Sistem Klasifikasi Dokumen Skripsi. Sistem di implementasi menggunakan PHP dan MySQL, dan diuji menggunakan *K-Fold Cross Validation* (*10 Fold*). Berdasarkan pada hasil uji Sistem didapatkan hasil tingkat akurasi sebesar 81%. Oleh karena itu, dapat disimpulkan bahwa Sistem Klasifikasi Dokumen Abstrak Karya Akhir Menggunakan Algoritma *Naïve Bayes* di Jurusan Teknik Elektro telah berhasil dikembangkan.

Kata kunci: sistem, *Naïve Bayes Classifier*, klasifikasi, dokumen, dan algoritma.

**COLLEGE STUDENT THESIS DOCUMENT CLASSIFICATION USING
NAÏVE BAYES CLASSIFIER (NBC) BASED ON ABSTRACT THESIS IN
THE DEPARTMENT OF ELECTRICAL ENGINEERING
STATE UNIVERSITY OF JAKARTA**


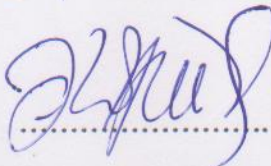
NUR INDAH PRATIWI

ABSTRACT

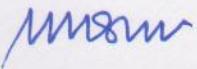
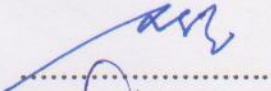
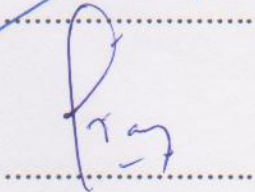
College student thesis document in the Department of Electrical Engineering, State University of Jakarta increases every year, the classification of documents becomes very important to organize the documents in order to facilitate the research. Development of document classification system aim to develop a system which can classify college student thesis document based on abstract using Naïve Bayes Classifier (NBC) algorithm. Thus, to facilitate the work of classifying documents in the Department of Electrical Engineering. This research using experimental method and using 100 abstracts of documents, 90 documents as a data train and 10 documents as data test. Data taken from college student thesis of Department of Electrical Engineering, State University of Jakarta since March 14, 2014 until March 27, 2014. After software development process, generated a classification system called Thesis Document Classification System. Implementations of system using PHP and MySQL, and tested using K-Fold Cross Validation (10 Fold). Based on the results of the test system showed an accuracy rate of 81%. Therefore, it can be concluded that the college student thesis document classification system using Naïve Bayes algorithm in the Department of Electrical Engineering has been successfully developed.

Keyword: system, *Naïve Bayes Classifier*, classification, document, dan algorithm.

LEMBAR PENGESAHAN

NAMA DOSEN	TANDA TANGAN	TANGGAL
Hamidillah Ajie, S.Si., M.T (Dosen Pembimbing I)		18/7-2014
Widodo, M.Kom (Dosen Pembimbing II)		17/7-2014

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
Drs. Wisnu Djatmiko, M. T. (Ketua Penguji)		22/07-2014
Prasetyo Wibowo Y., M. Eng (Dosen Penguji)		21/07-2014
Dra. Yuliatrisa Sastrawidjaya, M.Pd (Dosen Penguji Ahli)		17/7-2014

Tanggal Lulus:

HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Karya tulis skripsi saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di perguruan tinggi lain
2. Karya tulis ini adalah murni gagasan, rumusan, dan penelitian saya sendiri dengan arahan dosen pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, Juli 2014

Yang membuat pernyataan



Nur Indah Pratiwi
5235107412

KATA PENGANTAR

Segala puji dan syukur kehadiran Allah swt atas rahmat dan karuniaNya yang telah diberikan kepada peneliti, sehingga peneliti dapat menyelesaikan skripsi dengan judul: “Klasifikasi Dokumen Karya Akhir Mahasiswa Menggunakan Naïve Bayes Classifier (NBC) Berdasarkan Abstrak Karya Akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta”.

Dalam merencanakan, menyusun dan menyelesaikan penelitian skripsi, peneliti banyak menerima bantuan, bimbingan, dan motivasi serta dukungan dari berbagai pihak, oleh karena itu pada kesempatan ini peneliti bermaksud mengucapkan terima kasih yang sebesar-besarnya disampaikan kepada:

1. Bapak Hamidillah Ajie, S.Si., M.T selaku pembimbing I dan Bapak Widodo, M.Kom selaku pembimbing II yang telah memberikan motivasi, arahan dan kepercayaan kepada peneliti dalam menyelesaikan skripsi ini.
2. Orang tua yang tak hentinya memanjatkan doa dan memberikan semangat kepada peneliti dalam menyelesaikan skripsi ini.

Peneliti menyadari bahwa skripsi ini jauh dari kesempurnaan, untuk itu peneliti mohon maaf apabila terdapat kekurangan dan kesalahan baik dari isi maupun penelitian. Akhir kata peneliti berharap agar penelitian dan penyusunan proposal penelitian ini dapat bermanfaat bagi para pembaca dan semua pihak yang terkait.

Jakarta, Juli 2014

Peneliti

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	x
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	
1.1. Latar Belakang Masalah	1
1.2. Identifikasi Masalah	3
1.3. Pembatasan Masalah	3
1.4. Perumusan Masalah.....	4
1.5. Tujuan Umum Penelitian.....	4
1.6. Kegunaan Penelitian	4
BAB II KAJIAN TEORITIS DAN KERANGKA BERPIKIR	
2.1. Kajian Teoritis	5
2.1.1. <i>Natural Language Processing</i>	5
2.1.1.1. Information Retrieval	7

2.1.1.2.	Text Mining.....	9
2.1.1.3.	Ekstraksi Dokumen	13
2.1.2.	Data Mining	17
2.1.2.1.	Pengertian Data Mining	17
2.1.2.2.	Data	23
2.1.2.3.	Klasifikasi Dokumen.....	23
2.1.2.4.	Naïve Bayes Classifier	25
2.1.2.5	Kehandalan Naïve Bayes Classifier	31
2.2	Kerangka Berpikir	34

BAB III METODOLOGI PENELITIAN

3.1.	Tujuan Operasional Penelitian	38
3.2.	Tempat dan Waktu Penelitian	38
3.3.	Metode Penelitian.....	38
3.4.	Instrumen Penelitian.....	39
3.5.	Pelaksanaan Penelitian	39
3.5.1.	Proses Pengembangan Software	39
3.5.2.	Definisi Kebutuhan.....	40
3.5.3.	Desain dan Analisis Sistem	41
3.5.3.1.	Diagram Konteks Usulan	42
3.5.3.2.	Diagram Nol Sistem.....	42
3.5.3.3.	Diagram Proses 1 Level 1	43
3.5.3.4.	Diagram Proses 2 Level 1	45

3.5.3.5.	Flow Chart Diagram.....	46
3.5.3.6.	Desain Struktur Database.....	46
3.5.3.7.	Pembuatan Desain Tampilan Halaman Sistem.....	47
3.5.3.8.	Implementasi Tampilan Halaman Sistem	49
3.5.3.9.	Langkah Kerja Sistem	51
3.5.4.	Implementasi dan <i>Testing Unit</i>	52
3.5.5.	Integrasi dan <i>Testing System</i>	52
3.5.6.	Operasi dan <i>Maintenance</i>	54

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

4.1.	Hasil Penelitian.....	55
4.1.1.	Hasil Pengujian Perangkat Lunak.....	55
4.1.2.	Hasil Pengujian Algoritma <i>Naïve Bayes Classifier</i>	57
4.2.	Pembahasan	98
4.2.1.	Pembahasan Hasil Pengujian Perangkat Lunak.....	98
4.2.2.	Pembahasan Hasil Pengujian Algoritma <i>Naïve Bayes Classifier</i>	99

BAB V KESIMPULAN DAN SARAN

5.1.	Kesimpulan.....	102
5.2.	Implikasi Bidang Keteknikan dan Pendidikan	104
5.2.1.	Implikasi Bidang Keteknikan	104
5.2.2.	Implikasi Bidang Pendidikan.....	104
5.3.	Saran	104

DAFTAR PUSTAKA	106
LAMPIRAN	109
TENTANG PENULIS.....	150

DAFTAR TABEL

2.1.	Aturan Pemenggalan Imbuhan Enhanced Confix Stripping Stemmer	15
2.2.	Frekuensi Kata dalam Dokumen Data Set.....	26
2.3.	Hasil Perhitungan Probabilitas	28
2.4.	Data Sebuah Dokumen Tes	29
3.1.	Tahap Pengembangan Software	40
3.2.	Daftar Kebutuhan Fungsional.....	41
3.3.	Tabel kata dasar pada <i>database</i> karya akhir	46
3.4.	Tabel <i>stoplist</i> pada <i>database</i> karya akhir.....	47
3.5.	Tabel karya akhir pada <i>database</i> karya akhir	47
3.6.	Tabel upload pada <i>database</i> karya akhir	47
3.7.	Kriteria pengujian kebutuhan fungsional	53
4.1.	Hasil pengujian kebutuhan fungsional.....	56
4.2.	Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 1 (Fold 1)	58
4.3.	Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 1 (Fold 1)	60
4.4.	Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 2 (Fold 2)	62
4.5.	Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 2 (Fold 2)	64
4.6.	Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 3 (Fold 3)	66

4.7. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 3 (Fold 3)	68
4.8. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 4 (Fold 4)	70
4.9. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 4 (Fold 4)	72
4.10. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 5 (Fold 5)	74
4.11. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 5 (Fold 5)	76
4.12. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 6 (Fold 6)	78
4.13. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 6 (Fold 6)	80
4.14. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 7 (Fold 7)	82
4.15. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 7 (Fold 7)	84
4.16. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 8 (Fold 8)	86
4.17. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 8 (Fold 8)	88
4.18. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 9 (Fold 9)	90
4.19. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 9 (Fold 9)	92

4.20. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 10 (Fold 10)	94
4.21. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 10 (Fold 10)	96
5.1. Hasil Pengujian Sistem Klasifikasi Dokumen.....	102

DAFTAR GAMBAR

2.1	Proses Text Mining	11
2.2	Proses Case Folding dan Tokenizing.....	13
2.3.	Proses Filtering	14
2.4.	Proses Stemming	14
2.5.	Data Mining sebagai bagian dari proses Knowledge Discovery from Data.....	19
2.6.	Arsitektur Data Mining.....	20
2.7.	Data Mining sebagai pertemuan antar berbagai disiplin ilmu	22
2.8.	Bagan Kerangka Berpikir	34
2.9.	Gambaran Proses Sistem Klasifikasi Dokumen Abstrak.....	36
3.1.	Metode Waterfall	39
3.2.	Diagram Konteks Usulan Sistem Pengklasifikasi Dokumen menggunakan <i>Naïve Bayes Classifier</i>	42
3.3.	Diagram Nol Sistem Pengklasifikasi Dokumen menggunakan <i>Naïve Bayes Classifier</i>	42
3.4.	Diagram Proses 1 Level 1 Sistem Pengklasifikasi Dokumen menggunakan <i>Naïve Bayes Classifier</i>	44
3.5.	Diagram Proses 2 Level 1 Sistem Pengklasifikasi Dokumen menggunakan <i>Naïve Bayes Classifier</i>	45
3.6.	Flow Chart Diagram	46
3.7.	Desain tampilan halaman utama Sistem Pengklasifikasi Dokumen.....	48
3.8.	Tampilan halaman utama Sistem Pengklasifikasi Dokumen.....	49

3.9. Tampilan halaman result Sistem Pengklasifikasi Dokumen	50
---	----

DAFTAR LAMPIRAN

1.	Perhitungan manual probabilitas pada <i>Fold 1</i>	109
2.	Perhitungan manual probabilitas pada <i>Fold 2</i>	113
3.	Perhitungan manual probabilitas pada <i>Fold 3</i>	117
4.	Perhitungan manual probabilitas pada <i>Fold 4</i>	121
5.	Perhitungan manual probabilitas pada <i>Fold 5</i>	125
6.	Perhitungan manual probabilitas pada <i>Fold 6</i>	129
7.	Perhitungan manual probabilitas pada <i>Fold 7</i>	133
8.	Perhitungan manual probabilitas pada <i>Fold 8</i>	137
9.	Perhitungan manual probabilitas pada <i>Fold 9</i>	141
10.	Perhitungan manual probabilitas pada <i>Fold 10</i>	145
11.	Lembar Pernyataan Validasi Kategori Manual.....	149

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Setiap tahunnya jumlah dokumen karya akhir (skripsi, karya inovatif dan komprehensif) di Universitas Negeri Jakarta semakin bertambah. Mahasiswa yang akan menempuh jalur karya akhir tersebut akan mencari referensi dari karya akhir yang sudah ada. Tentunya dalam proses pencarian, mahasiswa akan dihadapkan oleh beberapa kategori dokumen karya akhir yang akan dipilih. Proses pencarian yang mudah akan membuat mahasiswa secara efisien mencari referensi atas kategori bidang ilmu yang membuatnya tertarik. Namun sebaliknya, jika proses pencarian berdasarkan kategori tidak terstruktur dengan baik, mahasiswa akan memakan waktu yang lama untuk mencari satu referensi. Sehingga, sistem pengklasifikasi karya akhir yang baik sangat dibutuhkan untuk membantu mahasiswa mengembangkan suatu arah penelitian. Dengan adanya pengklasifikasi dokumen, arah penelitian dari suatu Jurusan di Universitas Negeri Jakarta juga dapat dilihat. Sehingga, untuk melihat kemana arah penelitian atau kekuatan penelitian Jurusan di Universitas Negeri Jakarta dapat dilihat dari kategori karya akhir mahasiswa. Setiap karya akhir yang dihasilkan sebaiknya dipublikasikan. Oleh sebab itu, koleksi karya akhir akan sangat banyak. Proses pencarian menjadi hal yang sangat berat tanpa sistem yang baik.

Pengklasifikasian dokumen di Jurusan Teknik Elektro Universitas Negeri Jakarta (UNJ) saat ini masih dilakukan secara manual. Artinya, setiap mahasiswa

yang ingin mencari informasi pada salah satu judul skripsi harus membaca terlebih dahulu abstrak secara keseluruhan. Ketika mahasiswa mencari suatu informasi berbentuk dokumen dengan membacanya terlebih dahulu, hal tersebut dapat memakan waktu lebih lama jika informasi yang dicari tidak sesuai dengan apa yang ia harapkan. Sebuah sistem klasifikasi dokumen dapat dikembangkan untuk mengklasifikasikan abstrak sebuah karya akhir, yang selanjutnya untuk menentukan karya akhir tersebut termasuk dalam kategori apa.

Klasifikasi dokumen merupakan sebuah proses yang penting dalam *data mining*. Pada klasifikasi kelompok data (*class label*) yang sudah diketahui, sebuah data akan masuk ke dalam kelompok tertentu yang sebelumnya telah ditentukan. Manfaat dari klasifikasi dokumen adalah untuk mengorganisasikan dokumen. Dengan semakin meningkatkan jumlah dokumen yang bertambah setiap harinya, maka akan lebih mudah mencari informasi dari sebuah dokumen yang telah terorganisasi dan telah dikelompokkan menurut kategorinya masing-masing. Manfaat pentingnya klasifikasi dapat ditemui contohnya pada aplikasi *spam filtering*.

Naïve Bayes adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk *data mining*. Performa *Naïve Bayes* kompetitif dalam proses klasifikasi, walaupun menggunakan asumsi keindependenan atribut (tidak ada kaitan antar atribut).

Oleh karena itu, dibutuhkan suatu sistem yang dapat mengklasifikasikan dokumen berdasarkan abstrak karya akhir dengan cepat dan relevan sehingga *user* dapat menemukan informasi yang dicari dalam waktu yang efisien. Hal tersebut yang melatarbelakangi dibuatnya sebuah sistem klasifikasi dokumen menggunakan algoritma *Naïve Bayes Classifier (NBC)*.

1.2. Identifikasi Masalah

Berdasarkan latar belakang di atas, masalah yang dapat diidentifikasi sebagai berikut:

1. Bagaimana prinsip kerja *Naïve Bayes Classifier* dalam sistem pengklasifikasian suatu dokumen?
2. Bagaimana proses Klasifikasi Dokumen Karya Akhir Mahasiswa Menggunakan *Naïve Bayes Classifier (NBC)* Berdasarkan Abstrak Karya Akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta?
3. Apakah sistem pengklasifikasi yang menggunakan *Naïve Bayes Classifier* dapat melakukan pengklasifikasian dokumen?

1.3. Pembatasan Masalah

Melihat luasnya lingkup permasalahan yang diidentifikasi pada pembahasan sebelumnya, pembatasan masalah sangat penting untuk dilakukan.

Penelitian dibatasi pada:

1. *Data training* atau data latih menggunakan abstrak skripsi.
2. Dokumen abstrak pada penelitian ini adalah dokumen yang memiliki format data .doc, dan .docx

3. Dokumen yang digunakan untuk proses pengklasifikasian ini adalah dokumen abstrak yang menggunakan Bahasa Indonesia.
4. Kesalahan penulisan di dalam abstrak skripsi, akan diabaikan.

1.4. Perumusan Masalah

Berdasarkan proses latar belakang, identifikasi, dan pembatasan masalah, maka perumusan masalah yang akan dibahas pada penelitian ini adalah:

Bagaimana menghasilkan suatu sistem klasifikasi dokumen karya akhir mahasiswa menggunakan Naïve Bayes Classifier (NBC) berdasarkan abstrak karya akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta?

1.5. Tujuan Penelitian

Tujuan dari penelitian ini adalah mengembangkan sebuah sistem yang dapat mengklasifikasikan dokumen karya akhir mahasiswa berdasarkan abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier* (NBC).

1.6. Kegunaan Penelitian

Kegunaan dari penelitian ini dapat membantu memudahkan *user* mencari informasi yang telah diklasifikasikan menggunakan *Naïve Bayes Classifier* berdasarkan abstrak karya akhir mahasiswa di Jurusan Teknik Elektro Universitas Negeri Jakarta. Hasil dari sistem ini merupakan kategori yang dapat digunakan untuk pengembangan aplikasi selanjutnya.

BAB II

KAJIAN TEORITIS DAN KERANGKA BERPIKIR

2.1. Kajian Teoritis

2.1.1. Natural Language Processing

Natural Language Processing (NLP) adalah pemrosesan bahasa alami yang merupakan salah satu tujuan jangka panjang *Artificial Intelligence* (Kecerdasan Buatan) yang digunakan untuk membuat suatu program yang dapat memiliki kemampuan memahami bahasa manusia.¹

Tujuan dalam bidang *Natural Language* ini adalah melakukan proses pembuatan model komputasi dari bahasa, bahasa yang dihasilkan ini dapat dimengerti oleh komputer sehingga dapat terjadi suatu interaksi antara manusia dan komputer dengan perantara bahasa alami.

Ada 3 (tiga) aspek utama pada teori pemahaman mengenai *Natural Language*, yaitu:

1. **Sintaksis:** yaitu pemahaman tentang urutan kata dalam pembentukan kalimat dan hubungan antar kata tersebut dalam proses perubahan bentuk dari kalimat menjadi bentuk yang sistematis. Meliputi proses pengaturan tata letak suatu kata dalam kalimat akan membentuk kalimat yang dapat dikenali.
2. **Semantik:** yaitu pemetaan bentuk struktur sintaksis dengan memanfaatkan tiap kata ke dalam bentuk yang lebih mendasar dan tidak tergantung struktur kalimat. Semantic mempelajari arti suatu kata

¹ Raymond J Mooney et al, *UTexas: Natural Language Semantics using Distributional Semantics and Probabilistic Logic*, (United States: The University of Texas at Austin, 2014), h.1-2.

dan bagaimana sekumpulan arti kata tersebut membentuk suatu arti kata dari kalimat yang utuh. Dalam tingkatan ini belum tercakup konteks dari kalimat tersebut.

3. **Pragmatik:** pengetahuan pada tingkatan ini berkaitan dengan masing-masing konteks yang berbeda tergantung pada situasi dan tujuan pembuatan sistem.²

Berikut ini adalah bidang-bidang pengetahuan yang berhubungan dengan *Natural Language*³:

1. **Fonetik dan fonologi:** berhubungan dengan suara yang menghasilkan kata yang dapat dikenali. Bidang ini menjadi penting dalam proses aplikasi yang memakai metode *speech based system*.
2. **Morfologi:** yaitu pengetahuan tentang kata dan bentuknya dimanfaatkan untuk membedakan satu kata dengan lainnya. Pada tingkat ini juga dapat dipisahkan antara kata dan elemen lain seperti tanda baca. Contoh: Melarikan (word)
 Lari (root)
 Me- (prefix)
 -kan (suffix)

² D. Poole and Alan Mackworth, *Artificial Intelligence: Natural Language Understanding*, (Canada: Canada License, 2010), h.2.

³ Derwin Suhartono, *Natural Language Processing*, <http://socs.binus.ac.id/2013/06/22/natural-language-processing/>, 9 Maret 2014.

2.1.1.1. Information Retrieval

Sejak tahun 1940-an masalah penyimpanan informasi dan pengambilan informasi telah menarik perhatian. Masalah yang ada yaitu informasi yang akurat dan cepat akses menjadi semakin sulit. Salah satu efek dari hal ini adalah informasi yang relevan akan diabaikan karena tidak pernah ditemukan, yang pada akhirnya menyebabkan banyak duplikasi pekerjaan dan usaha. Dengan munculnya komputer, banyak ide telah diberikan untuk menggunakan komputer dalam menyediakan sistem pencarian yang cepat dan cerdas. Di perpustakaan, memiliki penyimpanan informasi, beberapa tugas seperti katalogisasi dan administrasi umum, telah berhasil diambil alih oleh komputer. Namun, masalah pencarian yang efektif sebagian besar masih belum terpecahkan. Seorang pengguna tidak mempunyai waktu atau tidak ingin menghabiskan waktu membaca koleksi seluruh dokumen, terlepas dari kenyataan bahwa mungkin secara fisik tidak mungkin baginya untuk melakukannya. Ketika komputer kecepatan tinggi menjadi tersedia untuk non-numerik kerja, banyak yang mengira bahwa komputer akan mampu membaca koleksi seluruh dokumen untuk mengekstrak dokumen yang relevan.

Pengertian dari kata *Information Retrieval* dapat berarti sangat luas. Dalam kasus pembelajaran, *Information Retrieval* dapat diartikan sebagai berikut: “*Information retrieval* adalah menemukan suatu bahan yang biasanya berbentuk dokumen dari suatu data yang tidak terstruktur yang dapat memenuhi kebutuhan

informasi dari sebuah penyimpanan yang besar dan biasanya disimpan di dalam komputer.”⁴

Information retrieval akan berhubungan dengan bagaimana cara untuk menyimpan, merepresentasikan, serta mengorganisasikan dan mengakses sebuah kebutuhan informasi. Kebanyakan cara yang digunakan untuk melakukan *information retrieval* adalah dengan menggunakan *keyword* yang ingin dicari, lalu dengan *keyword* tersebut, dibandingkan dengan isi dokumen, setelah itu dihasilkanlah dokumen-dokumen yang relevan dan tidak relevan.

Pada sistem *information retrieval*, terdapat banyak model yang dapat digunakan. Dengan banyaknya model-model tersebut untuk melakukan *retrieval* informasi, semakin banyak pula pertimbangan yang dilakukan untuk memilih dokumen yang tepat dengan implementasi *information retrieval* yang diinginkan peneliti. Cara yang paling mudah untuk mendapatkan hasil *retrieval* yang bagus adalah dengan menggabungkan semua fitur yang ada. Namun jika hal tersebut dilakukan, akan memakan waktu yang lebih lama dari proses *indexing* dan *retrieval*-nya itu sendiri.

Secara garis besar, permasalahan yang terjadi pada saat ini dalam masalah data adalah sebagai berikut:

1. Jumlah dokumen digital semakin bertambah dari segi kuantitas.
2. Isi dokumen digital yang semakin banyak, sehingga diperlukan metode paling efektif untuk mengatur dan me-*retriev* kembali data yang telah disimpan.

⁴ Christopher D Manning, *An Introduction to Information Retrieval*, (Cambridge: Cambridge University Press, 2009), h.1.

3. Kesalahan dalam pencarian karena penggunaan metode yang digunakan tidak sesuai.

Untuk menyelesaikan masalah-masalah di atas, maka diperlukan sebuah proses *information retrieval* yang sesuai agar pengguna dapat menemukan dokumen yang relevan sesuai dengan keinginan dan kebutuhannya. Jika menggunakan metode yang tepat, *information retrieval* akan sangat membantu dalam pencarian dokumen dengan kuantitas yang besar karena waktu akan jauh lebih efisien dan mengurangi tingkat kesalahan dalam pengembalian informasi.

2.1.1.2. Text Mining

Text mining atau yang biasa disebut dengan *Text Data Mining* (TDM) merupakan suatu proses pengambilan informasi dari teks yang terdapat di dalamnya⁵. Dengan *text mining*, dapat dicari kata-kata yang dapat mewakili isi dari suatu dokumen, lalu ditentukan kategorinya berdasarkan frekuensi kata-kata yang terdapat di dalamnya. Setelah suatu dokumen dilakukan analisis, maka akan muncul kategori olahraga, kesehatan, selebriti, kriminal, ekonomi, politik atau yang lain, dicocokkan dengan database kata kunci yang sebelumnya telah dibuat.

Klasifikasi dilakukan berdasarkan pembelajaran dari kumpulan dokumen untuk mendapatkan suatu pola tiap *class*. Pola dapat berupa suatu *rule* atau aturan. Pembelajaran untuk mendapatkan pola atau kriteria keputusan suatu *class* oleh komputer dilakukan dengan cara “mempelajari” secara otomatis dari data latih (*data train*). Diperlukan sejumlah dokumen (*training document*) yang sangat baik untuk tiap *class*. Semua algoritma klasifikasi dalam *text mining* mewakili

⁵ Ashok Srivastava & Mehran Sahami, *Text Mining: Classification, Clustering, and Applications*, (Florida: Chapman & Hall CRC Press, 2009), h.12

dokumen dalam suatu ruang dimensi yang tinggi. Untuk mengefisienkan, diperlukan pengurangan dimensi yang disebut dengan *feature selection*. *Feature selection* adalah proses pemilihan sebuah subset term yang muncul dalam himpunan training⁶. Klasifikasi teks hanya akan menggunakan hasil *feature selection*. Alasannya adalah agar metode pengklasifikasian lebih efisien dengan mengurangi ukuran *vocabulary* serta meningkatkan akurasi klasifikasi dengan membuang *feature noise*.

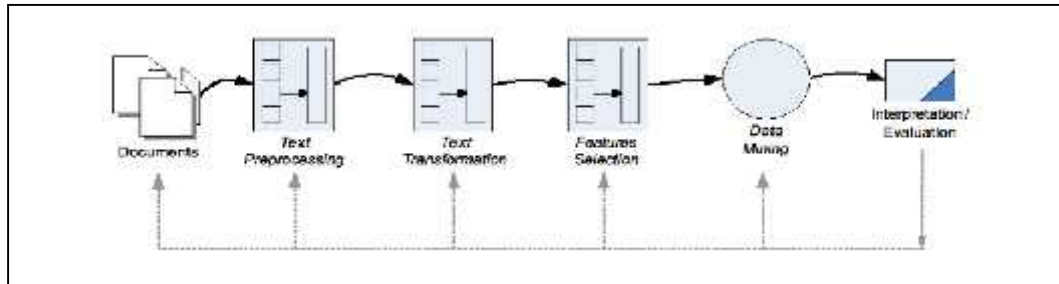
Proses *text mining* untuk data dokumen atau teks memerlukan lebih banyak tahapan, mengingat data teks memiliki karakteristik yang lebih kompleks daripada data biasa. Menurut Loretta Auvil dan Duane Sears Smith dari University of Illinois, karakteristik dokumen teks adalah:

1. Database teks yang berukuran besar,
2. Memiliki dimensi yang tinggi, yakni satu kata merupakan satu dimensi,
3. Mengandung kumpulan kata yang saling terkait (frase) dan antara kumpulan kata satu dengan lain dapat memiliki arti yang berbeda,
4. Banyak mengandung kata ataupun arti yang bias (*ambiguity*),
5. Dokumen email merupakan dokumen yang tidak memiliki struktur bahasa yang baku, karena di dalamnya terkadang muncul istilah seperti “r u there?” dan sebagainya.

Berdasarkan ketidakteraturan struktur data teks, maka proses *text mining* memerlukan beberapa tahap awal yang akan mengubah teks tersebut ke dalam

⁶ Sergious Theoridis & Konstantinos Koutroumbas, *Pattern Recognition: Fourth Edition*, (Oxford: Academic Press, 2009), h.261

bentuk yang jauh lebih terstruktur. Proses *text mining* dapat dijelaskan dengan gambar berikut.



Gambar 2.1. Proses Text Mining

Berikut ini adalah penjelasan tahap demi tahap proses *Text Mining* yang telah di gambarkan pada Gambar 2.1.

- a. *Documents*: dokumen dengan format .doc dan .docx yang tersedia untuk proses *text mining*.
- b. *Text Pre-processing*: sebelum divektorkan, pada bagian ini terlebih dahulu data diproses dengan tahapan: *case folding*, *tokenizing*, *filtering*, dan *stemming*. Inti dari proses ini adalah untuk mempersiapkan agar teks dapat diubah menjadi lebih terstruktur.
- c. *Text Transformation*: dimensi kamus yang telah dihasilkan dari token-token tentu akan berukuran sangat besar, sehingga perlu dilakukan proses transformasi untuk mengurangi ukuran dimensinya. Proses transformasi yang dapat diterapkan antara lain: *stopwords*, *frequent words*, dan pengurangan token (*Stemming*).
- d. *Feature Selection*: proses pemilihan subset dari istilah yang terjadi pada data latih dan menggunakan subset tersebut sebagai *feature* pada

klasifikasi teks. Ada tiga jenis metode yang bisa digunakan pada *feature selection* ini, yaitu⁷:

- 1) *Feature selection tipe wrapper*, yaitu melakukan *feature selection* bersamaan dengan pelaksanaan pemodelan. *Selection* tipe ini menggunakan suatu kriteria yang memanfaatkan rata-rata klasifikasi dari metode pengklasifikasi/pemodelan yang digunakan. Untuk mengurangi biaya komputasi, proses pemilihan umumnya dilakukan dengan memanfaatkan rata-rata klasifikasi dari metode pengklasifikasi/pemodelan untuk pemodelan dengan nilai terendah (misalnya dalam KNN, menggunakan nilai K terendah).
 - 2) *Feature selection tipe filter*: *feature selection* pada tipe ini hampir sama dengan tipe *wrapper* yaitu dengan menggunakan *intrinsic statistical properties* dari data. Perbedaannya, pada tipe *filter* pengkajian *feature* tidak dilakukan secara bersamaan dengan pemodelan.
 - 3) *Feature selection tipe embedded*: *feature selection* pada jenis ini memanfaatkan suatu *learning machine* dalam proses *feature selection*.
- e. *Data Mining*: setelah dokumen melewati proses untuk menjadikannya lebih terstruktur dari sebelumnya, maka pada proses ini dokumen tersebut digali informasi yang terdapat di dalamnya (pada kasus klasifikasi, perhitungan *mining* sesuai dengan metode yang sedang digunakan).

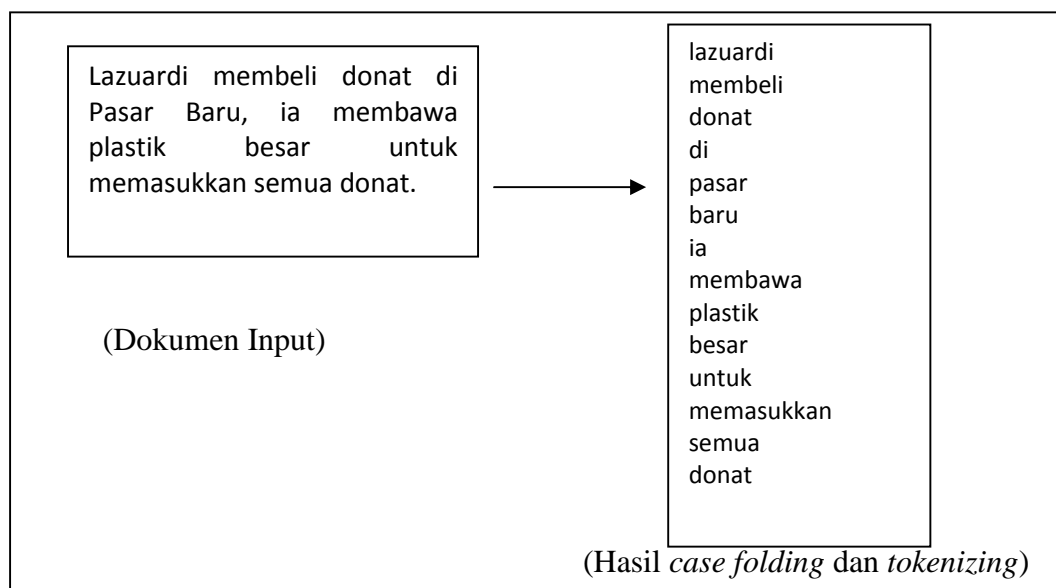
⁷ Yudi Agusta, *Feature Selection*, <http://yudiagusta.wordpress.com/2009/01/13/feature-selection/comment-page-1/>, 28 Mei 2014

- f. *Interpretation/Evaluation*: dari informasi yang digali pada proses *data mining*, maka selanjutnya didapatkan hasil akhir dari keseluruhan proses untuk melihat apakah hasil tersebut cocok dengan aplikasi yang telah dibuat, atau hasilnya tidak memuaskan namun signifikan.

2.1.1.3. Ekstraksi Dokumen

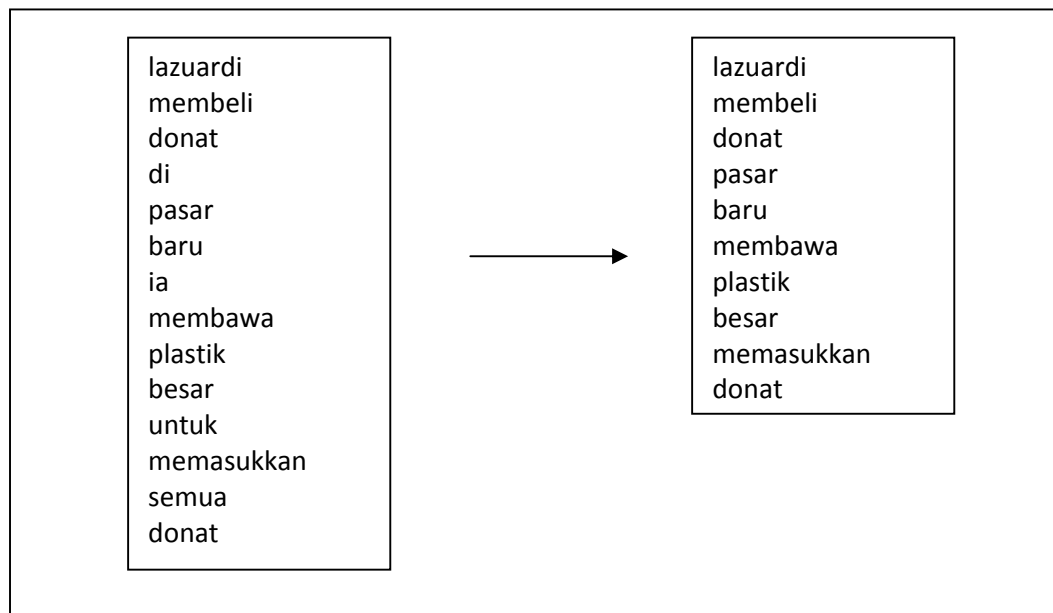
Sebuah dokumen digital mempunyai ketidakstrukturan teks karena dimensinya yang tinggi. Sebelum mengubahnya ke dalam bentuk yang jauh lebih terstruktur diperlukan proses *text mining* terhadap dokumen tersebut. Secara umum proses *text mining* melewati tahap: *Case folding*, *tokenizing*, *filtering*, dan *stemming*.

Case folding adalah mengubah seluruh huruf yang terdapat di dalam dokumen menjadi huruf kecil (*lowercase*). Sedangkan tahap *tokenizing* adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya. Proses *case folding* dan *tokenizing* dapat dilihat pada gambar berikut.



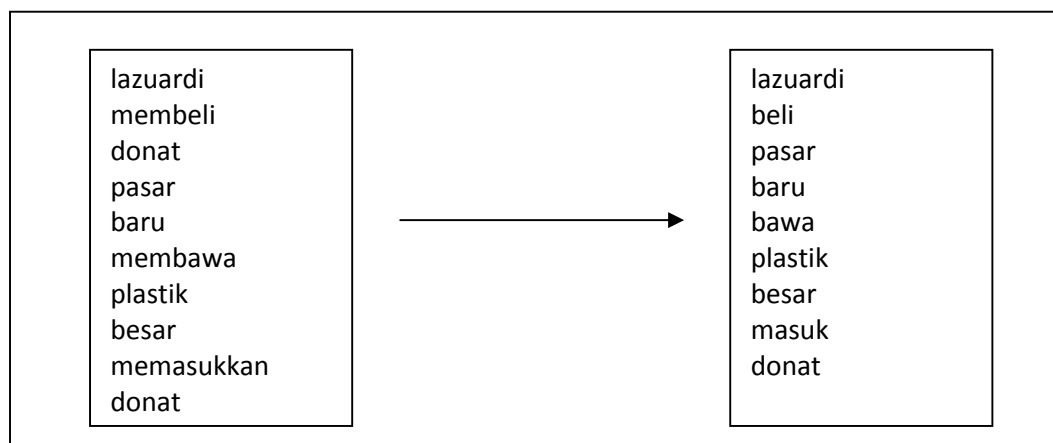
Gambar 2.2. Proses Case Folding dan Tokenizing

Filtering adalah tahap mengambil kata-kata penting dari hasil token. Bisa menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). Proses filtering dapat dilihat pada gambar di bawah ini:



Gambar 2.3. Proses Filtering

Stemming adalah tahap mencari root kata atau kata dasar dari tiap kata hasil filtering. Penghilangan duplikat pada kata-kata di dalam dokumen juga penting. Contoh hasil proses stemming dapat dilihat pada gambar di bawah ini:



Gambar 2.4. Proses Stemming

Stemming merupakan suatu proses yang terdapat dalam sistem IR yang mentransformasikan kata-kata yang terdapat dalam suatu dokumen ke kata-kata akar atau dasarnya (*root word*) dengan menggunakan aturan-aturan tertentu. Contohnya, kata melampaui, terlampaui, dilampaui akan di-*stemming* ke *root word*-nya, yaitu: lampau.

Stemming pada klasifikasi dokumen abstrak karya akhir menggunakan algoritma *Enhanced Confix Stripping Stemmer* oleh I Putu Adhi Kerta dkk. *ECS Stemmer* merupakan perbaikan dari algoritma *Confix Stripping Stemmer* oleh Jelita Asia. Algoritma *stemming* untuk Bahasa Indonesia pertama adalah algoritma Nazief-Adriani oleh Bobby Nazief dan Mirna Adriani pada tahun 1996. Algoritma *stemming* Nazief-Adriani untuk Bahasa Indonesia kemudian dilakukan perbaikan menjadi *CS Stemer* pada tahun 2007, kemudian *ECS Stemmer* pada tahun 2008. Berikut adalah daftar aturan pemenggalan imbuhan yang telah dilakukan perbaikan pada algoritma *Enhanced Confix Stripping Stemmer*.

Tabel 2.1. Aturan Pemenggalan Imbuhan *Enhanced Confix Stripping Stemmer*⁸

Aturan	Format Kata	Pemenggalan
1	berV ...	ber-V ... be-rV ...
2	berCAP ...	ber-CAP ... dimana C!=‘r’ & P!=‘er’
3	berCAerV ...	ber-CaerV ... dimana C!=‘r’
4	belajar	bel-ajar
5	beC ₁ erC ₂ ...	be-C ₁ erC ₂ ... dimana C ₁ !={ ‘r’ ‘l’ }
6	terV ...	ter-V ... te-rV ...
7	terCerV ...	ter-CerV ... dimana C!= ‘r’
8	terCP ...	ter-Cp ... dimana C!=‘r’ dan P!=‘er’
9	teC ₁ erC ₂ ...	te-C ₁ erC ₂ ... dimana C ₁ !=‘r’

⁸ Andita Dwi Yoga, *Implementasi Modifikasi Enhanced Confix Stripping Stripping Stemmer untuk Bahasa Indonesia dengan Metode Corpus Based Stemming: Seminar Tugas Akhir*, <http://digilib.its.ac.id/public/ITS-Undergraduate-14255-presentationpdf.pdf>, 10 Maret 2014

10	me {l r w y} V ...	me-{l r w y}V ...
11	mem{b f v} ...	mem-{b f v} ...
12	mempe ...	mem-pe ...
13	mem{rV V} ...	me-m{rV V}... me-p{rV V}...
14	men{c d j s z}...	men-{c d j s z}...
15	menV...	me-nV... me-tV
16	meng{g h q k}...	meng-{g h q k}...
17	mengV ...	meng-V ... meng-kV ... (mengV-... jika V='e')
18	menyV ...	meny-sV...
19	mempA ...	mem-pA ... dimana A!='e'
20	pe{w y}V ...	pe-{w y}V ...
21	perV ...	per-V ... pe-rV ...
Aturan	Format Kata	Pemenggalan
23	perCAP	per-CAP ... dimana C!='r' dan P!='er'
24	perCAerV ...	per-CAerV ... dimana C!='r'
25	pem{b f V} ...	pem-{b f V}...
26	pem{rV V} ...	pe-m{rV V} ... pe-p{rV V} ...
27	pen{c d j z} ...	pen-{c d j z} ...
28	penV ...	pe-nV ... pe-tV ...
29	pengC ...	peng-C ...
30	pengV ...	peng-V ... peng-kV ... (pengV- ... jika V='e')
31	penyV ...	peny-sV ...
32	peIV ...	pe-IV ... kecuali "pelajar" yang menghasilkan "ajar"
33	peCerV ...	per-erV ... dimana C!={r w y l m n}
34	peCP ...	pe-CP... dimana C!={r w y l m n} dan P!='er'
35	terC ₁ erC ₂ ...	ter-C ₁ erC ₂ ... dimana C ₁ !='r'
36	peC ₁ erC ₂ ...	pe-C ₁ erC ₂ ... dimana C ₁ !={r w y l m n}

Keterangan Tabel 2.1.:

C : huruf konsonan

V: huruf vocal

A: huruf vocal atau konsonan

P: partikel atau fragmen dari suatu kata, misalnya "er"

2.1.2. Data Mining

2.1.2.1. Pengertian Data Mining

Menurut Jiawei Han dan Michelin Kamber “*Data mining refers to extracting or “mining” knowledge from large amounts of data.*”⁹. Dari pernyataan tersebut dapat diartikan bahwa *Data Mining* merujuk pada suatu ilmu yang berkenaan dengan mengekstrak atau “menambang” pengetahuan dalam jumlah yang sangat besar.

Hampir sama dengan Jiawei Han dan Michelin Kamber. Pang-Ning Tan, Michael Steinbach, dan Vipin Kumar mengartikan “*Data mining is the process of automatically discovering useful information in large data repositories. Data mining techniques are deployed to scour large databases in order to find novel and useful patterns that might otherwise remain unknown*”¹⁰. Berdasarkan penjelasan tersebut dapat diartikan bahwa *data mining* merupakan proses otomatis untuk mencari informasi dalam suatu tempat penyimpanan data yang cukup besar. Teknik *data mining* juga digunakan untuk menemukan pola baru dan berguna untuk mengetahui pengetahuan yang belum diketahui sebelumnya.

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam menyimpan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.¹¹

⁹ Jiawei Han & Michelin Kamber, *Data Mining: Concepts and Technique, Second Edition*, (San Fransisco: Morgan Kaufman Publisher, 2006), h.5

¹⁰ Pang-Ning Tan, Michael Steinbach, & Vipin Kumar, *Introduction to Data Mining*, (Boston: Pearson Education, Inc, 2006), h.2

¹¹ Daniel T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*. (New Jersey: John Willey & Sons, Inc., 2005), h.2

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database* dan *data mining* juga merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar.¹²

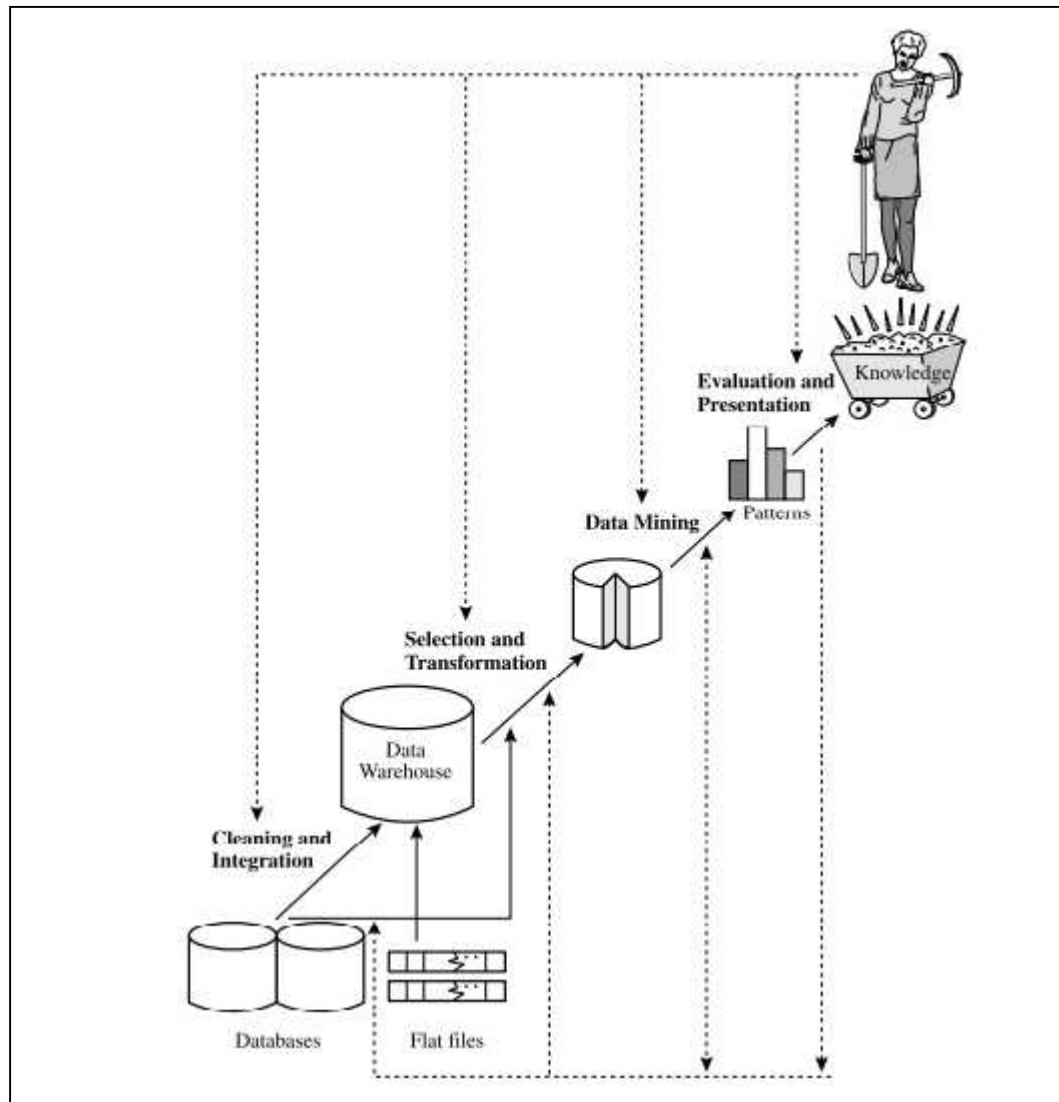
Berdasarkan pada definisi-definisi yang telah disebutkan oleh para ahli di atas, hal penting yang terkait dengan *data mining* adalah¹³:

- a. *Data mining* merupakan suatu proses otomatis terhadap data yang sudah ada, tetapi proses tradisional pun masih digunakan.
- b. Data yang akan diproses berupa data yang sangat besar.
- c. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Beberapa orang menyamakan *data mining* dengan *Knowledge Discovery from Data* (KDD). Karena yang sebenarnya *data mining* merupakan salah satu bagian dari proses KDD. Berikut adalah gambar proses dalam KDD yang di dalamnya terdapat proses *data mining* (Gambar 2.5).

¹² Kusriani & Emha Taufiq Luthfi, *Algoritma Data Mining*, (Yogyakarta: Penerbit Andi, 2009), h.3

¹³ *Ibid*, h.4

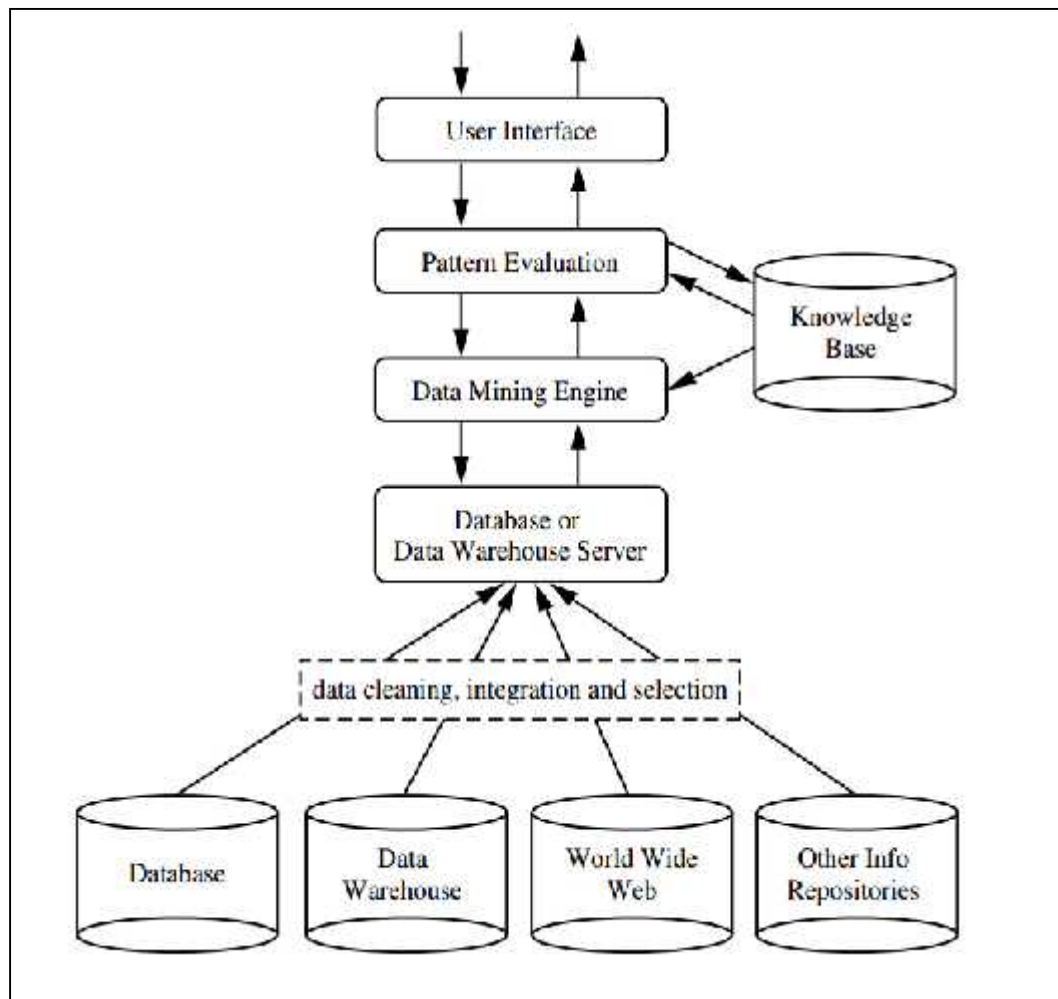


Gambar. 2.5. Data Mining Sebagai Bagian dari Proses Knowledge Discovery from Data¹⁴

Data mining yang merupakan salah satu bagian dari proses KDD (*Knowledge Discovery from Data*). Yang sumber datanya dapat berasal dari “*databases, data warehouse, or other information repositories*”.¹⁵

¹⁴ Jiawei Han & Michelin Kamber, Op. Cit, h.6

¹⁵ *Ibid*, h.7



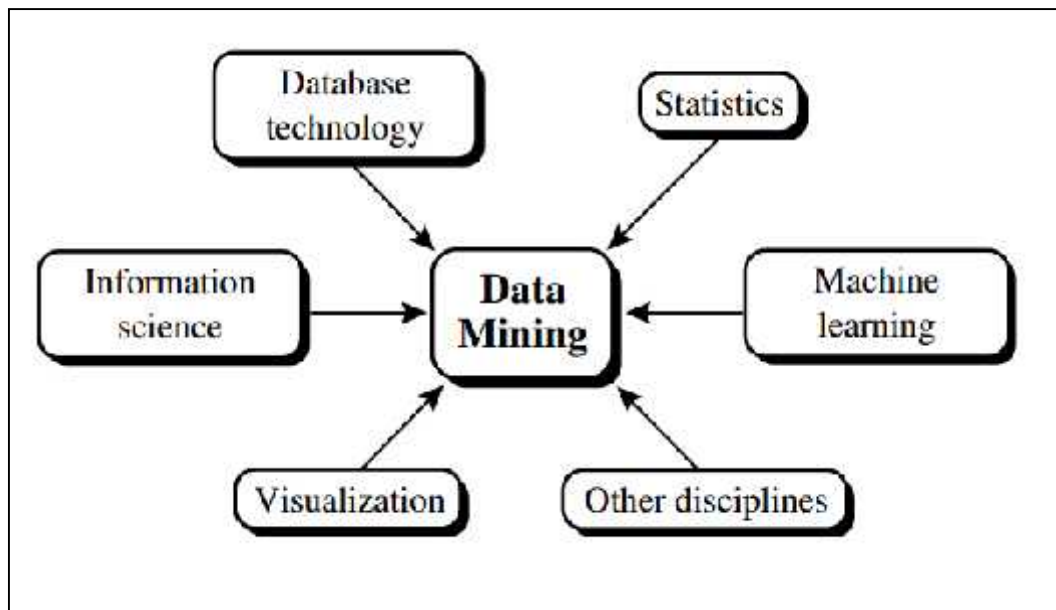
Gambar 2.6. Arsitektur Data Mining¹⁶

Berikut adalah penjelasan yang ada pada Gambar 2.6, yaitu:

- a. *Database, data warehouse, World Wide Web, or other information repository*: Merupakan satu atau beberapa *database, data warehouse, spreadsheets*, atau jenis lain dari gudang informasi.
- b. *Database or data warehouse server*: Database server dan *data warehouse server* bertanggung jawab untuk memberikan data yang relevan, berdasarkan pada permintaan pengguna *data mining*.

¹⁶ *Ibid*, h.8

- c. *Knowledge base*: Merupakan domain pengetahuan yang digunakan untuk memandu pencarian atau mengevaluasi hasil pola yang menarik.
- d. *Data mining engine*: *Data mining engine* sangat penting dalam sistem *data mining* dan idealnya terdiri dari satu set modul fungsi untuk tugas seperti karakterisasi, asosiasi dan analisis korelasi, klasifikasi, prediksi, analisis cluster, analisis *outlier*, dan analisis evolusi.
- e. *Pattern evaluation module*: *Pattern evaluation module* biasanya menggunakan langkah-langkah yang menarik dan interaktif dalam modul *data mining* sehingga fokus pencariannya pada pola yang menarik dapat diketahui.
- f. *User interface*: *User interface* merupakan komunikasi antara pengguna (*user*) dan sistem *data mining*, memungkinkan pengguna untuk berinteraksi dengan sistem untuk menentukan *query* atau tugas *data mining*, menyediakan informasi untuk membantu fokus pencarian, dan melakukan eksplorasi *data mining*



Gambar 2.7. Data mining sebagai pertemuan antar berbagai disiplin ilmu¹⁷

Menurut Jiawei Han dan Michelin Kamber dalam bukunya yang berjudul *Data Mining Concept and Techniques* menyebutkan bahwa dalam *data mining* melibatkan beberapa disiplin ilmu yaitu “*database and data warehouse technology, statistics, machine learning, high-performance computing, pattern recognition, neural networks, data visualization, information retrieval, image and signal processing, and spatial or temporal data analysis*”¹⁸. Dalam semua bidang tersebut sangat berperan dalam terbentuknya *data mining*. Sehingga *data mining* tidak bisa lepas dari bidang-bidang ilmu tersebut. Seperti ditunjuk pada Gambar 2.7.

Menurut Jiawei Han dan Michelin Kamber, tugas utama *data mining* secara umum dapat dibagi menjadi dua kategori yaitu: *descriptive* dan *predictive*. “*Descriptive mining tasks characterize the general properties of the data in the*

¹⁷ *Ibid*, h.29

¹⁸ *Ibid*, h.9

database. Predictive mining tasks perform inference on the current data in order to make predictions".¹⁹

2.1.2.2. Data

Menurut Pang-Ning Tan, Michael Steinbach, dan Vipin Kumar data adalah kumpulan objek data dan atribut data tersebut. Atribut sendiri merupakan properti atau karakteristik dari objek. Dan atribut juga dikenal sebagai variabel, *field*, karakteristik, atau fitur. Contohnya warna pada mobil, warna pada baju, dll.

2.1.2.3. Klasifikasi Dokumen

Klasifikasi merupakan salah satu tugas penting dalam *data mining*. Pada klasifikasi kelompok data (*class label*) yang sudah diketahui, sebuah data akan masuk ke dalam kelompok tertentu yang sebelumnya telah ditentukan. Setiap hari, jumlah dokumen semakin bertambah. Diantara berbagai bentuk informasi digital, diperkirakan 80% dokumen digital adalah dalam bentuk teks²⁰. Tingginya volume dokumen teks ini dikarenakan aktivitas yang terus meningkat dari berbagai sumber berita dan aktivitas penulisan dokumen akademis dari kegiatan riset, konferensi dan pertemuan-pertemuan ilmiah.

Oleh karena itu, klasifikasi dokumen merupakan masalah yang mendasar namun sangat penting karena manfaatnya dapat mengatasi permasalahan yang telah disebutkan sebelumnya. Sebuah dokumen dapat dikelompokkan ke dalam kategori tertentu berdasarkan kata-kata atau kalimat-kalimat yang ada di dalam dokumen tersebut. Kata atau kalimat yang ada di dalam dokumen memiliki makna

¹⁹ *Ibid*, h.21

²⁰ Ah-Hwee Tan, "Text Mining: The state of the art and the challenges", (Singapore, 1999), h.1

tertentu dan dapat digunakan untuk menentukan kategori dari dokumen tersebut. Performa pengklasifikasi biasanya diukur dan dinyatakan dengan galat.

Perhatikan contoh kalimat berikut:

1. Paling tidak akan ada 28 cabang olahraga yang akan dipelatinaskan hingga SEA Games 2015. Namun hanya 18 cabor yang dipersiapkan untuk Asian Games. [**Sport, Detik 20 Januari 2014**]
2. Kaveri adalah generasi terbaru chip tersebut. Performanya terbilang cukup mentereng, AMD mengklaim versi tertinggi chip tersebut yakni A10-8570K bisa dipakai memainkan Battlefield 4 dengan nyaman pada level medium di resolusi 720p. *Wow!* Luar biasa untuk ukuran GPU onboard. [**Teknologi, Detik 20 Januari 2014**]
3. Dengan dihapusnya peran guru inti dalam implementasi kurikulum 2013, Nuh berharap guru sasaran akan lebih paham konsep dan materi mengenai kurikulum tersebut. Pelatihan guru sasaran akan dilakukan langsung oleh instruktur nasional. [**Edukasi, Kompas 14 Januari 2014**]

Pada kalimat nomer 1, terdapat kata olahraga dan Asian Games. Kata-kata tersebut memiliki keterkaitan dengan masalah olahraga, sehingga dapat disimpulkan kalimat tersebut berkategori olahraga. Kata tersebut tidak dapat langsung dijadikan acuan, untuk melihat relevansi dokumen tersebut terhadap kategori, haruslah dilihat seluruh kata-kata yang terdapat di dalam dokumen tersebut.

Manfaat dari klasifikasi dokumen adalah untuk mengorganisasikan dokumen. Dengan semakin meningkatkan jumlah dokumen yang bertambah

setiap harinya, maka akan lebih mudah mencari informasi dari sebuah dokumen yang telah terorganisasi dan telah dikelompokkan menurut kategorinya masing-masing. Contoh aplikasi penggunaan klasifikasi dokumen teks yang banyak digunakan adalah *e-mail spam filtering*. Pada aplikasi *spam filtering* sebuah *e-mail* diklasifikasikan apakah *e-mail* tersebut termasuk *spam* atau tidak dengan memperhatikan kata-kata yang terdapat di dalam *e-mail* tersebut. Aplikasi ini telah digunakan oleh banyak *e-mail provider*.

Dokumen karya akhir mahasiswa pada penelitian ini adalah dokumen yang berasal dari isi abstrak skripsi mahasiswa jenjang S1 Jurusan Teknik Elektro UNJ sebagai bahan penelitian. Sedangkan klasifikasi dokumen karya akhir mahasiswa pada penelitian ini adalah mengklasifikasikan atau mengkategorikan isi abstrak skripsi jenjang S1 Jurusan Teknik Elektro UNJ berbentuk dokumen dengan format .doc atau .docx.

2.1.2.4. Naïve Bayes Classifier

Naïve bayes merupakan algoritma data mining untuk klasifikasi yang tidak menggunakan *rules* maupun *decision tree*. Karena itulah algoritma *Naïve Bayes* masuk ke dalam kelompok algoritma klasifikasi *non-rule based classification*.²¹

Naïve Bayes Classifier memilih klasifikasi yang paling mungkin (ditandai dengan variable V_{nb}) dan diberikan atribut yang bernilai a_1, a_2, \dots, a_n . Dituliskan rumus *clasisfier* sebagai berikut:

$$V_{nb} = \operatorname{argmax}_{v_j \in V} P(v_j \prod P(a_i | v_j))$$

²¹ Max Brammer, *Principles of Data Mining*, (London: Springer, 2007), h.24.

Pertimbangkan data set berikut ini. Ada 6 dokumen dimulai dari D0 ... D5 sebagai data set pelatihan. Misalnya, dari seluruh dokumen kata yang diambil dan dipertimbangkan hanya 6 kosakata: “kill”, “bomb”, “kidnap”, “music”, “movie”, dan “tv”. Ada dua kelas (kategori) dokumen: “Terrorism” dan “Entertainment”. Dokumen pemrosesan awal ditunjukkan dalam tabel berikut. Angka-angka dibawah ini menunjukkan frekuensi kata dalam dokumen. Sebagai contoh, kata “kill” ada dua kali dalam dokumen D0.

Tabel 2.2. Frekuensi Kata dalam Dokumen Data Set

Doc Latih	kill	bomb	kidnap	music	movie	tv	Class
D0	2	1	3	0	0	1	Terrorism
D1	1	1	1	0	0	0	Terrorism
D2	1	1	2	0	1	0	Terrorism
D3	0	1	0	2	1	1	Entertainment
D4	0	0	1	1	1	0	Entertainment
D5	0	0	0	2	2	0	Entertainment

Tahap 1: Membangun Model Naïve Bayes

$|V|$ = jumlah kosakata

$$P(C_i) = \frac{\text{Jumlah dari dokumen di dalam kelas}}{\text{Jumlah seluruh dokumen}}$$

$P(C_i)$ = prior probabilitas untuk setiap kelas

N_i = total frekuensi kata dari tiap kelas

$$N_{\text{terrorism}} = 2+1+3+1+1+1+1+1+1+2+1 = 15$$

$$N_{\text{entertainment}} = 1+2+1+1+1+1+1+2+2=12$$

$P(W_i|C_i)$ = probabilitas bersyarat dari terjadinya kata kunci diberikan sebuah kelas.

Sebagai contoh:

$$P(\text{kill} | \text{Terrorism}) = \frac{2+1+1}{15} = \frac{4}{15}$$

$$P(\text{kill} | \text{Entertainment}) = \frac{0+0+0}{15} = \frac{0}{12}$$

Untuk menghindari masalah “Zero Frequency”, sebaiknya menggunakan Laplace Estimator dengan rumus sebagai berikut:

$$P(X_{n+1} = 1 | X_1 + \dots + X_n = s) = \frac{s + 1}{n + 2}$$

Dengan asumsi distribusi seragam atas semua kata-kata sebagai berikut:

$$P(\text{kill} | \text{Terrorism}) = \frac{2+1+1+1}{15+|v|} = \frac{5}{21} = 0,2380$$

$$P(\text{kill} | \text{Entertainment}) = \frac{0+0+0+1}{12+|v|} = \frac{1}{18} = 0,0555$$

$$P(\text{bomb} | \text{Terrorism}) = \frac{1+1+1+1}{15+|v|} = \frac{4}{21} = 0,1904$$

$$P(\text{bomb} | \text{Entertainment}) = \frac{1+1}{12+|v|} = \frac{2}{18} = 0,1111$$

$$P(\text{kidnap} | \text{Terrorism}) = \frac{3+1+2+1}{15+|v|} = \frac{7}{21} = 0,3333$$

$$P(\text{kidnap} | \text{Entertainment}) = \frac{1+1}{12+|v|} = \frac{2}{18} = 0,1111$$

$$P(\text{music} | \text{Terrorism}) = \frac{1}{15+|v|} = \frac{1}{21} = 0,0476$$

$$P(\text{music} | \text{Entertainment}) = \frac{2+1+2+1}{12+|v|} = \frac{6}{18} = 0,3333$$

$$P(\text{movie} | \text{Terrorism}) = \frac{1+1}{15+|v|} = \frac{2}{21} = 0,0952$$

$$P(\text{movie} | \text{Entertainment}) = \frac{1+1+2+1}{12+|v|} = \frac{5}{18} = 0,2777$$

$$P(\text{TV} | \text{Terrorism}) = \frac{1+1}{15+|v|} = \frac{2}{21} = 0,0952$$

$$P(\text{TV} | \text{Entertainment}) = \frac{1+1}{15+|v|} = \frac{2}{18} = 0,1111$$

Tabel 2.3. Hasil Perhitungan Probabilitas

C	P(kill Ci)	P(bomb Ci)	P(kidnap Ci)	P(music Ci)	P(movie Ci)	P(TV Ci)
1	0,2380	0,1904	0,3333	0,0476	0,0952	0,0952
2	0,0555	0,1111	0,1111	0,3333	0,2777	0,1111

Keterangan : C = Class. 1 = Terrorism. 2 = Entertainment

Tahap 2: Mengklasifikasikan Dokumen Tes

Untuk mengklasifikasikan dokumen tes D_t , hal yang harus dilakukan adalah mengkalkulasikan “posterior” probabilitas, $P(C_i | W)$ untuk setiap kelas berikut ini:

Tabel 2.4. Data Sebuah Dokumen Tes

Test Doc	kill	bomb	kidnap	music	movie	TV	Class
Dt	2	1	2	0	0	1	?

$$P(c_i | W) = P(c_i) \times \prod_{j=1}^n P(w_j | c_i)$$

$$P(\text{Terrorism} | W) = P(\text{Terrorism}) \times P(\text{kill} | \text{Terrorism}) \times P(\text{bomb} | \text{Terrorism}) \times \\ P(\text{kidnap} | \text{Terrorism}) \times P(\text{music} | \text{Terrorism}) \times P(\text{movie} | \\ \text{Terrorism}) \times P(\text{TV} | \text{Terrorism})$$

$$= 0,5 \times 0,2380^2 \times 0,1904^1 \times 0,3333^2 \times 0,0476^0 \times 0,0952^0 \times \\ 0,0952^1$$

$$= 0,5 \times 0,0566 \times 0,1904 \times 0,1110 \times 1 \times 1 \times 0,0952$$

$$= 5,7 \times 10^{-5}$$

$$P(\text{Entertainment} | W) = P(\text{Entertainment}) \times P(\text{kill} | \text{Entertainment}) \times P(\text{bomb} | \\ \text{Entertainment}) \times P(\text{kidnap} | \text{Entertainment}) \times P(\text{music} | \\ \text{Entertainment}) \times P(\text{movie} | \text{Entertainment}) \times P(\text{TV} | \\ \text{Entertainment})$$

$$= 0,5 \times 0,555^2 \times 0,1111^1 \times 0,3333^0 \times 0,2777^0 \times 0,1111^1$$

$$= 0,5 \times 0,0030 \times 0,1111 \times 0,0123 \times 1 \times 1 \times 0,1111$$

$$= 2,27 \times 10^{-7}$$

Karena $P(\text{Terrorism} | W)$ memiliki nilai tertinggi, oleh karena itu Dt diklasifikasikan menjadi "Terrorism". Ini masuk akal karena Dt mengandung

banyak kata yang berhubungan dengan terorisme seperti “kill”, “bomb” dan “kidnap”.

$$P(\text{Terorisme} | W) = \log(0.5 \times 0.23802 \times 0.19041 \times 0.33332 \times 0.04760 \times 0.09520 \times 0.09521)$$

$$= \log(0.5) + 2 \log(0.2380) + 1 \log(0.1904) + 2 \log(0.3333) + 0 \log(0.0476) + 0 \log(0.0952) + 1 \log(0.0952)$$

$$= -0.3010 - 1.2468 - 0.7203 - 0.9543 + 0 + 0 - 1.0213$$

$$= -4.2437$$

$$P(\text{Entertainment} | W) = \log(0.5 \times 0.05552 \times 0.11111 \times 0.11112 \times 0.33330 \times 0.27770 \times 0.11111)$$

$$= \log(0.5) + 2 \log(0.0555) + 1 \log(0.1111) + 2 \log(0.1111) + 0 \log(0.3333) + 0 \log(0.2777) + 1 \log(0.1111)$$

$$= -0.3010 - 2.511 - 0.9542 - 1.9085 + 0 + 0 - 0.9542$$

$$= -6.6289$$

Sekali lagi, karena $P(\text{Terorisme} | W)$ memiliki nilai yang lebih tinggi, karena itu D_t diklasifikasikan menjadi "Terorisme". Alternatif algoritma, dapat di jelaskan sebagai berikut:

$$\log(x*y) = \log(x) + \log(y) \text{ and } \log xy = y * \log(x)$$

2.1.2.5. Keandalan Naïve Bayes Classifier

Domingos dan Pazzani (1997) dalam *paper* nya yang berjudul “*On the Optimality of the Simple Bayesian Classifier under Zero-One Loss*” menjelaskan performa *Naïve Bayes* dalam fungsi *Zero-One Loss*. Fungsi *Zero-One Loss* ini mendefinisikan *error* hanya sebagai pengklasifikasi yang salah. Fungsi *Zero-One Loss* tidak memberi nilai suatu kesalahan perhitungan peluang selama peluang maksimum ditugaskan ke dalam kelas yang benar²². Ini berarti bahwa *Naïve Bayes* dapat mengubah peluang *posterior* dari tiap kelas, tetapi kelas dengan nilai peluang *posterior* maksimum jarang diubah. Sebagai contoh, diasumsikan peluang sebenarnya dari $P(+ | E) = 0,9$ dan $P(- | E) = 0,1$, sedangkan peluang yang dihasilkan oleh *Naïve Bayes* adalah $P(+ | E) = 0,6$ dan $P(- | E) = 0,4$. Nilai peluang tersebut berbeda jauh, namun pilihan kelas [+] tetap tidak berpengaruh.

Domingos dan Pazzani (1997) dalam *paper* nya juga menjelaskan:

“The Bayesian classifier may often be a better classifier than more powerful alternatives when the sample size is small, even in domains where its learning model is not the most appropriate one, and verified this by means of experiments in artificial domains. We also verified that searching for attribute dependences is not necessarily the best approach to improving the Bayesian classifier’s performance.” Yang dapat di artikan *Naïve Bayes Classifier* mungkin sering menjadi *Classifier* lebih baik dari alternatif yang lebih kuat ketika ukuran sampel kecil, bahkan dalam domain dimana model pembelajaran adalah bukan yang paling tepat, dan diverifikasi ini dengan cara eksperimen dalam domain buatan. Pada *paper* tersebut juga diverifikasi bahwa mencari ketergantungan

²² Pedro Domingos & Michael Pazzani, “On the Optimality of the Simple Bayesian Classifier under Zero-One Loss”, (Netherlands, 1997), h.25

atribut belum tentu pendekatan terbaik untuk meningkatkan kinerja *Classifier Bayesian*.

Aplikasi nyata dari penggunaan metode *Naïve Bayes Classifier* untuk klasifikasi dokumen dalam bidang teknologi informasi adalah penyaringan spam pada layanan surat elektronik (*e-mail*) yang biasa disebut dengan *spam filtering*. Spam adalah penyalahgunaan sistem pesan elektronik (termasuk media penyiaran dan sistem pengiriman digital) untuk mengirim berita iklan dan keperluan lainnya secara massal. Penyaringan surat elektronik spam (selanjutnya disebut “spam” saja) perlu dilakukan agar hanya informasi.

Metoda *Naïve Bayes classifier* merupakan metoda klasifikasi yang berdasar kepada teorema bayes, sebuah teorema yang terkenal di dalam bidang ilmu probabilitas. Selain itu, metoda ini turut didukung oleh ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Metoda ini sangat luas dipakai dalam berbagai bidang, khususnya dalam proses klasifikasi dokumen. Klasifikasi ini merupakan salah satu teknik dalam data mining yang merupakan kegiatan penunjang dalam bidang sistem informasi. Seperti halnya metoda-metoda lain, metoda *Naïve Bayes classifier* ini tidaklah 100% sempurna. Ada banyak kelebihan dan kekurangan dari metoda ini, yang dapat menjadi dasar bahan kajian lebih lanjut untuk mendapatkan atau mengembangkan metoda klasifikasi lain, yang dapat bekerja dengan lebih efektif dan efisien, serta mengurangi jumlah titik kelemahan yang dapat disalahgunakan oleh orang lain.

Metode *Naïve Bayes Classifier* dipilih karena *Naive bayesian filtering* memiliki kelebihan dibandingkan dengan metoda *filtering* yang lain, diantaranya adalah²³:

1. *Bayesian filter* memiliki komputasi yang mudah.
2. *Bayesian* memeriksa *email* secara keseluruhan yaitu memeriksa token di *database spam* maupun *legitimate*.
3. *Bayesian filtering* termasuk dalam *supervised learning* yaitu secara otomatis akan melakukan proses *learning* dari *email* yang masuk.
4. *Bayesian filtering* cocok diterapkan di level aplikasi *client/individual user*.
5. *Bayesian filtering* cocok diterapkan pada *binary class* yaitu klasifikasi ke dalam dua kelas.
6. Metode ini *multilingual* dan internasional. *Bayesian filtering* menggenerate token dengan pengenalan karakter sehingga mampu diimplementasikan pada *email* dengan bahasa apapun.

²³ M. Rachli, "Email Filtering Menggunakan Naive Bayesian", (Bandung: Institut Teknologi Bandung, 2007) h.23

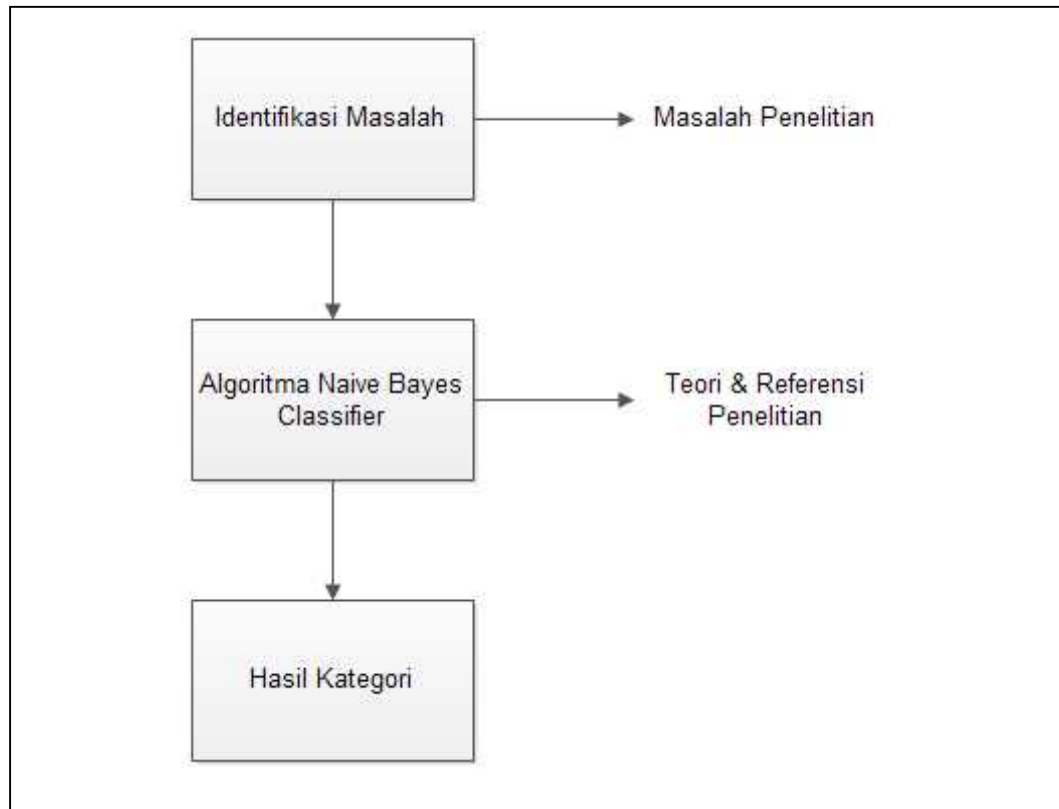
2.2. Kerangka Berpikir

Pengklasifikasian dokumen di Jurusan Teknik Elektro Universitas Negeri Jakarta saat ini masih dilakukan secara manual. Selain itu, dokumen karya kahir mahasiswa yang ada di Jurusan Teknik Elektro UNJ pun, setiap tahunnya semakin bertambah. Pencarian informasi akan memakan waktu yang lama jika membaca seluruh dokumen. Oleh karena itu, dibutuhkan solusi pengklasifikasi dokumen karya akhir. Untuk merealisasikan sebuah sistem yang mampu mengklasifikasikan dokumen abstrak karya akhir harus didasarkan prosedur yang diawali dengan mengidentifikasi masalah yang bertujuan untuk mengangkat masalah yang akan diteliti dan dicari solusinya.

Algoritma *Naïve Bayes Classifier* merupakan salah satu algoritma *data mining* untuk klasifikasi, serta merupakan algoritma yang tidak menggunakan *rules* ataupun *decision tree*. Selain itu, NBC memiliki komputasi yang mudah, serta memiliki tingkat akurasi tinggi dan *error rate* yang minimum. Algoritma NBC juga mampu menggenerasikan *token* dengan pengenalan karakter sehingga mampu diimplementasikan pada *token* dengan bahasa Indonesia. Sehingga cocok pada kasus dokumen karya akhir yang menggunakan bahasa Indonesia sebagai data penelitian.

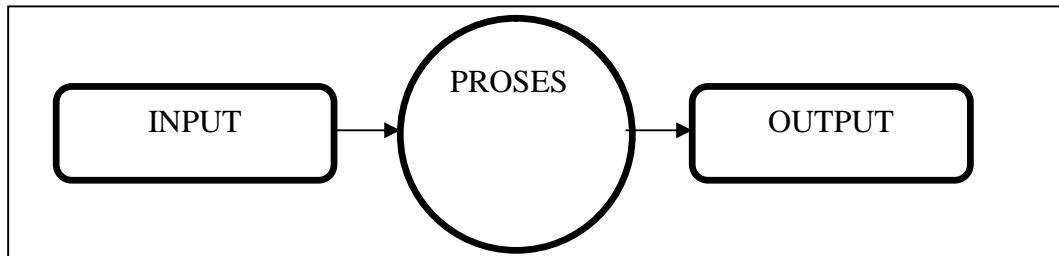
Untuk membantu *user* mengklasifikasikan dokumen karya akhir, diperlukan sebuah sistem klasifikasi dengan menerapkan algoritma *Naïve Bayes Classifier*. Keluaran yang akan dihasilkan oleh sistem adalah sebuah kategori yang dapat menggeneralisasikan isi abstrak sebagai data penelitian. Diharapkan dengan mengembangkan Sistem Klasifikasi Dokumen Karya Akhir, akan dapat mengorganisasikan serta mengklasifikasikan dokumen karya akhir dengan tingkat

akurasi yang baik. Bagan kerangka berpikir pada penelitian ini diilustrasikan pada Gambar 2.8.



Gambar 2.8. Bagan Kerangka Berpikir

Proses klasifikasi dokumen menggunakan algoritma *Naïve Bayes* selanjutnya dapat dilihat pada Gambar 2.9.



Gambar 2.9 Gambaran Proses Sistem Klasifikasi Dokumen Abstrak

Keterangan:

Input: Memilih dokumen abstrak yang akan di klasifikasikan

Proses:

1. Melakukan tahap *case folding*
2. Melakukan tahap *tokenizing*
3. Melakukan tahap *stemming*
4. Melakukan tahap *remove duplicate*, yaitu penghilangan duplikasi kata
5. Melakukan proses *analyzing* dengan *Naïve Bayes Classifier*

Output:

1. Menampilkan hasil akhir perhitungan algoritma
2. Menampilkan hasil waktu eksekusi
3. Menampilkan kategori dokumen: Media Pembelajaran, Penelitian Pendidikan, Produk Rekayasa, dan Analisa.

Setelah program klasifikasi dokumen abstrak karya akhir selesai dibuat, atau kemudian dapat disebut dengan sistem klasifikasi dokumen abstrak karya akhir. Kemudian, dilakukan pengujian terhadap sistem tersebut. Setelah semua proses pengujian telah dilalui, dan sistem tersebut dinyatakan layak, maka sistem tersebut dapat digunakan.

BAB III

METODOLOGI PENELITIAN

3.1. Tujuan Operasional Penelitian

Tujuan operasional dari penelitian adalah untuk mengembangkan aplikasi klasifikasi dokumen abstrak yang dapat digunakan untuk menemukan *class* atau kategori dokumen abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier*.

3.2. Tempat dan Waktu Penelitian

Penelitian dilakukan di Laboratorium Multimedia Jurusan Teknik Elektro Universitas Negeri Jakarta sejak bulan Maret 2014 hingga Juni 2014. data dokumen abstrak karya akhir mahasiswa Jurusan Teknik Elektro di ambil pada tanggal 27 Maret 2014.

3.3. Metode Penelitian

Dalam penelitian ini, menggunakan metode penelitian eksperimen dan untuk pengembangan perangkat lunak (software) menggunakan metode *waterfall* (sekuensial linier).

3.4. Instrumen Penelitian

Berikut ini adalah instrumen yang digunakan dalam penelitian:

1. Perangkat Keras

- a. Processor Intel® Core™ i5-2430M GeForce GT540M @ 2.4 GHz.
- b. Memori 4 GB DDR3
- c. Monitor 14"

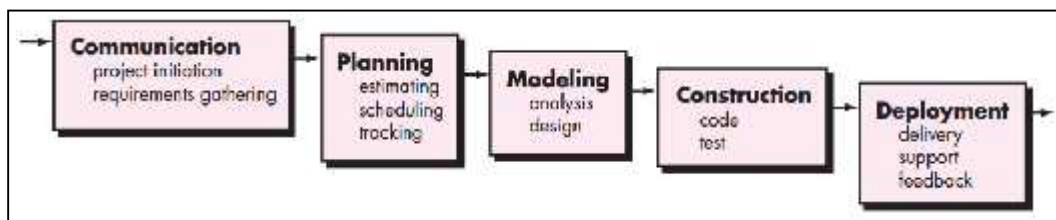
2. Perangkat Lunak

- a. Java *jdk-7u25-windows-x64*.
- b. IDE *netbeans-7.2-ml-windows*.
- c. *XAMPP for windows x64 1.8.2*
- d. *MySQL Database 5.5.32*
- e. *PHP 5.4.22*

3.5. Pelaksanaan Penelitian

3.5.1. Proses Pengembangan Software

Proses penerapan metode *waterfall* (sekuensial linier) akan di jelaskan pada Gambar 3.1 dan masing-masing tahapan dalam fase metode tersebut akan dijelaskan pada Tabel 3.1. Dalam pengembangan metode *waterfall*, definisi kebutuhan pada tahap *Communication* sangat penting, jika pada tahap tersebut tidak dilakukan dengan baik, maka tahap selanjutnya akan mengalami kesulitan.



Gambar 3.1. Metode *Waterfall*

Tabel 3.1. Tahap Pengembangan *Software*

No.	Tahap <i>Waterfall</i>	Proses yang dilakukan
1.	<i>Communication</i>	Identifikasi masalah yang akan dipecahkan dilakukan dengan cara mendefinisikan kebutuhan dan menguraikan deskripsi perangkat lunak.
2.	<i>Planning</i>	Membuat <i>time-table</i> serta membuat dokumentasi dari setiap proses.
3.	<i>Modeling</i>	Pembuatan <i>flow-chart diagram</i> , <i>Entity Relationship Diagram</i> untuk <i>database</i> serta pembuatan design GUI (<i>Graphics User Interface</i>).
4.	<i>Construction</i>	Pengkodean sistem perangkat lunak sesuai dengan design yang telah di rancang.
5.	<i>Deployment</i>	Mengoperasikan <i>software</i> , melakukan pemeliharaan <i>software</i> serta mendapatkan umpan balik dari <i>software</i> tersebut.

3.5.2. Definisi Kebutuhan

Pada tahap definisi kebutuhan ini penulis melakukan pengamatan terhadap banyaknya dokumen abstrak yang ada di Jurusan Teknik Elektro Universitas Negeri Jakarta. Aplikasi klasifikasi dokumen abstrak karya akhir bertujuan untuk menjawab akan kebutuhan *software* yang dapat mengotomatisasi proses pengklasifikasian atau pengkategorisasian dokumen abstrak karya akhir.

Kebutuhan fungsional harus mendefinisikan aksi dasar yang harus diambil oleh sistem untuk menerima dan memproses *input* berupa dokumen teks dengan format *.doc* dan *.docx* sehingga menghasilkan *output*. Kebutuhan fungsional tersebut meliputi *content* hingga bagian *administrator*.

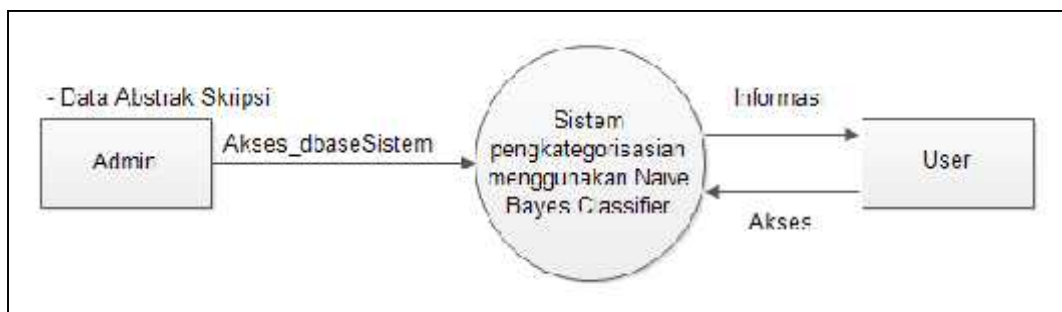
Tabel 3.2. Daftar Kebutuhan Fungsional

Aspek	Kebutuhan Fungsional
<i>Content</i>	Suatu aplikasi perangkat lunak untuk membantu dalam mengklasifikasikan dokumen karya akhir berdasarkan abstrak karya akhir pada sistem <i>Naïve Bayes Classifier</i> .
<i>User</i>	<i>User</i> dapat mengklasifikasikan karya akhir dengan menginputkan dokumen bertipe .doc dan .docx ke dalam sistem NBC.
<i>Admin</i>	Admin dapat mengedit data, <i>menginput</i> dan mengelola data karya akhir di dalam database sistem NBC
	Admin dapat memberikan hak akses kepada <i>User</i> yang ingin mengklasifikasikan karya akhir berdasarkan abstrak.

3.5.3. Desain dan Analisis Sistem

Pada tahapan desain dan analisis sistem, penulis merancang desain sistem dengan menggunakan *Data Flow Diagram* (DFD) serta *Entity Relationship Diagram* (ERD) sebagai rancangan sistem *database*. Data Flow Diagram dapat digunakan untuk merepresentasikan sebuah sistem atau perangkat lunak pada beberapa level abstraksi. DFD dapat dibagi menjadi beberapa level yang lebih detail untuk merepresentasikan aliran informasi atau fungsi yang lebih detail.

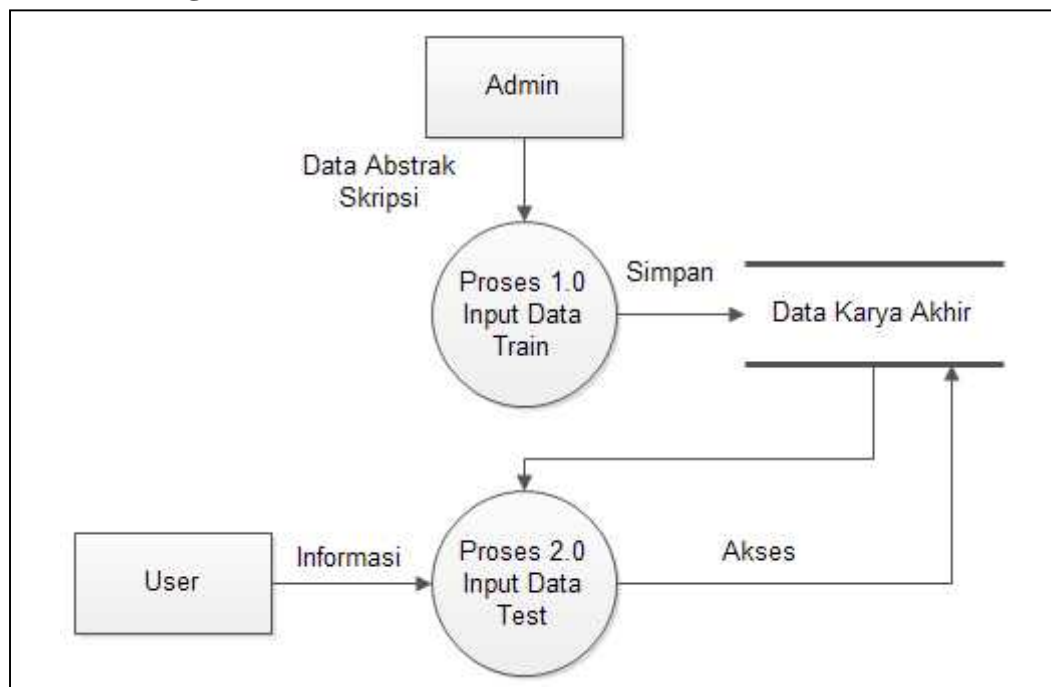
3.5.3.1. Diagram Konteks Usulan



Gambar 3.2. Diagram Konteks Usulan Sistem Pengklasifikasi Dokumen menggunakan *Naïve Bayes Classifier*

Diagram di atas menunjukkan Sistem pengkategorori dokumen menggunakan *Naïve Bayes Classifier*. Pada diagram konteks, terdapat dua entitas yaitu Admin dan User, dan terdapat 1 proses yaitu Sistem pengkategorisasian menggunakan *Naïve Bayes Classifier*. Terdapat 2 arus data yang saling berhubungan yaitu data informasi dan data karya akhir.

3.5.3.2. Diagram Nol Sistem



Gambar 3.3. Diagram Nol Sistem Pengklasifikasi Dokumen menggunakan *Naïve Bayes Classifier*

Keterangan:

a. Proses 1.0

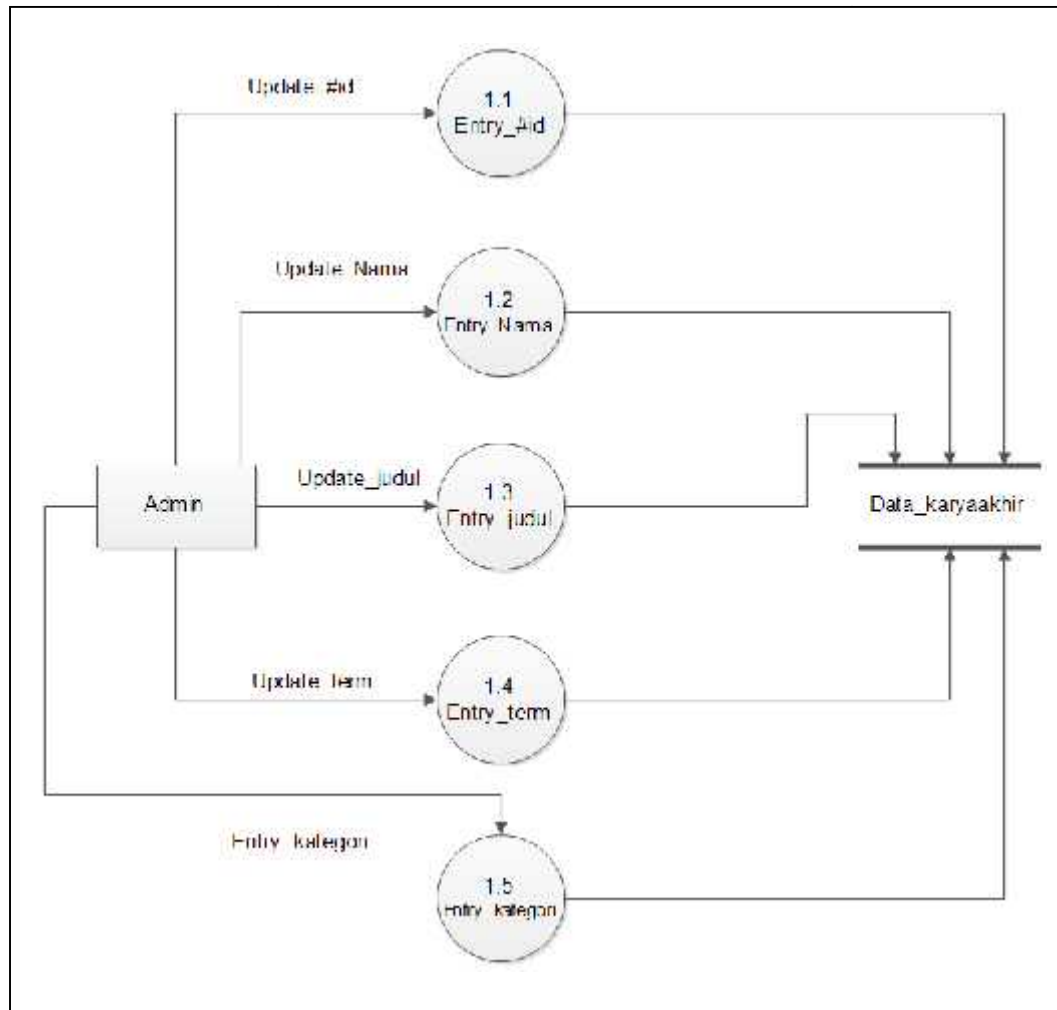
Pada proses 1.0 Pengolahan Data yang meliputi Data Abstrak Skripsi beserta informasi yang mengenai Data tersebut diolah oleh Admin dan disimpan di dalam database yang telah tersedia.

b. Proses 2.0

Proses 2.0 User dapat menginput sebuah dokumen abstrak bertipe .doc dan .docx sehingga, dokumen tersebut dapat diproses selanjutnya untuk kemudian diketahui kategori yang sesuai dengan dokumen tersebut.

3.5.3.3. Diagram Proses 1 Level 1

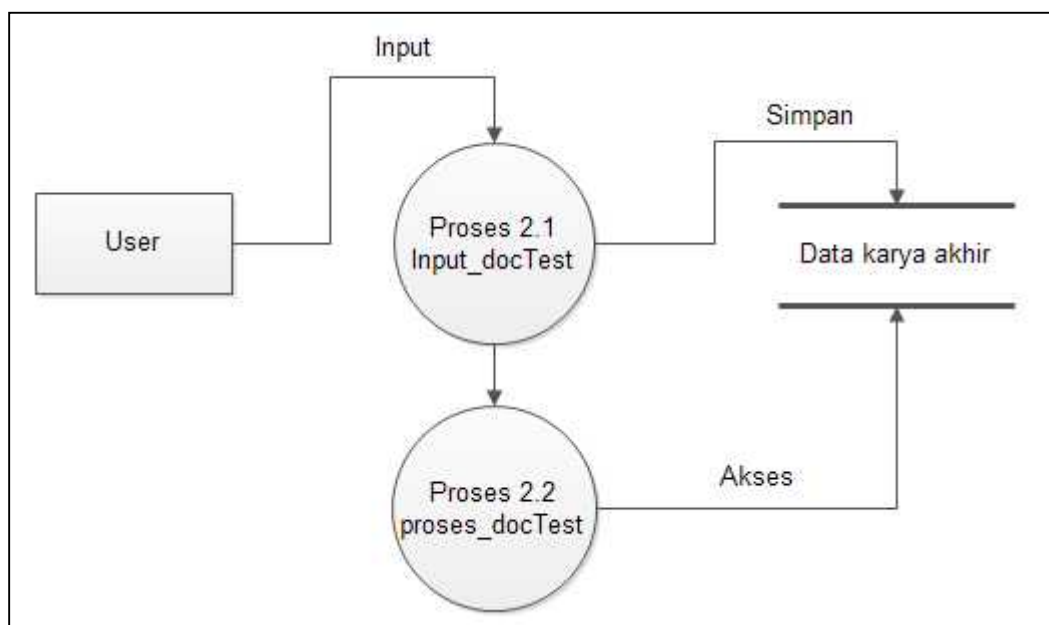
Pada diagram proses 1 level 1 menunjukkan penggambaran lengkap tentang proses entri data ke dalam sistem *database* yang meliputi id dokumen, nama penulis, judul karya akhir, hingga kategori dari dokumen tersebut. *Admin* dapat melakukan entri data mengenai dokumen abstrak yang bersangkutan, sedangkan *user* hanya dapat *input data test*, yang hasil dari *data test* tersebut bergantung pada *data train*. Diagram proses 1 level 1 digambarkan secara lengkap pada Gambar 3.4 di bawah ini.



Gambar 3.4. Diagram Proses 1 Level 1 Sistem Pengklasifikasi Dokumen menggunakan *Naïve Bayes Classifier*

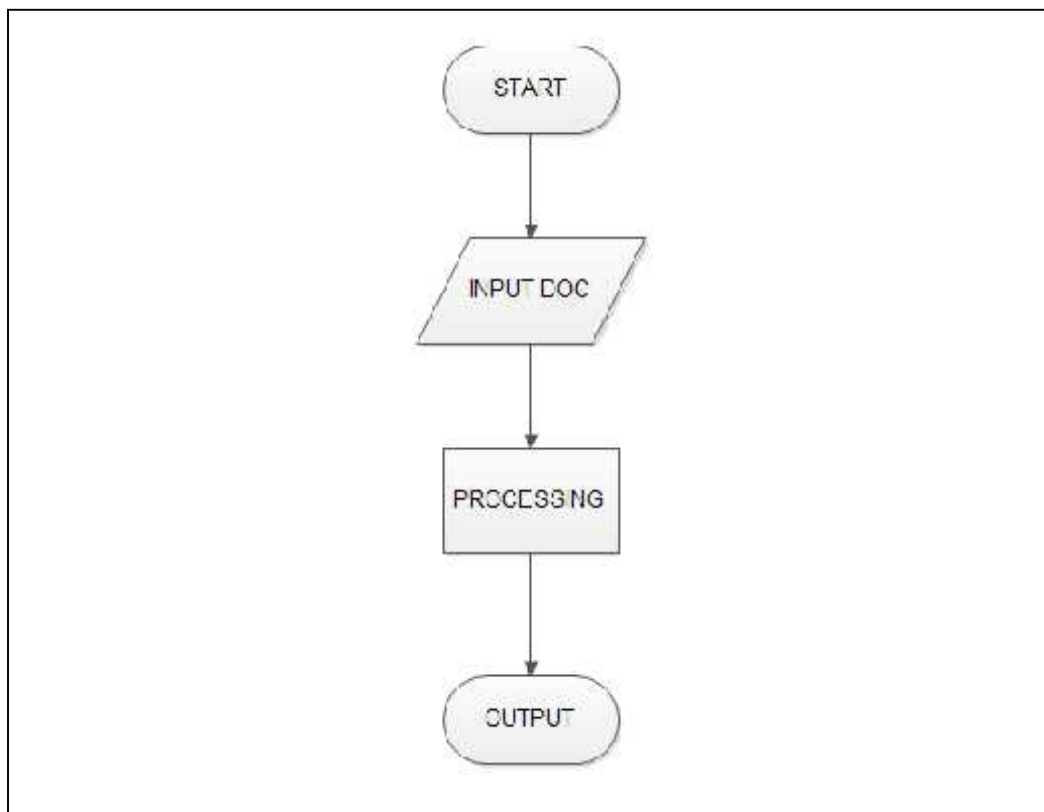
3.5.3.4. Diagram Proses 2 Level 1

Pada diagram proses 2 level 1 menunjukkan *User* dapat meng-*input* dokumen abstrak yang sesuai dengan ketentuan. Sehingga, *User* mendapatkan informasi berupa suatu kategori dari dokumen yang telah di *input* sebelumnya. Diagram proses 2 level 1 ditunjukkan pada Gambar 3.5 di bawah ini.



Gambar 3.5. Diagram Proses 2 Level 1 Sistem Pengklasifikasi Dokumen menggunakan *Naïve Bayes Classifier*

3.5.3.5. Flow Chart Diagram



Gambar 3.6. *Flow Chart Diagram*

3.5.3.6. Desain Struktur Database

Tabel berikut adalah desain *database* karya akhir yang akan dikembangkan dalam sistem klasifikasi dokumen.

Tabel 3.3. Tabel kata dasar pada *database* karya akhir

No.	Nama atribut	Tipe atribut	Keterangan
1.	id_katadasar	INT (10)	AUTO_INCREMENT
2.	katadasar	VARCHAR (20)	NOT NULL
3.	tipe_katadasar	VARCHAR (20)	NOT NULL

Tabel 3.4. Tabel *stoplist* pada *database* karya akhir

No.	Nama atribut	Tipe atribut	Keterangan
1.	id_stoplist	INT (10)	AUTO_INCREMENT
2.	stoplist	VARCHAR (50)	NOT NULL

Tabel 3.5. Tabel karya akhir pada *database* karya akhir

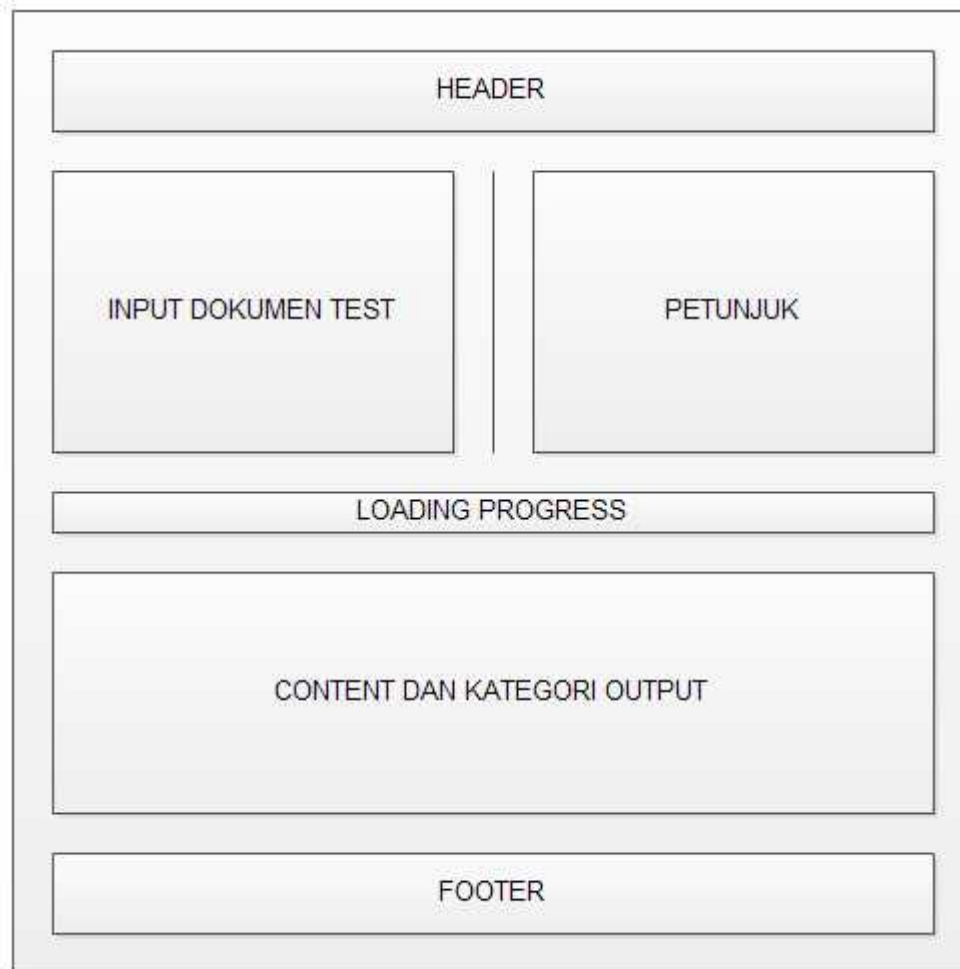
No.	Nama atribut	Tipe atribut	Keterangan
1.	id_karyaakhir	INT (15)	AUTO_INCREMENT
2.	nama_penulis	VARCHAR (100)	NOT NULL
3.	judul_karyaakhir	VARCHAR (500)	NOT NULL
4.	isi_abstrak	LONGTEXT	NOT NULL
5.	kategori	VARCHAR (30)	NOT NULL

Tabel 3.6. Tabel *upload* pada *database* karya akhir

No.	Nama atribut	Tipe atribut	Keterangan
1.	id_karyaakhir	INT(15)	AUTO_INCREMENT
2.	filename	LONGTEXT	NOT NULL

3.5.3.7. Pembuatan Desain Tampilan Halaman Sistem

Hal utama yang dibutuhkan dalam tampilan sebuah sistem (dalam hal ini berbasis web), adalah desain yang sederhana dan *user friendly* bagi para penggunanya. Sehingga, *user* dapat menggunakannya tanpa perlu kesulitan. Pada Gambar 3.7. merupakan desain sketsa tampilan halaman utama Sistem klasifikasi Dokumen.



Gambar 3.7. Desain tampilan halaman utama Sistem Pengklasifikasi Dokumen

Keterangan:

1. Header

Berisikan nama Sistem Kategori Dokumen.

2. Input Dokumen Test

Berisikan *field* untuk *browse* dokumen bertipe .doc atau .docx yang akan di-*input* oleh *user*.

3. Petunjuk

Berisikan langkah-langkah cara mengklasifikasikan dokumen abstrak karya akhir.

4. Content

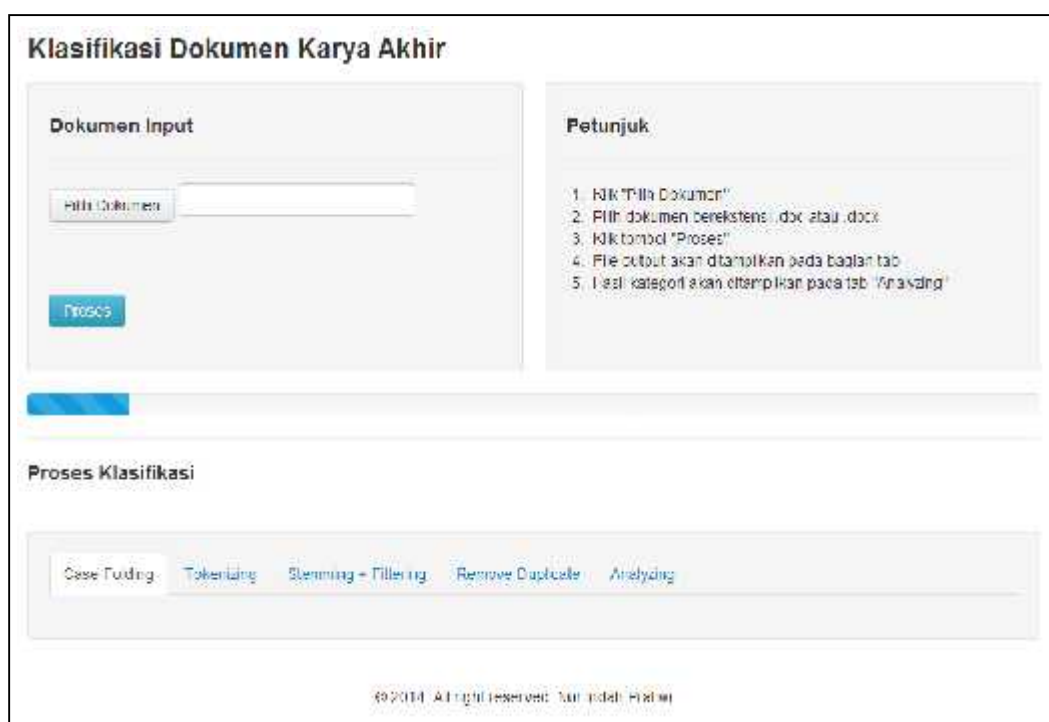
Berisikan proses pemecahan *string* yang terdapat dalam dokumen sehingga dapat dihitung menggunakan *Naïve Bayes Classifier*. Serta, berisikan keluaran hasil dari pemecahan string, perhitungan *Naïve Bayes Classifier* terhadap dokumen yang telah di-*input*, dan kategori.

5. Footer

Berisikan tahun pembuatan sistem dan nama pembuatnya.

3.5.3.8. Implementasi Tampilan Halaman Sistem

Pada tahap implementasi tampilan, akan memperlihatkan hasil halaman utama sistem yang telah dibuat sesuai dengan desain sketsa pada Gambar 3.7. Pada Gambar 3.8. dibawah ini, menunjukkan tampilan halaman utama Sistem Kategori Dokumen.



Gambar 3.8. Tampilan halaman utama Sistem Pengklasifikasi Dokumen

3.5.3.9. Langkah Kerja Sistem

Setelah melalui tahap pembuatan desain alur data, desain *database* dan desain tampilan sistem, maka tahap selanjutnya adalah proses *coding*. Berikut adalah langkah kerja dalam pengembangan sistem untuk mengklasifikasi dokumen:

1. Mengumpulkan 100 dokumen berbentuk abstrak skripsi
2. Abstrak skripsi tersebut di *filter* sehingga isi abstrak saja yang diambil
3. Pembuatan desain alur data
4. Pembuatan desain *database*
5. Pembuatan desain tampilan sistem
6. Proses *coding* sistem, membuat fungsi *case folding*, *tokenizing*, *filtering*, *stemming* dan *remove duplicate*
7. Input dokumen isi abstrak skripsi ke dalam sistem
8. Menghitung jumlah kata yang muncul
9. Menentukan kata yang akan menjadi acuan
10. Proses perhitungan manual kata tersebut untuk menentukan peluang dalam 10 kali uji
11. Nilai peluang yang dihasilkan akan menjadi variabel pada sistem klasifikasi untuk menghitung peluang pada masing-masing kategori
12. Pembuatan fungsi *analyzing* dengan menggunakan algoritma NBC
13. Nilai peluang yang sudah didapatkan akan dimasukkan ke dalam sistem untuk digabungkan dengan fungsi *analyzing* yang telah dibuat
14. Nilai peluang tertinggi akan di tampilkan oleh sistem

3.5.4. Implementasi dan *Testing Unit*

Pada tahap implementasi dan *testing unit* kode-kode pemrograman dilakukan. Sistem klasifikasi akan dipecah menjadi beberapa modul kecil untuk mendapatkan kata-kata yang kemudian akan digabungkan dan dihitung dalam *Naïve Bayes Classifier*.

Bahasa Pemrograman menggunakan PHP versi 5.4.22, serta menggunakan jQuery versi 1.8, MySQL versi 5.5.34 yang digunakan sebagai *database* untuk menyimpan data dengan menggunakan *interface* phpmyadmin versi 4.0.9.

3.5.5. Integrasi dan *Testing System*

Pada tahap ini dilakukan penggabungan modul yang dipecah menjadi lebih kecil pada tahap sebelumnya. Model tahap integrasi dan pengujian sistem menggunakan model *black box*, yaitu pengujian yang digunakan untuk menguji fungsi-fungsi khusus dari perangkat lunak yang dirancang. Kebenaran perangkat lunak yang diuji hanya dilihat berdasarkan keluaran yang dihasilkan dari data atau kondisi masukan yang diberikan untuk fungsi yang ada, tanpa melihat bagaimana proses untuk mendapatkan keluaran tersebut. Dari keluaran yang dihasilkan, kemampuan program dalam memenuhi kebutuhan pemakai dapat diukur sekaligus dapat diketahui kesalahannya.

Pengujian kebutuhan fungsional dilakukan dengan menggunakan scenario proses yang telah ditentukan pada tabel 3.7.

Tabel 3.7. Kriteria pengujian kebutuhan fungsional.

No.	Skenario proses	Tampilan Pada Sistem
1.	<i>User</i> tidak memilih file. Mengklik tombol “Proses”. Field kosong.	
2.	<i>User</i> tidak memilih file. Mengklik tab “Tokenizing”. Field kosong.	
3.	<i>User</i> tidak memilih file. Mengklik tab “Stemming+Filtering”. Field kosong.	
4.	<i>User</i> tidak memilih file. Mengklik tab “Remove Duplicate”. Field kosong.	
5.	<i>User</i> tidak memilih file. Mengklik tab “Analyzing”. Field kosong.	
6.	<i>User</i> memilih file bertipe image. Mengklik tombol “Proses”.	
7.	<i>User</i> memilih file bertipe video. Mengklik tombol “Proses”.	
8.	<i>User</i> memilih file dokumen berekstensi xlsx. Mengklik tombol “Proses”.	
9.	<i>User</i> memilih file dokumen berekstensi ppt. Mengklik tombol “Proses”.	
10.	<i>User</i> memilih file dokumen berekstensi pdf. Mengklik tombol “Proses”.	
11.	<i>User</i> memilih file dokumen berekstensi txt. Mengklik tombol “Proses”.	
12.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”.	
13.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”.	
14.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Tokenizing”.	
15.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Stemming + Filtering”.	
16.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Remove Duplicate”.	

No.	Skenario proses	Tampilan Pada Sistem
17.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Analyze”.	
18.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Tokenizing” saat masih loading.	
19.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Stemming+Filtering” saat masih loading.	
20.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Remove Duplicate” saat masih loading.	
21.	<i>User</i> memilih file. Mengklik tombol “Proses”. Mengklik tab “Analyzing” saat masih loading.	
22.	<i>User</i> memilih file dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tombol “Ulangi Proses”.	

3.5.6. Operasi dan *Maintenance*

Pemeliharaan dilakukan untuk memperbaiki jika sistem tidak berjalan sesuai dengan fungsinya yang kemungkinan disebabkan karena perubahan atau memperbaharui kode pemrograman, meskipun sudah dilakukan *testing* sebelumnya.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Hasil Penelitian

Setelah melalui beberapa tahap pengembangan *software*, mulai dari pengumpulan data abstrak skripsi, proses *filtering* menjadi isi abstrak, hingga proses *preprocessing* serta proses *coding* perangkat lunak, maka dihasilkan sebuah sistem klasifikasi yang bernama Sistem Klasifikasi Dokumen Skripsi.

4.1.1. Hasil Pengujian Perangkat Lunak

Setelah semua proses pengembangan *software* selesai dilakukan, maka perlu dilakukan pengujian. Pengujian dilakukan untuk mengetahui apakah *software* yang telah dikembangkan telah layak digunakan oleh *User* atau tidak. Apabila *software* belum layak digunakan oleh *User*, maka *software* tersebut perlu dilakukan perbaikan agar sesuai dengan tujuan pembuatan. Setelah semua proses perbaikan telah selesai dilakukan dan *software* tersebut telah dinyatakan layak, maka *software* tersebut telah selesai dikembangkan dan *software* tersebut siap digunakan.

Berdasarkan penjelasan sebelumnya, maka perlu dijelaskan bagaimana melakukan pengujian *software* tersebut. Pada tabel 4.1. merupakan hasil pengujian *software* yang telah dilakukan berdasarkan pada uraian yang ada pada Bab III.

Tabel. 4.1. Hasil pengujian kebutuhan fungsional

No.	Skenario proses	Tampilan Pada Sistem
1.	<i>User</i> tidak memilih file. Mengklik tombol “Proses”. Field kosong.	Error: undefined extension
2.	<i>User</i> tidak memilih file. Mengklik tab “Tokenizing”. Field kosong.	Error: Invalid file type
3.	<i>User</i> tidak memilih file. Mengklik tab “Stemming+Filtering”. Field kosong.	Error: Invalid file type
4.	<i>User</i> tidak memilih file. Mengklik tab “Remove Duplicate”. Field kosong.	Error: Invalid file type
5.	<i>User</i> tidak memilih file. Mengklik tab “Analyzing”. Field kosong.	Error: Invalid file type
6.	<i>User</i> memilih file bertipe image. Mengklik tombol “Proses”.	Error: Invalid file type
7.	<i>User</i> memilih file bertipe video. Mengklik tombol “Proses”.	Error: Invalid file type
8.	<i>User</i> memilih file dokumen berekstensi xlsx. Mengklik tombol “Proses”.	Error: Invalid file type
9.	<i>User</i> memilih file dokumen berekstensi ppt. Mengklik tombol “Proses”.	Error: Invalid file type
10.	<i>User</i> memilih file dokumen berekstensi pdf. Mengklik tombol “Proses”.	Error: File not exist
11.	<i>User</i> memilih file dokumen berekstensi txt. Mengklik tombol “Proses”.	Error: Invalid file type
12.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”.	Case folding berisi teks dari file yang pilih
13.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”.	Case folding berisi teks dari file yang pilih
14.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Tokenizing”.	Menampilkan kata dengan urutan ke bawah (Satu baris satu kata)
15.	<i>User</i> memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Stemming + Filtering”.	Menampilkan kata dasar dari kata yang ditampilkan pada tab tokenizing

No.	Skenario proses	Tampilan Pada Sistem
16.	User memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Remove Duplicate”.	Menghilangkan kata yang duplikat serta menghitung jumlah kata
17.	User memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Analyze”.	Menampilkan jumlah kata term pilihan serta menampilkan perhitungan, serta menampilkan kategori
18.	User memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Tokenizing” saat masih loading.	Buffer. Tab dapat terbuka setelah proses selesai
19.	User memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Stemming+Filtering” saat masih loading.	Buffer. Tab dapat terbuka setelah proses selesai
20.	User memilih dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tab “Remove Duplicate” saat masih loading.	Buffer. Tab dapat terbuka setelah proses selesai
21.	User memilih file. Mengklik tombol “Proses”. Mengklik tab “Analyzing” saat masih loading.	Buffer. Tab dapat terbuka setelah proses selesai
22.	User memilih file dokumen berekstensi doc atau docx. Mengklik tombol “Proses”. Mengklik tombol “Ulangi Proses”.	Sistem menuju index.php dan semua field menjadi kosong kembali

4.1.2. Hasil Pengujian Algoritma *Naïve Bayes Classifier*

Untuk dapat mengetahui tingkat akurasi hasil prediksi terhadap aplikasi *data mining* yang telah dikembangkan, maka perlu diadakan *testing*. Tabel di bawah ini merupakan tabel data *testing* pengujian sistem berdasarkan data yang di ambil pada dokumen abstrak di Jurusan Teknik Elektro Universitas Negeri

Jakarta. Pengujian sistem menggunakan metode *K-Fold Cross Validation* dengan membagi data menjadi 10 bagian (10 Fold). Perbandingan *data train* dan *data tes* sebesar 9:1.

Tabel 4.2. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 1 (Fold 1)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	42	0	2	0	0	3	0	2	Analisa
2	43	0	2	0	1	0	0	1	Analisa
3	49	0	7	2	0	3	0	0	Analisa
4	55	0	0	0	0	3	0	1	Analisa
5	60	0	0	1	0	1	0	13	Analisa
6	61	0	3	5	0	0	0	2	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	5	14	0	0	6	1	10	2	Media Pembelajaran
28	13	9	0	1	6	0	11	0	Media Pembelajaran
29	23	8	0	0	1	3	4	2	Media Pembelajaran
30	29	7	0	0	7	0	7	0	Media Pembelajaran
31	38	4	0	2	1	0	2	0	Media Pembelajaran
32	56	3	0	0	3	0	0	1	Media Pembelajaran
33	84	5	0	0	4	0	0	2	Media Pembelajaran
34	90	1	0	0	8	0	0	1	Media Pembelajaran

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
35	96	6	0	0	0	1	4	2	Media Pembelajaran
36	100	5	0	0	4	1	1	0	Media Pembelajaran
37	8	0	0	0	1	0	4	0	Penelitian Pendidikan
38	9	0	0	0	7	0	7	1	Penelitian Pendidikan
39	11	0	2	0	5	0	6	2	Penelitian Pendidikan
40	12	0	0	0	0	0	9	3	Penelitian Pendidikan
41	16	0	0	0	8	1	12	4	Penelitian Pendidikan
42	17	0	1	0	2	0	9	0	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	14	0	6	0	0	0	0	0	Produk Rekayasa
62	15	0	3	5	0	0	0	0	Produk Rekayasa
63	19	0	0	3	0	0	0	1	Produk Rekayasa
64	20	3	3	0	0	0	0	2	Produk Rekayasa
65	24	0	7	0	0	0	0	2	Produk Rekayasa
66	26	0	0	0	0	2	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.2:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.3. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 1 (Fold 1)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
25	0	1	0	0	0	0	0	Analisa	Produk Rekayasa	3,1731 sec
39	0	0	0	0	0	0	1	Analisa	Analisa	4,6442 sec
40	0	1	0	0	2	0	2	Analisa	Analisa	4,7752 sec
1	7	3	1	5	0	0	0	Media Pemb.	Media Pemb.	3,5672 sec
4	0	0	2	4	0	9	3	Media Pemb.	Penelitian Pend.	5,3283 sec
6	0	3	0	11	0	16	1	Penelitian Pend.	Penelitian Pend.	4,9792 sec
7	0	0	0	3	1	10	1	Penelitian Pend.	Penelitian Pend.	4,6722 sec
2	0	5	0	0	0	0	1	Produk Rekayasa	Produk Rekayasa	4,4422 sec
3	0	10	0	0	1	0	0	Produk Rekayasa	Produk Rekayasa	4,1112 sec
10	0	0	0	0	0	0	5	Produk Rekayasa	Analisa	3,2571 sec

Keterangan Kolom di Tabel 4.3:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.3.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{7}{10} \times 100\% = 70\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{3}{10} \times 100\% = 30\%$$

Dari Tabel 4.3 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 1 dapat menghasilkan nilai akurasi sebesar 70% untuk 30 *Data Test* Uji ke – 1 serta nilai *error* sebesar 30%. dengan rata-rata waktu eksekusi 4,2949 Sekon.

Tabel 4.4. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 2 (Fold 2)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	55	0	0	0	0	3	0	1	Analisa
5	60	0	0	1	0	1	0	13	Analisa
6	61	0	3	5	0	0	0	2	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	13	9	0	1	6	0	11	0	Media Pembelajaran
30	23	8	0	0	1	3	4	2	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	12	0	0	0	0	0	9	3	Penelitian Pendidikan
41	16	0	0	0	8	1	12	4	Penelitian Pendidikan
42	17	0	1	0	2	0	9	0	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	20	3	3	0	0	0	0	2	Produk Rekayasa
65	24	0	7	0	0	0	0	2	Produk Rekayasa
66	26	0	0	0	0	2	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.4:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Tabel 4.5. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 2 (Fold 2)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
42	0	2	0	0	3	0	2	Analisa	Analisa	6,3053 sec
43	0	2	0	1	0	0	1	Analisa	Analisa	5,6473 sec
49	0	7	2	0	3	0	0	Analisa	Produk Rekayasa	4,0902 sec
5	14	0	0	6	1	10	2	Media Pemb.	Media Pemb.	4,9412 sec
8	0	0	0	1	0	4	0	Penelitian Pend.	Penelitian Pend.	3,9672 sec
9	0	0	0	7	0	7	1	Penelitian Pend.	Penelitian Pend.	4,6752 sec
11	0	2	0	5	0	6	2	Penelitian Pend.	Penelitian Pend.	3,3301 sec
14	0	6	0	0	0	0	0	Produk Rekayasa	Produk Rekayasa	3,7572 sec
15	0	3	5	0	0	0	0	Produk Rekayasa	Produk Rekayasa	4,2142 sec
19	0	0	3	0	0	0	1	Produk Rekayasa	Produk Rekayasa	3,9292 sec

Keterangan Kolom di Tabel 4.5:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.5.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{9}{10} \times 100\% = 90\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{1}{10} \times 100\% = 10\%$$

Dari Tabel 4.5 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 2 dapat menghasilkan nilai akurasi sebesar 90% untuk 30 *Data Test* Uji ke – 2 serta nilai *error* sebesar 10%. dengan rata-rata waktu eksekusi 4,4857 Sekon.

Tabel 4.6. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 3 (Fold 3)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	23	8	0	0	1	3	4	2	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.6:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.7. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 3 (Fold 3)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
55	0	0	0	0	3	0	1	Analisa	Analisa	4,2352 sec
60	0	0	1	0	1	0	13	Analisa	Analisa	5,2423 sec
61	0	3	5	0	0	0	2	Analisa	Produk Rekayasa	4,0932 sec
13	9	0	1	6	0	11	0	Media Pemb.	Media Pemb.	4,3722 sec
12	0	0	0	0	0	9	3	Penelitian Pend.	Penelitian Pend.	4,6532 sec
16	0	0	0	8	1	12	4	Penelitian Pend.	Penelitian Pend.	5,3893 sec
17	0	1	0	2	0	9	0	Penelitian Pend.	Penelitian Pend.	4,3502 sec
20	3	3	0	0	0	0	2	Produk Rekayasa	Analisa	5,1392 sec
24	0	7	0	0	0	0	2	Produk Rekayasa	Produk Rekayasa	3,5162 sec
26	0	0	0	0	2	0	1	Produk Rekayasa	Analisa	3,7252 sec

Keterangan Kolom di Tabel 4.7:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.7.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{7}{10} \times 100\% = 70\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{3}{10} \times 100\% = 30\%$$

Dari Tabel 4.7 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 3 dapat menghasilkan nilai akurasi sebesar 70% untuk 30 *Data Test* Uji ke – 3 serta nilai *error* sebesar 30%. dengan rata-rata waktu eksekusi 3,9557 Sekon.

Tabel 4.8. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 4 (Fold 4)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.8:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.9. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 4 (Fold 4)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
62	0	1	0	0	0	0	0	Analisa	Produk Rekayasa	3,8452 sec
63	0	0	0	0	2	0	2	Analisa	Analisa	4,3412 sec
64	0	0	0	0	1	0	2	Analisa	Analisa	4,4182 sec
23	8	0	0	1	3	4	2	Media Pemb.	Media Pemb.	4,8182 sec
18	0	0	0	4	0	12	1	Penelitian Pend.	Penelitian Pend.	4,6622 sec
21	0	0	0	2	1	7	0	Penelitian Pend.	Penelitian Pend.	4,4182 sec
22	0	3	1	8	1	15	4	Penelitian Pend.	Penelitian Pend.	5,5393 sec
27	0	1	0	0	0	0	0	Produk Rekayasa	Produk Rekayasa	2,8931 sec
28	0	0	0	0	0	2	1	Produk Rekayasa	Penelitian Pend.	4,2802 sec
30	0	2	0	0	0	0	0	Produk Rekayasa	Produk Rekayasa	3,5282 sec

Keterangan Kolom di Tabel 4.9:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.9.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{9}{10} \times 100\% = 90\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{1}{10} \times 100\% = 10\%$$

Dari Tabel 4.9 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 4 dapat menghasilkan nilai akurasi sebesar 90% untuk 30 *Data Test* Uji ke – 4 serta nilai *error* sebesar 10%. dengan rata-rata waktu eksekusi 4,2744 Sekon.

Tabel 4.10. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 5 (Fold 5)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	36	0	5	0	0	0	10	2	Penelitian Pendidikan
50	44	0	0	0	0	2	14	1	Penelitian Pendidikan
51	46	0	0	0	0	1	2	1	Penelitian Pendidikan
52	50	0	0	0	3	1	7	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.10:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.11. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 5 (Fold 5)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
66	0	7	0	0	0	0	1	Analisa	Produk Rekayasa	3,6382 sec
67	0	4	0	0	3	0	2	Analisa	Analisa	5,5113 sec
68	0	0	2	0	1	0	1	Analisa	Analisa	4,5742 sec
29	7	0	0	7	0	7	0	Media Pemb.	Media Pemb.	4,4762 sec
32	0	0	0	11	1	9	0	Penelitian Pend.	Penelitian Pend.	5,5623 sec
33	0	0	0	5	0	6	2	Penelitian Pend.	Penelitian Pend.	5,1892 sec
31	0	1	2	0	0	0	2	Produk Rekayasa	Produk Rekayasa	3,7752 sec
34	0	1	0	0	2	0	1	Produk Rekayasa	Analisa	3,8312 sec
35	0	4	4	0	2	0	2	Produk Rekayasa	Produk Rekayasa	3,9952 sec
37	0	2	1	0	0	0	2	Produk Rekayasa	Produk Rekayasa	3,7892 sec

Keterangan Kolom di Tabel 4.11:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.11.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{8}{10} \times 100\% = 80\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{2}{10} \times 100\% = 20\%$$

Dari Tabel 4.11 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 5 dapat menghasilkan nilai akurasi sebesar 80% untuk 30 *Data Test* Uji ke – 5 serta nilai *error* sebesar 20%. dengan rata-rata waktu eksekusi 4,0553 Sekon.

Tabel 4.12. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 6 (Fold 6)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	46	0	0	0	0	1	2	1	Penelitian Pendidikan
52	50	0	0	0	3	1	7	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	31	0	1	2	0	0	0	2	Produk Rekayasa
75	34	0	1	0	0	2	0	1	Produk Rekayasa
76	35	0	4	4	0	2	0	2	Produk Rekayasa
77	37	0	2	1	0	0	0	2	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.12:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.13. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 6 (Fold 6)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
70	0	0	0	0	2	0	2	Analisa	Analisa	41672
71	0	0	0	0	0	0	1	Analisa	Analisa	39972
72	0	1	4	0	0	0	0	Analisa	Produk Rekayasa	42232
38	4	0	2	1	0	2	0	Media Pemb.	Media Pemb.	46692
36	0	5	0	0	0	10	2	Penelitian Pend.	Penelitian Pend.	48962
44	0	0	0	0	2	14	1	Penelitian Pend.	Penelitian Pend.	43762
41	0	0	0	0	0	0	1	Produk Rekayasa	Analisa	33571
45	0	4	4	0	0	0	2	Produk Rekayasa	Produk Rekayasa	40912
47	0	0	0	0	1	0	2	Produk Rekayasa	Analisa	41282
48	0	4	0	0	0	1	1	Produk Rekayasa	Produk Rekayasa	47322

Keterangan Kolom di Tabel 4.13:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.13.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{7}{10} \times 100\% = 70\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{3}{10} \times 100\% = 30\%$$

Dari Tabel 4.13 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 6 dapat menghasilkan nilai akurasi sebesar 70% untuk 30 *Data Test* Uji ke – 6 serta nilai *error* sebesar 30%. dengan rata-rata waktu eksekusi 4,2637 Sekon.

Tabel 4.14. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 7 (Fold 7)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	36	0	5	0	0	0	10	2	Penelitian Pendidikan
52	44	0	0	0	0	2	14	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	31	0	1	2	0	0	0	2	Produk Rekayasa
75	34	0	1	0	0	2	0	1	Produk Rekayasa
76	35	0	4	4	0	2	0	2	Produk Rekayasa
77	37	0	2	1	0	0	0	2	Produk Rekayasa
78	41	0	0	0	0	0	0	1	Produk Rekayasa
79	45	0	4	4	0	0	0	2	Produk Rekayasa
80	47	0	0	0	0	1	0	2	Produk Rekayasa
81	48	0	4	0	0	0	1	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.14:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.15. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 7 (Fold 7)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
73	0	0	0	0	1	0	2	Analisa	Analisa	4,5362 sec
78	0	0	1	0	1	0	2	Analisa	Analisa	4,3102 sec
79	0	1	0	0	1	0	2	Analisa	Analisa	4,2162 sec
56	3	0	0	3	0	0	1	Media Pemb.	Media Pemb.	4,1272 sec
46	0	0	0	0	1	2	1	Penelitian Pend.	Penelitian Pend.	3,7032 sec
50	0	0	0	3	1	7	1	Penelitian Pend.	Penelitian Pend.	4,3202 sec
51	0	3	1	0	0	0	1	Produk Rekayasa	Produk Rekayasa	5,3083 sec
52	0	0	4	0	0	0	0	Produk Rekayasa	Produk Rekayasa	3,7032 sec
57	0	0	1	0	0	0	1	Produk Rekayasa	Produk Rekayasa	4,6232 sec
58	0	0	7	0	0	0	1	Produk Rekayasa	Produk Rekayasa	3,9962 sec

Keterangan Kolom di Tabel 4.15:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.15.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{10}{10} \times 100\% = 100\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{0}{10} \times 100\% = 0\%$$

Dari Tabel 4.15 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 7 dapat menghasilkan nilai akurasi sebesar 100% untuk 30 *Data Test* Uji ke – 7 serta nilai *error* sebesar 0%. dengan rata-rata waktu eksekusi 4,2844 Sekon.

Tabel 4.16. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 8 (Fold 8)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	56	3	0	0	3	0	0	1	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	36	0	5	0	0	0	10	2	Penelitian Pendidikan
52	44	0	0	0	0	2	14	1	Penelitian Pendidikan
53	46	0	0	0	0	1	2	1	Penelitian Pendidikan
54	50	0	0	0	3	1	7	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	52	0	0	4	0	0	0	0	Produk Rekayasa
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.16:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Tabel 4.17. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 8 (Fold 8)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
80	0	0	0	0	0	0	1	Analisa	Analisa	5,0392 sec
81	0	0	0	0	1	0	1	Analisa	Analisa	3,9962 sec
83	0	7	1	0	0	0	0	Analisa	Produk Rekayasa	4,1272 sec
84	5	0	0	4	0	0	2	Media Pemb.	Media Pemb.	3,9632 sec
59	0	5	0	0	0	0	0	Produk Rekayasa	Produk Rekayasa	4,4152 sec
74	0	3	0	0	0	0	1	Produk Rekayasa	Produk Rekayasa	3,3251 sec
77	0	0	2	0	0	0	0	Produk Rekayasa	Produk Rekayasa	5,3083 sec
53	1	0	0	0	1	3	1	Penelitian Pend.	Media Pemb.	3,5492 sec
54	0	0	0	7	0	0	1	Penelitian Pend.	Penelitian Pend.	4,9732 sec
65	0	0	0	1	1	0	1	Penelitian Pend.	Analisa	3,9312 sec

Keterangan Kolom di Tabel 4.17:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.17.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{8}{10} \times 100\% = 80\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{2}{10} \times 100\% = 20\%$$

Dari Tabel 4.17 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 8 dapat menghasilkan nilai akurasi sebesar 80% untuk 30 *Data Test* Uji ke – 8 serta nilai *error* sebesar 20%. dengan rata-rata waktu eksekusi 4,2628 Sekon.

Tabel 4.18. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 9 (Fold 9)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	80	0	0	0	0	0	0	1	Analisa
23	81	0	0	0	0	1	0	1	Analisa
24	83	0	7	1	0	0	0	0	Analisa
25	91	0	0	0	0	0	0	1	Analisa
26	93	0	1	0	0	2	0	0	Analisa
27	94	0	1	0	0	2	0	1	Analisa
28	1	7	3	1	5	0	0	0	Media Pembelajaran
29	4	0	0	2	4	0	9	3	Media Pembelajaran
30	5	14	0	0	6	1	10	2	Media Pembelajaran
31	13	9	0	1	6	0	11	0	Media Pembelajaran
32	23	8	0	0	1	3	4	2	Media Pembelajaran
33	29	7	0	0	7	0	7	0	Media Pembelajaran
34	38	4	0	2	1	0	2	0	Media Pembelajaran
35	56	3	0	0	3	0	0	1	Media Pembelajaran
36	84	5	0	0	4	0	0	2	Media Pembelajaran
37	96	6	0	0	0	1	4	2	Media Pembelajaran
38	100	5	0	0	4	1	1	0	Media Pembelajaran
39	6	0	3	0	11	0	16	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	7	0	0	0	3	1	10	1	Penelitian Pendidikan
41	8	0	0	0	1	0	4	0	Penelitian Pendidikan
42	9	0	0	0	7	0	7	1	Penelitian Pendidikan
43	11	0	2	0	5	0	6	2	Penelitian Pendidikan
44	12	0	0	0	0	0	9	3	Penelitian Pendidikan
45	16	0	0	0	8	1	12	4	Penelitian Pendidikan
46	17	0	1	0	2	0	9	0	Penelitian Pendidikan
47	18	0	0	0	4	0	12	1	Penelitian Pendidikan
48	21	0	0	0	2	1	7	0	Penelitian Pendidikan
49	22	0	3	1	8	1	15	4	Penelitian Pendidikan
50	32	0	0	0	11	1	9	0	Penelitian Pendidikan
51	33	0	0	0	5	0	6	2	Penelitian Pendidikan
52	36	0	5	0	0	0	10	2	Penelitian Pendidikan
53	44	0	0	0	0	2	14	1	Penelitian Pendidikan
54	46	0	0	0	0	1	2	1	Penelitian Pendidikan
55	50	0	0	0	3	1	7	1	Penelitian Pendidikan
56	53	1	0	0	0	1	3	1	Penelitian Pendidikan
57	54	0	0	0	7	0	0	1	Penelitian Pendidikan
58	65	0	0	0	1	1	0	1	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	52	0	0	4	0	0	0	0	Produk Rekayasa
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	59	0	5	0	0	0	0	0	Produk Rekayasa
86	74	0	3	0	0	0	0	1	Produk Rekayasa
87	77	0	0	2	0	0	0	0	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Keterangan Kolom di Tabel 4.18:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Tabel 4.19. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 9 (Fold 9)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
85	0	0	0	0	2	0	1	Analisa	Analisa	3,9632 sec
88	0	0	0	6	0	0	1	Analisa	Penelitian Pend.	4,3202 sec
90	1	0	0	8	0	0	1	Media Pemb.	Media Pemb.	3,9622 sec
69	0	0	0	0	0	2	0	Penelitian Pend.	Penelitian Pend.	5,3083 sec
75	0	0	0	5	2	5	1	Penelitian Pend.	Penelitian Pend.	4,3212 sec
76	0	2	0	11	2	9	1	Penelitian Pend.	Penelitian Pend.	3,5132 sec
89	0	0	0	0	4	5	2	Penelitian Pend.	Penelitian Pend.	5,4173 sec
82	0	1	4	0	0	0	0	Produk Rekayasa	Produk Rekayasa	3,7472 sec
86	0	2	4	0	0	0	0	Produk Rekayasa	Produk Rekayasa	4,4442 sec
87	0	2	0	0	0	0	2	Produk Rekayasa	Analisa	5,1742 sec

Keterangan Kolom di Tabel 4.19:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.19.

$$\text{Accuracy} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{8}{10} \times 100\% = 80\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictons}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{2}{10} \times 100\% = 20\%$$

Dari Tabel 4.19 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 9 dapat menghasilkan nilai akurasi sebesar 80% untuk 30 *Data Test* Uji ke – 9 serta nilai *error* sebesar 20%. dengan rata-rata waktu eksekusi 4,4171 Sekon.

Tabel 4.20. Pencarian Probabilitas dengan Menggunakan 90 Data Train dengan Data Tes Uji ke – 10 (Fold 10)

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	80	0	0	0	0	0	0	1	Analisa
23	81	0	0	0	0	1	0	1	Analisa
24	83	0	7	1	0	0	0	0	Analisa
25	85	0	0	0	0	2	0	1	Analisa
26	88	0	0	0	6	0	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	56	3	0	0	3	0	0	1	Media Pembelajaran
35	84	5	0	0	4	0	0	2	Media Pembelajaran
36	90	1	0	0	8	0	0	1	Media Pembelajaran
37	6	0	3	0	11	0	16	1	Penelitian Pendidikan
38	7	0	0	0	3	1	10	1	Penelitian Pendidikan
39	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
40	9	0	0	0	7	0	7	1	Penelitian Pendidikan
41	11	0	2	0	5	0	6	2	Penelitian Pendidikan
42	12	0	0	0	0	0	9	3	Penelitian Pendidikan
43	16	0	0	0	8	1	12	4	Penelitian Pendidikan
44	17	0	1	0	2	0	9	0	Penelitian Pendidikan
45	18	0	0	0	4	0	12	1	Penelitian Pendidikan
46	21	0	0	0	2	1	7	0	Penelitian Pendidikan
47	22	0	3	1	8	1	15	4	Penelitian Pendidikan
48	32	0	0	0	11	1	9	0	Penelitian Pendidikan
49	33	0	0	0	5	0	6	2	Penelitian Pendidikan
50	36	0	5	0	0	0	10	2	Penelitian Pendidikan
51	44	0	0	0	0	2	14	1	Penelitian Pendidikan
52	46	0	0	0	0	1	2	1	Penelitian Pendidikan
53	50	0	0	0	3	1	7	1	Penelitian Pendidikan
54	53	1	0	0	0	1	3	1	Penelitian Pendidikan
55	54	0	0	0	7	0	0	1	Penelitian Pendidikan
56	65	0	0	0	1	1	0	1	Penelitian Pendidikan
57	69	0	0	0	0	0	2	0	Penelitian Pendidikan
58	75	0	0	0	5	2	5	1	Penelitian Pendidikan
59	76	0	2	0	11	2	9	1	Penelitian Pendidikan
60	89	0	0	0	0	4	5	2	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
82	52	0	0	4	0	0	0	0	Produk Rekayasa
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	59	0	5	0	0	0	0	0	Produk Rekayasa
86	74	0	3	0	0	0	0	1	Produk Rekayasa
87	77	0	0	2	0	0	0	0	Produk Rekayasa
88	82	0	1	4	0	0	0	0	Produk Rekayasa
89	86	0	2	4	0	0	0	0	Produk Rekayasa
90	87	0	2	0	0	0	0	2	Produk Rekayasa

Keterangan Kolom di Tabel 4.20:

#1 : Atribut "Media"

#5 : Atribut "Analisis"

#2 : Atribut "Sistem"

#6 : Atribut "Ajar"

#3 : Atribut "Alat"

#7 : Atribut "Metode"

#4 : Atribut "Pembelajaran"

Tabel 4.21. Perbandingan Class Prediksi dan Class Sistem dengan Menggunakan 10 Data Tes Uji ke – 10 (Fold 10)

id	#1	#2	#3	#4	#5	#6	#7	Class Prediksi	Class Sistem	Waktu
91	0	0	0	0	0	0	1	Analisa	Analisa	3,8882 sec
93	0	1	0	0	2	0	0	Analisa	Analisa	3,7472 sec
94	0	1	0	0	2	0	1	Analisa	Analisa	4,4442 sec
96	6	0	0	0	1	4	2	Media Pemb.	Media Pemb.	4,0202 sec
100	5	0	0	4	1	1	0	Media Pemb.	Media Pemb.	5,4283 sec
92	0	1	0	8	0	7	2	Penelitian Pend.	Penelitian Pend.	3,6632 sec
97	0	0	0	4	1	7	1	Penelitian Pend.	Penelitian Pend.	3,7832 sec
95	0	0	8	0	0	0	2	Produk Rekayasa	Produk Rekayasa	2,9211 sec
98	0	6	3	0	0	0	1	Produk Rekayasa	Produk Rekayasa	4,8052 sec
99	0	2	4	0	0	0	0	Produk Rekayasa	Produk Rekayasa	4,8832 sec

Keterangan Kolom di Tabel 4.21:

#1	: Atribut "Media"	#5	: Atribut "Analisis"
#2	: Atribut "Sistem"	#6	: Atribut "Ajar"
#3	: Atribut "Alat"	#7	: Atribut "Metode"
#4	: Atribut "Pembelajaran"		

Perhitungan Akurasi Sistem pada Tabel 4.21.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Akurasi} = \frac{10}{10} \times 100\% = 100\%$$

$$\text{Error rate} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

$$\text{Nilai Error} = \frac{0}{10} \times 100\% = 00\%$$

Dari Tabel 4.21 didapatkan hasil bahwa dari 90 *Data Train* Uji ke – 10 dapat menghasilkan nilai akurasi sebesar 100% untuk 30 *Data Test* Uji ke – 10 serta nilai *error* sebesar 0%. dengan rata-rata waktu eksekusi 3,7800 Sekon.

4.2. Pembahasan

4.2.1. Pembahasan Hasil Pengujian Perangkat Lunak

Berdasarkan Tabel 4.1 dapat diketahui bahwa sistem klasifikasi dokumen berdasarkan abstrak karya akhir mahasiswa Jurusan Teknik Elektro telah dilakukan pengujian menggunakan metode *black box*. Pada sistem klasifikasi dokumen ini terdapat dua user yang berperan dalam menggunakan sistem.

User pertama adalah *user* biasa, *user* ini dapat menginput data berupa dokumen abstrak karya akhir yang telah diformat khusus. *User* yang memiliki pemahaman mengenai proses *Text Mining* akan sangat sesuai untuk sistem klasifikasi dokumen ini. Hasil kategori pada sistem klasifikasi dokumen akan ditampilkan pada suatu *Interface* khusus berbasis web, yang akan mengolah hasil kategori tersebut menjadi lebih cepat tanpa proses *Text Mining* di dalamnya.

User kedua adalah *admin*, yang dapat mengakses *database* abstrak karya akhir, yang mempunyai hak akses untuk mengatur *database*. *User admin* sangat sesuai apabila diberikan pada staf IT yang mengetahui konsep penggunaan sistem dan manajemen *database* khususnya *database MySQL*.

Jadi, berdasarkan pada penjelasan di atas dan kesesuaian dengan Tabel 4.1. bahwa sistem klasifikasi dokumen telah layak digunakan oleh *user* sesuai dengan hasil pengujian sistem klasifikasi dan sesuai dengan analisis kebutuhan *user*.

4.2.2. Pembahasan Hasil Pengujian Algoritma *Naïve Bayes Classifier*

Untuk pembagian *data* pada sistem klasifikasi dokumen adalah sebagai berikut, yaitu:

1. Total seluruh dokumen abstrak yang bersumber dari Jurusan Teknik Elektro Universitas Negeri Jakarta adalah 100 dokumen abstrak.
2. 100 dokumen abstrak tersebut mengalami format khusus yang telah dijelaskan pada Bab III.
3. Pengujian sistem menggunakan metode *K-Fold Cross Validation* dengan membagi data menjadi 10 bagian (*10 Fold*) dengan mempertahankan perbandingan pembagian dokumen *data train* dan *data test* sebesar 9:1.

Pengujian dan besaran angka akurasi dijelaskan sebagai berikut:

- a. Uji ke – 1: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 70% serta nilai *error* sebesar 30%.
- b. Uji ke – 2: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 90% serta nilai *error* sebesar 10%.
- c. Uji ke – 3: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 70% serta nilai *error* sebesar 30%.
- d. Uji ke – 4: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 80% serta nilai *error* sebesar 20%.

- e. Uji ke – 5: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 70% serta nilai *error* sebesar 30%.
 - f. Uji ke – 6: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 100% serta nilai *error* sebesar 0%.
 - g. Uji ke – 7: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 70% serta nilai *error* sebesar 30%.
 - h. Uji ke – 8: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 72% serta nilai *error* sebesar 28%.
 - i. Uji ke – 9: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 80% serta nilai *error* sebesar 20%.
 - j. Uji ke – 10: *Data Train* berjumlah 90 dokumen dan *Data Test* berjumlah 10 dokumen, dengan tingkat akurasi algoritma *Naïve Bayes* sebesar 100% serta nilai *error* sebesar 0%.
4. *Data class* dibagi menjadi empat *class* yaitu kelas pertama dengan nama “Analisis”, kelas kedua dengan nama “Media Pembelajaran”, kelas kedua dengan nama “Penelitian Pendidikan” dan kelas keempat dengan nama “Produk Rekayasa”.

Sesuai dengan Tabel Pengujian, bahwa terdapat waktu eksekusi yang berbeda dari masing-masing dokumen *test*. Menunjukkan bahwa kecepatan proses *Text Mining* sangat bergantung pada proses pemecahan kata di dalam dokumen yang telah di *input*.

Untuk mencari tingkat akurasi pada sistem klasifikasi dokumen, terlebih dahulu sistem klasifikasi dokumen memproses *data test*. Setelah diketahui nilai prediksinya, kemudian membandingkan nilai prediksinya dengan *class label* yang sudah diprediksi. Tingkat akurasi dapat diketahui dengan membagi nilai prediksi yang benar dengan nilai total dari *data testing*. Sedangkan nilai *error* dapat diketahui dengan membagi nilai prediksi yang salah dengan total *data testing*.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian pada Bab IV, maka pada penelitian ini dihasilkan sebuah *software* yang dapat mengklasifikasikan dokumen berbentuk isi abstrak skripsi untuk mengetahui termasuk ke dalam kategori apa isi abstrak tersebut. Berdasarkan pembahasan di Bab IV, sumber data dokumen abstrak pada penelitian klasifikasi dokumen berasal dari Jurusan Teknik Elektro Universitas Negeri Jakarta.

Jumlah data yang diambil berjumlah 100 dokumen, 90 dokumen abstrak digunakan sebagai *data train*, dan 10 dokumen abstrak digunakan sebagai *data test*. Pengujian sistem dilakukan dengan metode *cross validation* dengan hasil sebagai berikut:

Tabel. 5.1 Hasil Pengujian Sistem Klasifikasi Dokumen

Uji Ke -	Akurasi	Rata-Rata Waktu Eksekusi
1	70%	4,2949 sekon
2	90%	4,4857 sekon
3	70%	3,9557 sekon
4	80%	4,2744 sekon
5	80%	4,0553 sekon
6	70%	4,2637 sekon
7	100%	4,2844 sekon
8	70%	4,2628 sekon
9	80%	4,4171 sekon
10	100%	3,7800 sekon
Rata-rata	81%	4,2074 sekon

Dari Tabel 5.1 setelah melakukan pengujian terhadap sistem, didapatkan rata-rata akurasi Sistem Klasifikasi Dokumen sebesar 81% dan rata-rata waktu eksekusi sebesar 4,2074 sekon. Akurasi sebesar 81% merupakan tingkat akurasi yang cukup baik.

Nilai *prior* probabilitas di dapat berdasarkan pada tujuh kata di dalam dokumen abstrak yaitu: “Media”, “Sistem”, “Alat”, “Pembelajaran”, “Analisa”, “Ajar”, dan “Metode”. Kemudian keluaran dari sistem merupakan kategori yang dibagi menjadi empat kategori, yaitu: “Analisis”, “Media Pembelajaran”, “Penelitian Pendidikan” dan “Produk Rekayasa”.

Dalam penelitian sistem klasifikasi dokumen, sistem menggunakan algoritma *Naïve Bayes*, bahasa pemograman PHP, metodologi *waterfall* dan MySQL sebagai *database*. Setelah proses pengembangan telah selesai dilakukan, kemudian sistem dilakukan pengujian kebutuhan fungsional *user* dan didapatkan hasil yang sesuai dengan criteria pengujian kebutuhan *user*.

Berdasarkan pada nilai akurasi yang di dapat dan pada pengujian kebutuhan fungsional *user*, maka dapat disimpulkan bahwa pengembangan sistem Klasifikasi Dokumen Karya Akhir Mahasiswa Menggunakan *Naïve Bayes Classifier* (NBC) Berdasarkan Abstrak Karya Akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta telah berhasil dikembangkan serta dapat digunakan sebagai *tools* untuk mengklasifikasikan dokumen abstrak karya akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta.

5.2. Implikasi Bidang Keteknikan dan Pendidikan

5.2.1. Implikasi Bidang Keteknikan

Dengan adanya aplikasi klasifikasi dokumen abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier* dapat digunakan menjadi bahan demonstrasi dalam mata kuliah Rekayasa Perangkat Lunak (*Software Engineering*), Rekayasa Sistem (*System Engineering*), Kecerdasan Buatan (*Artificial Intelligence*), dan *Data Mining*. Aplikasi ini juga dapat menjadi salah satu contoh aplikasi *Natural Language Processing* yang dapat dikembangkan lagi.

5.2.2. Implikasi Bidang Pendidikan

Dalam bidang pendidikan, khususnya dalam lingkup Jurusan Teknik Elektro Universitas Negeri Jakarta, aplikasi klasifikasi dokumen abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier* ini dapat mendukung pengklasifikasi dokumen abstrak karya akhir di Jurusan Teknik Elektro.

5.3. Saran

Dalam penyusunan skripsi klasifikasi dokumen abstrak karya akhir menggunakan algoritma *Naïve Bayes Classifier* banyak ditemui hambatan yaitu data yang didapatkan tidak sesuai dengan target yang diharapkan. Dokumen abstrak karya akhir Jurusan Teknik Elektro memang sangat terbatas jumlahnya, mengingat dokumen abstrak tersebut sebagian dalam bentuk *hardcopy*, tidak dalam bentuk *softcopy*. Maka, diharapkan dalam penelitian selanjutnya target dokumen, baik dokumen set untuk *data training* dan *data test*. Hambatan lainnya adalah penentuan kata di dalam abstrak skripsi yang akan menjadi acuan untuk

melakukan perhitungan ke dalam sistem NBC. Kata tersebut harus mewakili kategori yang menjadi hasil dari sistem.

Untuk penelitian selanjutnya diharapkan dapat mencoba algoritma yang lain, yang kemungkinan mempunyai tingkat akurasi yang lebih tinggi. Sehingga mempunyai hasil prediksi yang jauh lebih baik.

DAFTAR PUSTAKA

- Agusta, Yudi. 2009. *Feature Selection*. [terhubung berkala].
<http://yudiagusta.wordpress.com/2009/01/13/feature-selection/comment-page-1/> [28 Mei 2014]
- Asian, Jelita. 2007. *Stemming Indonesian*. Jurnal. Australia: School of Computer Science and Information Technology, RMIT University.
- Brammer, Max. 2007. *Principles of Data Mining*. London: Springer
- Destuardi, dkk. 2009. Klasifikasi Emosi untuk Teks Bahasa Indonesia Menggunakan Metode Naïve Bayes. Seminar Nasional Pascasarjana, Institut Teknologi Sepuluh Nopember.
- Domingos, Pedro & Michael Pazzani. 1997. "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss". Netherlands.
- Hamzah, Amir. 2012. *Klasifikasi Teks dengan Naïve Bayes Classifier (NBC) untuk Pengelompokan Teks Berita dan Abstract Akademis*. Fakultas Teknologi Industri, Institut Sains dan Teknologi AKPRIND. Yogyakarta.
- Han, J., & Kamber, M. 2006. *Data Mining: Concepts and Technique, Second Edition*. San Fransisco: Morgan Kaufman Publisher.
- Kao, Anne & Stephen R. Poteet. 2007. *Natural Language Processing and Text Mining*. London: Springer.
- Kusrini, & Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: Penerbit Andi.
- Laplace Estimator*. [terhubung berkala]. http://en.wikipedia.org/wiki/Laplace_estimator [4 Juni 2014]
- Larose, D. T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. New Jersey: John Willey & Sons, Inc.
- Liyantanto. 2010. *Kata Dasar Bahasa Indonesia*. [terhubung berkala].
<http://liyantanto.wordpress.com/2010/12/06/kata-dasar-bahasa-indonesia/>
 [14 Maret 2014].

- Mahendra, I Putu Adhi Kerta. 2008. *Enhanced Confix Stripping Stemmer And Ants Algorithm For Classifying News Document In Indonesian Language*. Jurnal. Surabaya: Fakultas Teknologi Informasi, Institut Teknologi Surabaya.
- Manning, Christopher D. 2009. *An Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Mooney, Raymond J et al. 2014. *UTexas: Natural Language Semantics using Distributional Semantics and Probabilistic Logic*. United States: The University of Texas at Austin.
- Natural Language Processing. 2009. [terhubung berkala] <http://irdaloves.blogspot.com/2009/03/natural-language-processing-nlp.html> [21 Mei 2014]
- Pengembangan Software dengan Metode Waterfall*. 2013. [terhubung berkala]. <http://www.etunas.com/web/pengembangan-software-dengan-metode-waterfall.htm> [20 Maret 2014]
- Poole, D & Alan Mackworth. 2010. *Artificial Intelligence: Natural Language Understanding*. Canada: Canada License.
- Pressman, Roger S. 2010. *Software Engineering: A Practitioner's Approach, Seventh Edition*. New York: McGraw – Hill.
- Pradipta, Sukmana. 2012. *Information Retrieval*. [terhubung berkala]. <http://sukmanapradibtha.wordpress.com/2012/10/02/information-retrieval-stemming-untuk-bahasa-indonesia/> [7 Maret 2014].
- Proses Morfologi*. 2011. [terhubung berkala]. <http://bahasaindosugik.blogspot.com/2011/11/proses-morfologi.html> [14 Maret 2014].
- Rachli, M. 2007. "Email Filtering Menggunakan Naive Bayesian". Bandung: Institut Teknologi Bandung.

- Shadiq, M. Ammar. 2009. Keoptimalan Naïve Bayes dalam Klasifikasi. Jurnal. Bandung: Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, UPI.
- Srivastava, Ashok & Mehran Sahami. 2009. *Text Mining: Classification, Clustering, and Applications*. Florida: Chapman & Hall (CRC Press).
- Suhartono, Derwin. 2013. *Natural Language Processing*. <http://socs.binus.ac.id/2013/06/22/natural-language-processing>. [9 Maret 2014]
- Tan, Ah-Hwee. 1999. "Text Mining: The state of the art and the challenges". Singapore.
- Tan, P. – N., Steinbach, M., & Kumar, V. 2006. *Introduction to Data Mining*. Boston: Pearson Education, Inc.
- Theoridis, Sergious & Konstantinous Koutroumbas. 2009. *Pattern Recognition: Fourth Edition*. Oxford: Academic Press.
- Tim Penyusun. 2012. Buku Pedoman Skripsi/Komprehensif/Karya Inovatif. Jakarta: Universitas Negeri Jakarta.
- Witten, Ian H, dkk. 2005. *Data Mining: Practical Machine Learning Tools and Technique (Second Edition)*. San Frascisco: Morgan Kauffmann Publishers.
- Wu, Xindong. 2009. *The Top Ten Algorithms in Data Mining*. United States: CRC Press.
- Yoga, Andita Dwi. 2010. *Implementasi Modifikasi Enhanced Confix Stripping Stripping Stemmer untuk Bahasa Indonesia dengan Metode Corpus Based Stemming: Seminar Tugas Akhir*. <http://digilib.its.ac.id/public/ITS-Undergraduate-14255-presentationpdf.pdf>. [10 Maret 2014]

LAMPIRAN

1. Perhitungan manual probabilitas pada *Fold 1*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	42	0	2	0	0	3	0	2	Analisa
2	43	0	2	0	1	0	0	1	Analisa
3	49	0	7	2	0	3	0	0	Analisa
4	55	0	0	0	0	3	0	1	Analisa
5	60	0	0	1	0	1	0	13	Analisa
6	61	0	3	5	0	0	0	2	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	5	14	0	0	6	1	10	2	Media Pembelajaran
28	13	9	0	1	6	0	11	0	Media Pembelajaran
29	23	8	0	0	1	3	4	2	Media Pembelajaran
30	29	7	0	0	7	0	7	0	Media Pembelajaran
31	38	4	0	2	1	0	2	0	Media Pembelajaran
32	56	3	0	0	3	0	0	1	Media Pembelajaran
33	84	5	0	0	4	0	0	2	Media Pembelajaran
34	90	1	0	0	8	0	0	1	Media Pembelajaran
35	96	6	0	0	0	1	4	2	Media Pembelajaran
36	100	5	0	0	4	1	1	0	Media Pembelajaran

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
37	8	0	0	0	1	0	4	0	Penelitian Pendidikan
38	9	0	0	0	7	0	7	1	Penelitian Pendidikan
39	11	0	2	0	5	0	6	2	Penelitian Pendidikan
40	12	0	0	0	0	0	9	3	Penelitian Pendidikan
41	16	0	0	0	8	1	12	4	Penelitian Pendidikan
42	17	0	1	0	2	0	9	0	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	14	0	6	0	0	0	0	0	Produk Rekayasa
62	15	0	3	5	0	0	0	0	Produk Rekayasa
63	19	0	0	3	0	0	0	1	Produk Rekayasa
64	20	3	3	0	0	0	0	2	Produk Rekayasa
65	24	0	7	0	0	0	0	2	Produk Rekayasa
66	26	0	0	0	0	2	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	131	133
Media Pembelajaran	10	$= \frac{10}{90}$ $= 0,1111$	160	162
Penelitian Pendidikan	24	$= \frac{24}{90}$ $= 0,2667$	326	328
Produk Rekayasa	28	$= \frac{28}{90}$ $= 0,3111$	161	163
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{133}$ $= 0,0075$	38	$= \frac{38}{133}$ $= 0,2857$	17	$= \frac{17}{133}$ $= 0,1278$
Media Pembelajaran	63	$= \frac{63}{162}$ $= 0,3889$	1	$= \frac{1}{162}$ $= 0,0062$	4	$= \frac{4}{162}$ $= 0,0247$
Penelitian Pendidikan	2	$= \frac{2}{328}$ $= 0,0061$	15	$= \frac{15}{328}$ $= 0,0457$	2	$= \frac{2}{328}$ $= 0,0061$
Produk Rekayasa	4	$= \frac{4}{163}$ $= 0,0245$	63	$= \frac{63}{163}$ $= 0,3865$	58	$= \frac{58}{163}$ $= 0,3558$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/133$ $= 0,0602$	30	$= 30/133$ $= 0,2256$
Media Pembelajaran	41	$= 41/162$ $= 0,2531$	7	$= 7/162$ $= 0,0432$
Penelitian Pendidikan	93	$= 93/328$ $= 0,2835$	20	$= 20/328$ $= 0,0610$
Produk Rekayasa	1	$= 1/163$ $= 0,0061$	8	$= 8/163$ $= 0,0491$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/133$ $= 0,0075$	43	$= 43/133$ $= 0,3233$
Media Pembelajaran	40	$= 40/162$ $= 0,2469$	11	$= 11/162$ $= 0,0679$
Penelitian Pendidikan	168	$= 168/328$ $= 0,5122$	33	$= 33/328$ $= 0,1006$
Produk Rekayasa	4	$= 4/163$ $= 0,0245$	30	$= 30/163$ $= 0,1840$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0075	0,2857	0,1278	0,0602	0,2256	0,0075	0,3233
Media Pembelajaran	0,1111	0,3889	0,0062	0,0247	0,2531	0,0432	0,2469	0,0679
Penelitian Pendidikan	0,2667	0,0061	0,0457	0,0061	0,2835	0,0610	0,5122	0,1006
Produk Rekayasa	0,3111	0,0245	0,3865	0,3558	0,0061	0,0491	0,0245	0,1840

2. Perhitungan manual probabilitas pada *Fold 2*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	55	0	0	0	0	3	0	1	Analisa
5	60	0	0	1	0	1	0	13	Analisa
6	61	0	3	5	0	0	0	2	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	13	9	0	1	6	0	11	0	Media Pembelajaran
30	23	8	0	0	1	3	4	2	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	12	0	0	0	0	0	9	3	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	16	0	0	0	8	1	12	4	Penelitian Pendidikan
42	17	0	1	0	2	0	9	0	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	20	3	3	0	0	0	0	2	Produk Rekayasa
65	24	0	7	0	0	0	0	2	Produk Rekayasa
66	26	0	0	0	0	2	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	115	117
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	161	163
Penelitian Pendidikan	23	$= \frac{23}{90}$ $= 0,2556$	337	339
Produk Rekayasa	30	$= \frac{30}{90}$ $= 0,3333$	165	167
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{117}$ $= 0,0085$	29	$= \frac{29}{117}$ $= 0,2479$	15	$= \frac{15}{117}$ $= 0,1282$
Media Pembelajaran	56	$= \frac{56}{163}$ $= 0,3436$	4	$= \frac{4}{163}$ $= 0,0245$	7	$= \frac{7}{163}$ $= 0,0429$
Penelitian Pendidikan	2	$= \frac{2}{339}$ $= 0,0059$	16	$= \frac{16}{339}$ $= 0,0472$	2	$= \frac{2}{339}$ $= 0,0059$
Produk Rekayasa	4	$= \frac{4}{167}$ $= 0,0240$	69	$= \frac{69}{167}$ $= 0,4132$	50	$= \frac{50}{167}$ $= 0,2994$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	7	$= 7/117$ $= 0,0598$	26	$= 26/117$ $= 0,2222$
Media Pembelajaran	44	$= 44/163$ $= 0,2699$	6	$= 6/163$ $= 0,0368$
Penelitian Pendidikan	94	$= 94/339$ $= 0,2773$	21	$= 21/339$ $= 0,0619$
Produk Rekayasa	1	$= 1/167$ $= 0,0060$	9	$= 9/167$ $= 0,0539$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/117$ $= 0,0085$	43	$= 43/117$ $= 0,3675$
Media Pembelajaran	39	$= 39/163$ $= 0,2393$	12	$= 12/163$ $= 0,0736$
Penelitian Pendidikan	177	$= 177/339$ $= 0,5221$	32	$= 32/339$ $= 0,0944$
Produk Rekayasa	4	$= 4/167$ $= 0,0240$	35	$= 35/167$ $= 0,2096$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0075	0,0085	0,2479	0,1282	0,0598	0,2222	0,0085
Media Pembelajaran	0,1222	0,3889	0,3436	0,0245	0,0429	0,2699	0,0368	0,2393
Penelitian Pendidikan	0,2556	0,0061	0,0059	0,0472	0,0059	0,2773	0,0619	0,5221
Produk Rekayasa	0,3333	0,0245	0,0240	0,4132	0,2994	0,0060	0,0539	0,0240

3. Perhitungan manual probabilitas pada *Fold 3*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	62	0	1	0	0	0	0	0	Analisa
8	63	0	0	0	0	2	0	2	Analisa
9	64	0	0	0	0	1	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	23	8	0	0	1	3	4	2	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	18	0	0	0	4	0	12	1	Penelitian Pendidikan
44	21	0	0	0	2	1	7	0	Penelitian Pendidikan
45	22	0	3	1	8	1	15	4	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	27	0	1	0	0	0	0	0	Produk Rekayasa
68	28	0	0	0	0	0	2	1	Produk Rekayasa
69	30	0	2	0	0	0	0	0	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	109	111
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	167	169
Penelitian Pendidikan	23	$= \frac{23}{90}$ $= 0,2556$	323	325
Produk Rekayasa	30	$= \frac{30}{90}$ $= 0,3333$	163	165
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{111}$ $= 0,0090$	37	$= \frac{37}{111}$ $= 0,3333$	11	$= \frac{11}{111}$ $= 0,0991$
Media Pembelajaran	61	$= \frac{61}{169}$ $= 0,3609$	4	$= \frac{4}{169}$ $= 0,0237$	6	$= \frac{6}{169}$ $= 0,0355$
Penelitian Pendidikan	2	$= \frac{2}{325}$ $= 0,0062$	17	$= \frac{17}{325}$ $= 0,0523$	2	$= \frac{2}{325}$ $= 0,0062$
Produk Rekayasa	1	$= \frac{1}{165}$ $= 0,0061$	68	$= \frac{68}{165}$ $= 0,4121$	58	$= \frac{58}{165}$ $= 0,3515$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= \frac{8}{111}$ $= 0,0721$	28	$= \frac{28}{111}$ $= 0,2523$
Media Pembelajaran	44	$= \frac{44}{169}$ $= 0,2604$	7	$= \frac{7}{169}$ $= 0,0414$
Penelitian Pendidikan	97	$= \frac{97}{325}$ $= 0,2985$	20	$= \frac{20}{325}$ $= 0,0615$
Produk Rekayasa	1	$= \frac{1}{165}$ $= 0,0061$	7	$= \frac{7}{165}$ $= 0,0424$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{111}$ $= 0,0090$	30	$= \frac{30}{111}$ $= 0,2703$
Media Pembelajaran	38	$= \frac{38}{169}$ $= 0,2249$	14	$= \frac{14}{169}$ $= 0,0828$
Penelitian Pendidikan	164	$= \frac{164}{325}$ $= 0,5046$	28	$= \frac{28}{325}$ $= 0,0862$
Produk Rekayasa	4	$= \frac{4}{165}$ $= 0,0242$	31	$= \frac{31}{165}$ $= 0,1879$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0090	0,3333	0,0991	0,0721	0,2523	0,0090	0,2703
Media Pembelajaran	0,1222	0,3609	0,0237	0,0355	0,2604	0,0414	0,2249	0,0828
Penelitian Pendidikan	0,2556	0,0062	0,0523	0,0062	0,2985	0,0615	0,5046	0,0862
Produk Rekayasa	0,3333	0,0061	0,4121	0,3515	0,0061	0,0424	0,0242	0,1879

4. Perhitungan manual probabilitas pada *Fold 4*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	66	0	7	0	0	0	0	1	Analisa
11	67	0	4	0	0	3	0	2	Analisa
12	68	0	0	2	0	1	0	1	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	29	7	0	0	7	0	7	0	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	32	0	0	0	11	1	9	0	Penelitian Pendidikan
47	33	0	0	0	5	0	6	2	Penelitian Pendidikan
48	36	0	5	0	0	0	10	2	Penelitian Pendidikan
49	44	0	0	0	0	2	14	1	Penelitian Pendidikan
50	46	0	0	0	0	1	2	1	Penelitian Pendidikan
51	50	0	0	0	3	1	7	1	Penelitian Pendidikan
52	53	1	0	0	0	1	3	1	Penelitian Pendidikan
53	54	0	0	0	7	0	0	1	Penelitian Pendidikan
54	65	0	0	0	1	1	0	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	31	0	1	2	0	0	0	2	Produk Rekayasa
71	34	0	1	0	0	2	0	1	Produk Rekayasa
72	35	0	4	4	0	2	0	2	Produk Rekayasa
73	37	0	2	1	0	0	0	2	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	130	132
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	176	178
Penelitian Pendidikan	23	$= \frac{23}{90}$ $= 0,2556$	313	315
Produk Rekayasa	30	$= \frac{30}{90}$ $= 0,3333$	177	179
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{132}$ $= 0,0076$	39	$= \frac{39}{132}$ $= 0,2955$	17	$= \frac{17}{132}$ $= 0,1288$
Media Pembelajaran	62	$= \frac{62}{178}$ $= 0,3483$	4	$= \frac{4}{178}$ $= 0,0225$	7	$= \frac{7}{178}$ $= 0,0393$
Penelitian Pendidikan	2	$= \frac{2}{315}$ $= 0,0063$	15	$= \frac{15}{315}$ $= 0,0476$	1	$= \frac{1}{315}$ $= 0,0032$
Produk Rekayasa	4	$= \frac{4}{179}$ $= 0,0223$	75	$= \frac{75}{179}$ $= 0,4190$	58	$= \frac{58}{179}$ $= 0,3240$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/132$ $= 0,0606$	29	$= 29/132$ $= 0,2197$
Media Pembelajaran	49	$= 49/178$ $= 0,2753$	4	$= 4/178$ $= 0,0225$
Penelitian Pendidikan	93	$= 93/315$ $= 0,2952$	19	$= 19/315$ $= 0,0603$
Produk Rekayasa	1	$= 1/179$ $= 0,0056$	9	$= 9/179$ $= 0,0503$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/132$ $= 0,0076$	42	$= 42/132$ $= 0,3182$
Media Pembelajaran	45	$= 45/178$ $= 0,2528$	12	$= 12/178$ $= 0,0674$
Penelitian Pendidikan	160	$= 160/315$ $= 0,5079$	30	$= 30/315$ $= 0,0952$
Produk Rekayasa	2	$= 2/179$ $= 0,0112$	35	$= 35/179$ $= 0,1955$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0076	0,2955	0,1288	0,0606	0,2197	0,0076	0,3182
Media Pembelajaran	0,1222	0,3483	0,0225	0,0393	0,2753	0,0225	0,2528	0,0674
Penelitian Pendidikan	0,2556	0,0063	0,0476	0,0032	0,2952	0,0603	0,5079	0,0952
Produk Rekayasa	0,3333	0,0223	0,4190	0,3240	0,0056	0,0503	0,0112	0,1955

5. Perhitungan manual probabilitas pada *Fold 5*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	70	0	0	0	0	2	0	2	Analisa
14	71	0	0	0	0	0	0	1	Analisa
15	72	0	1	4	0	0	0	0	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	38	4	0	2	1	0	2	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	36	0	5	0	0	0	10	2	Penelitian Pendidikan
50	44	0	0	0	0	2	14	1	Penelitian Pendidikan
51	46	0	0	0	0	1	2	1	Penelitian Pendidikan
52	50	0	0	0	3	1	7	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	41	0	0	0	0	0	0	1	Produk Rekayasa
75	45	0	4	4	0	0	0	2	Produk Rekayasa
76	47	0	0	0	0	1	0	2	Produk Rekayasa
77	48	0	4	0	0	0	1	1	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	117	119
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	173	175
Penelitian Pendidikan	24	$= \frac{24}{90}$ $= 0,2667$	338	340
Produk Rekayasa	29	$= \frac{29}{90}$ $= 0,3222$	157	159
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{119}$ $= 0,0084$	29	$= \frac{29}{119}$ $= 0,2437$	15	$= \frac{15}{119}$ $= 0,1261$
Media Pembelajaran	63	$= \frac{63}{175}$ $= 0,3600$	4	$= \frac{4}{175}$ $= 0,0229$	7	$= \frac{7}{175}$ $= 0,0400$
Penelitian Pendidikan	2	$= \frac{2}{340}$ $= 0,0059$	18	$= \frac{18}{340}$ $= 0,0529$	2	$= \frac{2}{340}$ $= 0,0059$
Produk Rekayasa	4	$= \frac{4}{159}$ $= 0,0252$	70	$= \frac{70}{159}$ $= 0,4403$	51	$= \frac{51}{159}$ $= 0,3208$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= \frac{8}{119}$ $= 0,0672$	28	$= \frac{28}{119}$ $= 0,2353$
Media Pembelajaran	43	$= \frac{43}{175}$ $= 0,2457$	7	$= \frac{7}{175}$ $= 0,0400$
Penelitian Pendidikan	91	$= \frac{91}{340}$ $= 0,2676$	20	$= \frac{20}{340}$ $= 0,0588$
Produk Rekayasa	1	$= \frac{1}{159}$ $= 0,0063$	5	$= \frac{5}{159}$ $= 0,0314$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{119}$ $= 0,0084$	42	$= \frac{42}{119}$ $= 0,3529$
Media Pembelajaran	42	$= \frac{42}{175}$ $= 0,2400$	14	$= \frac{14}{175}$ $= 0,0800$
Penelitian Pendidikan	179	$= \frac{179}{340}$ $= 0,5265$	33	$= \frac{33}{340}$ $= 0,0971$
Produk Rekayasa	4	$= \frac{4}{159}$ $= 0,0252$	29	$= \frac{29}{159}$ $= 0,1824$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0084	0,2437	0,1261	0,0672	0,2353	0,0084	0,3529
Media Pembelajaran	0,1222	0,3600	0,0229	0,0400	0,2457	0,0400	0,2400	0,0800
Penelitian Pendidikan	0,2667	0,0059	0,0529	0,0059	0,2676	0,0588	0,5265	0,0971
Produk Rekayasa	0,3222	0,0252	0,4403	0,3208	0,0063	0,0314	0,0252	0,1824

6. Perhitungan manual probabilitas pada *Fold 6*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	73	0	0	0	0	1	0	2	Analisa
17	78	0	0	1	0	1	0	2	Analisa
18	79	0	1	0	0	1	0	2	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	56	3	0	0	3	0	0	1	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	46	0	0	0	0	1	2	1	Penelitian Pendidikan
52	50	0	0	0	3	1	7	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	31	0	1	2	0	0	0	2	Produk Rekayasa
75	34	0	1	0	0	2	0	1	Produk Rekayasa
76	35	0	4	4	0	2	0	2	Produk Rekayasa
77	37	0	2	1	0	0	0	2	Produk Rekayasa
78	51	0	3	1	0	0	0	1	Produk Rekayasa
79	52	0	0	4	0	0	0	0	Produk Rekayasa
80	57	0	0	1	0	0	0	1	Produk Rekayasa
81	58	0	0	7	0	0	0	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	128	130
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	185	187
Penelitian Pendidikan	24	$= \frac{24}{90}$ $= 0,2667$	338	340
Produk Rekayasa	29	$= \frac{29}{90}$ $= 0,3222$	163	165
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{130}$ $= 0,0077$	39	$= \frac{39}{130}$ $= 0,3000$	13	$= \frac{13}{130}$ $= 0,1000$
Media Pembelajaran	66	$= \frac{66}{187}$ $= 0,3529$	4	$= \frac{4}{187}$ $= 0,0214$	5	$= \frac{5}{187}$ $= 0,0267$
Penelitian Pendidikan	2	$= \frac{2}{340}$ $= 0,0059$	13	$= \frac{13}{340}$ $= 0,0382$	2	$= \frac{2}{340}$ $= 0,0059$
Produk Rekayasa	4	$= \frac{4}{165}$ $= 0,0242$	70	$= \frac{70}{165}$ $= 0,4242$	54	$= \frac{54}{165}$ $= 0,3273$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/130$ $= 0,0615$	30	$= 30/130$ $= 0,2308$
Media Pembelajaran	49	$= 49/187$ $= 0,2620$	7	$= 7/187$ $= 0,0374$
Penelitian Pendidikan	107	$= 107/340$ $= 0,3147$	19	$= 19/340$ $= 0,0559$
Produk Rekayasa	1	$= 1/165$ $= 0,0061$	8	$= 8/165$ $= 0,0485$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/130$ $= 0,0077$	43	$= 43/130$ $= 0,3308$
Media Pembelajaran	47	$= 47/187$ $= 0,2513$	14	$= 14/187$ $= 0,0749$
Penelitian Pendidikan	170	$= 170/340$ $= 0,5000$	32	$= 32/340$ $= 0,0941$
Produk Rekayasa	3	$= 3/165$ $= 0,0182$	30	$= 30/165$ $= 0,1818$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0077	0,3000	0,1000	0,0615	0,2308	0,0077	0,3308
Media Pembelajaran	0,1222	0,3529	0,0214	0,0267	0,2620	0,0374	0,2513	0,0749
Penelitian Pendidikan	0,2667	0,0059	0,0382	0,0059	0,3147	0,0559	0,5000	0,0941
Produk Rekayasa	0,3222	0,0242	0,4242	0,3273	0,0061	0,0485	0,0182	0,1818

7. Perhitungan manual probabilitas pada *Fold 7*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	80	0	0	0	0	0	0	1	Analisa
20	81	0	0	0	0	1	0	1	Analisa
21	83	0	7	1	0	0	0	0	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	84	5	0	0	4	0	0	2	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	36	0	5	0	0	0	10	2	Penelitian Pendidikan
52	44	0	0	0	0	2	14	1	Penelitian Pendidikan
53	53	1	0	0	0	1	3	1	Penelitian Pendidikan
54	54	0	0	0	7	0	0	1	Penelitian Pendidikan
55	65	0	0	0	1	1	0	1	Penelitian Pendidikan
56	69	0	0	0	0	0	2	0	Penelitian Pendidikan
57	75	0	0	0	5	2	5	1	Penelitian Pendidikan
58	76	0	2	0	11	2	9	1	Penelitian Pendidikan
59	89	0	0	0	0	4	5	2	Penelitian Pendidikan
60	92	0	1	0	8	0	7	2	Penelitian Pendidikan
61	97	0	0	0	4	1	7	1	Penelitian Pendidikan
62	2	0	5	0	0	0	0	1	Produk Rekayasa
63	3	0	10	0	0	1	0	0	Produk Rekayasa
64	10	0	0	0	0	0	0	5	Produk Rekayasa
65	14	0	6	0	0	0	0	0	Produk Rekayasa
66	15	0	3	5	0	0	0	0	Produk Rekayasa
67	19	0	0	3	0	0	0	1	Produk Rekayasa
68	20	3	3	0	0	0	0	2	Produk Rekayasa
69	24	0	7	0	0	0	0	2	Produk Rekayasa
70	26	0	0	0	0	2	0	1	Produk Rekayasa
71	27	0	1	0	0	0	0	0	Produk Rekayasa
72	28	0	0	0	0	0	2	1	Produk Rekayasa
73	30	0	2	0	0	0	0	0	Produk Rekayasa
74	31	0	1	2	0	0	0	2	Produk Rekayasa
75	34	0	1	0	0	2	0	1	Produk Rekayasa
76	35	0	4	4	0	2	0	2	Produk Rekayasa
77	37	0	2	1	0	0	0	2	Produk Rekayasa
78	41	0	0	0	0	0	0	1	Produk Rekayasa
79	45	0	4	4	0	0	0	2	Produk Rekayasa
80	47	0	0	0	0	1	0	2	Produk Rekayasa
81	48	0	4	0	0	0	1	1	Produk Rekayasa
82	59	0	5	0	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	74	0	3	0	0	0	0	1	Produk Rekayasa
84	77	0	0	2	0	0	0	0	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	127	129
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	187	189
Penelitian Pendidikan	24	$= \frac{24}{90}$ $= 0,2667$	356	358
Produk Rekayasa	29	$= \frac{29}{90}$ $= 0,3222$	164	166
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{129}$ $= 0,0078$	39	$= \frac{39}{129}$ $= 0,3023$	16	$= \frac{16}{129}$ $= 0,1240$
Media Pembelajaran	67	$= \frac{67}{189}$ $= 0,3545$	4	$= \frac{4}{189}$ $= 0,0212$	7	$= \frac{7}{189}$ $= 0,0370$
Penelitian Pendidikan	2	$= \frac{2}{358}$ $= 0,0056$	18	$= \frac{18}{358}$ $= 0,0503$	2	$= \frac{2}{358}$ $= 0,0056$
Produk Rekayasa	4	$= \frac{4}{166}$ $= 0,0241$	75	$= \frac{75}{166}$ $= 0,4518$	45	$= \frac{45}{166}$ $= 0,2711$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/129$ $= 0,0620$	29	$= 29/129$ $= 0,2248$
Media Pembelajaran	47	$= 47/189$ $= 0,2487$	7	$= 7/189$ $= 0,0370$
Penelitian Pendidikan	104	$= 104/358$ $= 0,2905$	19	$= 19/358$ $= 0,0531$
Produk Rekayasa	1	$= 1/166$ $= 0,0060$	9	$= 9/166$ $= 0,0542$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/129$ $= 0,0078$	40	$= 40/129$ $= 0,3101$
Media Pembelajaran	49	$= 49/189$ $= 0,2593$	13	$= 13/189$ $= 0,0688$
Penelitian Pendidikan	185	$= 185/358$ $= 0,5168$	33	$= 33/358$ $= 0,0922$
Produk Rekayasa	4	$= 4/166$ $= 0,0241$	33	$= 33/166$ $= 0,1988$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0078	0,3023	0,1240	0,0620	0,2248	0,0078	0,3101
Media Pembelajaran	0,1222	0,3545	0,0212	0,0370	0,2487	0,0370	0,2593	0,0688
Penelitian Pendidikan	0,2667	0,0056	0,0503	0,0056	0,2905	0,0531	0,5168	0,0922
Produk Rekayasa	0,3222	0,0241	0,4518	0,2711	0,0060	0,0542	0,0241	0,1988

8. Perhitungan manual probabilitas pada *Fold 8*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	85	0	0	0	0	2	0	1	Analisa
23	88	0	0	0	6	0	0	1	Analisa
24	91	0	0	0	0	0	0	1	Analisa
25	93	0	1	0	0	2	0	0	Analisa
26	94	0	1	0	0	2	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	56	3	0	0	3	0	0	1	Media Pembelajaran
35	90	1	0	0	8	0	0	1	Media Pembelajaran
36	96	6	0	0	0	1	4	2	Media Pembelajaran
37	100	5	0	0	4	1	1	0	Media Pembelajaran
38	6	0	3	0	11	0	16	1	Penelitian Pendidikan
39	7	0	0	0	3	1	10	1	Penelitian Pendidikan
40	8	0	0	0	1	0	4	0	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	9	0	0	0	7	0	7	1	Penelitian Pendidikan
42	11	0	2	0	5	0	6	2	Penelitian Pendidikan
43	12	0	0	0	0	0	9	3	Penelitian Pendidikan
44	16	0	0	0	8	1	12	4	Penelitian Pendidikan
45	17	0	1	0	2	0	9	0	Penelitian Pendidikan
46	18	0	0	0	4	0	12	1	Penelitian Pendidikan
47	21	0	0	0	2	1	7	0	Penelitian Pendidikan
48	22	0	3	1	8	1	15	4	Penelitian Pendidikan
49	32	0	0	0	11	1	9	0	Penelitian Pendidikan
50	33	0	0	0	5	0	6	2	Penelitian Pendidikan
51	36	0	5	0	0	0	10	2	Penelitian Pendidikan
52	44	0	0	0	0	2	14	1	Penelitian Pendidikan
53	46	0	0	0	0	1	2	1	Penelitian Pendidikan
54	50	0	0	0	3	1	7	1	Penelitian Pendidikan
55	69	0	0	0	0	0	2	0	Penelitian Pendidikan
56	75	0	0	0	5	2	5	1	Penelitian Pendidikan
57	76	0	2	0	11	2	9	1	Penelitian Pendidikan
58	89	0	0	0	0	4	5	2	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa
82	52	0	0	4	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	82	0	1	4	0	0	0	0	Produk Rekayasa
86	86	0	2	4	0	0	0	0	Produk Rekayasa
87	87	0	2	0	0	0	0	2	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= \frac{26}{90}$ $= 0,2889$	127	129
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	183	185
Penelitian Pendidikan	23	$= \frac{23}{90}$ $= 0,2556$	355	357
Produk Rekayasa	30	$= \frac{30}{90}$ $= 0,3333$	172	174
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{129}$ $= 0,0078$	33	$= \frac{33}{129}$ $= 0,2558$	16	$= \frac{16}{129}$ $= 0,1240$
Media Pembelajaran	65	$= \frac{65}{185}$ $= 0,3514$	4	$= \frac{4}{185}$ $= 0,0216$	7	$= \frac{7}{185}$ $= 0,0378$
Penelitian Pendidikan	1	$= \frac{1}{357}$ $= 0,0028$	18	$= \frac{18}{357}$ $= 0,0504$	2	$= \frac{2}{357}$ $= 0,0056$
Produk Rekayasa	4	$= \frac{4}{174}$ $= 0,0230$	70	$= \frac{70}{174}$ $= 0,4023$	56	$= \frac{56}{174}$ $= 0,3218$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/129$ $= 0,0620$	31	$= 31/129$ $= 0,2403$
Media Pembelajaran	46	$= 46/185$ $= 0,2486$	7	$= 7/185$ $= 0,0378$
Penelitian Pendidikan	99	$= 99/357$ $= 0,2773$	19	$= 19/357$ $= 0,0532$
Produk Rekayasa	1	$= 1/174$ $= 0,0057$	9	$= 9/174$ $= 0,0517$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/129$ $= 0,0078$	44	$= 44/129$ $= 0,3411$
Media Pembelajaran	49	$= 49/185$ $= 0,2649$	12	$= 12/185$ $= 0,0649$
Penelitian Pendidikan	191	$= 191/357$ $= 0,5350$	32	$= 32/357$ $= 0,0896$
Produk Rekayasa	4	$= 4/174$ $= 0,0230$	35	$= 35/174$ $= 0,2011$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0078	0,2558	0,1240	0,0620	0,2403	0,0078	0,3411
Media Pembelajaran	0,1222	0,3514	0,0216	0,0378	0,2486	0,0378	0,2649	0,0649
Penelitian Pendidikan	0,2556	0,0028	0,0504	0,0056	0,2773	0,0532	0,5350	0,0896
Produk Rekayasa	0,3333	0,0230	0,4023	0,3218	0,0057	0,0517	0,0230	0,2011

9. Perhitungan manual probabilitas pada *Fold 9*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	80	0	0	0	0	0	0	1	Analisa
23	81	0	0	0	0	1	0	1	Analisa
24	83	0	7	1	0	0	0	0	Analisa
25	91	0	0	0	0	0	0	1	Analisa
26	93	0	1	0	0	2	0	0	Analisa
27	94	0	1	0	0	2	0	1	Analisa
28	1	7	3	1	5	0	0	0	Media Pembelajaran
29	4	0	0	2	4	0	9	3	Media Pembelajaran
30	5	14	0	0	6	1	10	2	Media Pembelajaran
31	13	9	0	1	6	0	11	0	Media Pembelajaran
32	23	8	0	0	1	3	4	2	Media Pembelajaran
33	29	7	0	0	7	0	7	0	Media Pembelajaran
34	38	4	0	2	1	0	2	0	Media Pembelajaran
35	56	3	0	0	3	0	0	1	Media Pembelajaran
36	84	5	0	0	4	0	0	2	Media Pembelajaran
37	96	6	0	0	0	1	4	2	Media Pembelajaran
38	100	5	0	0	4	1	1	0	Media Pembelajaran
39	6	0	3	0	11	0	16	1	Penelitian Pendidikan
40	7	0	0	0	3	1	10	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	8	0	0	0	1	0	4	0	Penelitian Pendidikan
42	9	0	0	0	7	0	7	1	Penelitian Pendidikan
43	11	0	2	0	5	0	6	2	Penelitian Pendidikan
44	12	0	0	0	0	0	9	3	Penelitian Pendidikan
45	16	0	0	0	8	1	12	4	Penelitian Pendidikan
46	17	0	1	0	2	0	9	0	Penelitian Pendidikan
47	18	0	0	0	4	0	12	1	Penelitian Pendidikan
48	21	0	0	0	2	1	7	0	Penelitian Pendidikan
49	22	0	3	1	8	1	15	4	Penelitian Pendidikan
50	32	0	0	0	11	1	9	0	Penelitian Pendidikan
51	33	0	0	0	5	0	6	2	Penelitian Pendidikan
52	36	0	5	0	0	0	10	2	Penelitian Pendidikan
53	44	0	0	0	0	2	14	1	Penelitian Pendidikan
54	46	0	0	0	0	1	2	1	Penelitian Pendidikan
55	50	0	0	0	3	1	7	1	Penelitian Pendidikan
56	53	1	0	0	0	1	3	1	Penelitian Pendidikan
57	54	0	0	0	7	0	0	1	Penelitian Pendidikan
58	65	0	0	0	1	1	0	1	Penelitian Pendidikan
59	92	0	1	0	8	0	7	2	Penelitian Pendidikan
60	97	0	0	0	4	1	7	1	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa
82	52	0	0	4	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	59	0	5	0	0	0	0	0	Produk Rekayasa
86	74	0	3	0	0	0	0	1	Produk Rekayasa
87	77	0	0	2	0	0	0	0	Produk Rekayasa
88	95	0	0	8	0	0	0	2	Produk Rekayasa
89	98	0	6	3	0	0	0	1	Produk Rekayasa
90	99	0	2	4	0	0	0	0	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	27	$= \frac{27}{90}$ $= 0,3000$	128	130
Media Pembelajaran	11	$= \frac{11}{90}$ $= 0,1222$	184	186
Penelitian Pendidikan	22	$= \frac{22}{90}$ $= 0,2444$	321	323
Produk Rekayasa	30	$= \frac{30}{90}$ $= 0,3333$	168	170
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= \frac{1}{130}$ $= 0,0077$	40	$= \frac{40}{130}$ $= 0,3077$	17	$= \frac{17}{130}$ $= 0,1308$
Media Pembelajaran	69	$= \frac{69}{186}$ $= 0,3710$	4	$= \frac{4}{186}$ $= 0,0215$	7	$= \frac{7}{186}$ $= 0,0376$
Penelitian Pendidikan	2	$= \frac{2}{323}$ $= 0,0062$	16	$= \frac{16}{323}$ $= 0,0495$	2	$= \frac{2}{323}$ $= 0,0062$
Produk Rekayasa	4	$= \frac{4}{170}$ $= 0,0235$	73	$= \frac{73}{170}$ $= 0,4294$	50	$= \frac{50}{170}$ $= 0,2941$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	2	$= 2/130$ $= 0,0154$	30	$= 30/130$ $= 0,2308$
Media Pembelajaran	42	$= 42/186$ $= 0,2258$	7	$= 7/186$ $= 0,0376$
Penelitian Pendidikan	91	$= 91/323$ $= 0,2817$	13	$= 13/323$ $= 0,0402$
Produk Rekayasa	1	$= 1/170$ $= 0,0059$	9	$= 9/170$ $= 0,0529$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/130$ $= 0,0077$	44	$= 44/130$ $= 0,3385$
Media Pembelajaran	49	$= 49/186$ $= 0,2634$	13	$= 13/186$ $= 0,0699$
Penelitian Pendidikan	173	$= 173/323$ $= 0,5356$	31	$= 31/323$ $= 0,0960$
Produk Rekayasa	4	$= 4/170$ $= 0,0235$	34	$= 34/170$ $= 0,2000$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,3000	0,0077	0,3077	0,1308	0,0154	0,2308	0,0077	0,3385
Media Pembelajaran	0,1222	0,3710	0,0215	0,0376	0,2258	0,0376	0,2634	0,0699
Penelitian Pendidikan	0,2444	0,0062	0,0495	0,0062	0,2817	0,0402	0,5356	0,0960
Produk Rekayasa	0,3333	0,0235	0,4294	0,2941	0,0059	0,0529	0,0235	0,2000

10. Perhitungan manual probabilitas pada *Fold 10*

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
1	25	0	1	0	0	0	0	0	Analisa
2	39	0	0	0	0	0	0	1	Analisa
3	40	0	1	0	0	2	0	2	Analisa
4	42	0	2	0	0	3	0	2	Analisa
5	43	0	2	0	1	0	0	1	Analisa
6	49	0	7	2	0	3	0	0	Analisa
7	55	0	0	0	0	3	0	1	Analisa
8	60	0	0	1	0	1	0	13	Analisa
9	61	0	3	5	0	0	0	2	Analisa
10	62	0	1	0	0	0	0	0	Analisa
11	63	0	0	0	0	2	0	2	Analisa
12	64	0	0	0	0	1	0	2	Analisa
13	66	0	7	0	0	0	0	1	Analisa
14	67	0	4	0	0	3	0	2	Analisa
15	68	0	0	2	0	1	0	1	Analisa
16	70	0	0	0	0	2	0	2	Analisa
17	71	0	0	0	0	0	0	1	Analisa
18	72	0	1	4	0	0	0	0	Analisa
19	73	0	0	0	0	1	0	2	Analisa
20	78	0	0	1	0	1	0	2	Analisa
21	79	0	1	0	0	1	0	2	Analisa
22	80	0	0	0	0	0	0	1	Analisa
23	81	0	0	0	0	1	0	1	Analisa
24	83	0	7	1	0	0	0	0	Analisa
25	85	0	0	0	0	2	0	1	Analisa
26	88	0	0	0	6	0	0	1	Analisa
27	1	7	3	1	5	0	0	0	Media Pembelajaran
28	4	0	0	2	4	0	9	3	Media Pembelajaran
29	5	14	0	0	6	1	10	2	Media Pembelajaran
30	13	9	0	1	6	0	11	0	Media Pembelajaran
31	23	8	0	0	1	3	4	2	Media Pembelajaran
32	29	7	0	0	7	0	7	0	Media Pembelajaran
33	38	4	0	2	1	0	2	0	Media Pembelajaran
34	56	3	0	0	3	0	0	1	Media Pembelajaran
35	84	5	0	0	4	0	0	2	Media Pembelajaran
36	90	1	0	0	8	0	0	1	Media Pembelajaran
37	6	0	3	0	11	0	16	1	Penelitian Pendidikan
38	7	0	0	0	3	1	10	1	Penelitian Pendidikan
39	8	0	0	0	1	0	4	0	Penelitian Pendidikan
40	9	0	0	0	7	0	7	1	Penelitian Pendidikan

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
41	11	0	2	0	5	0	6	2	Penelitian Pendidikan
42	12	0	0	0	0	0	9	3	Penelitian Pendidikan
43	16	0	0	0	8	1	12	4	Penelitian Pendidikan
44	17	0	1	0	2	0	9	0	Penelitian Pendidikan
45	18	0	0	0	4	0	12	1	Penelitian Pendidikan
46	21	0	0	0	2	1	7	0	Penelitian Pendidikan
47	22	0	3	1	8	1	15	4	Penelitian Pendidikan
48	32	0	0	0	11	1	9	0	Penelitian Pendidikan
49	33	0	0	0	5	0	6	2	Penelitian Pendidikan
50	36	0	5	0	0	0	10	2	Penelitian Pendidikan
51	44	0	0	0	0	2	14	1	Penelitian Pendidikan
52	46	0	0	0	0	1	2	1	Penelitian Pendidikan
53	50	0	0	0	3	1	7	1	Penelitian Pendidikan
54	53	1	0	0	0	1	3	1	Penelitian Pendidikan
55	54	0	0	0	7	0	0	1	Penelitian Pendidikan
56	65	0	0	0	1	1	0	1	Penelitian Pendidikan
57	69	0	0	0	0	0	2	0	Penelitian Pendidikan
58	75	0	0	0	5	2	5	1	Penelitian Pendidikan
59	76	0	2	0	11	2	9	1	Penelitian Pendidikan
60	89	0	0	0	0	4	5	2	Penelitian Pendidikan
61	2	0	5	0	0	0	0	1	Produk Rekayasa
62	3	0	10	0	0	1	0	0	Produk Rekayasa
63	10	0	0	0	0	0	0	5	Produk Rekayasa
64	14	0	6	0	0	0	0	0	Produk Rekayasa
65	15	0	3	5	0	0	0	0	Produk Rekayasa
66	19	0	0	3	0	0	0	1	Produk Rekayasa
67	20	3	3	0	0	0	0	2	Produk Rekayasa
68	24	0	7	0	0	0	0	2	Produk Rekayasa
69	26	0	0	0	0	2	0	1	Produk Rekayasa
70	27	0	1	0	0	0	0	0	Produk Rekayasa
71	28	0	0	0	0	0	2	1	Produk Rekayasa
72	30	0	2	0	0	0	0	0	Produk Rekayasa
73	31	0	1	2	0	0	0	2	Produk Rekayasa
74	34	0	1	0	0	2	0	1	Produk Rekayasa
75	35	0	4	4	0	2	0	2	Produk Rekayasa
76	37	0	2	1	0	0	0	2	Produk Rekayasa
77	41	0	0	0	0	0	0	1	Produk Rekayasa
78	45	0	4	4	0	0	0	2	Produk Rekayasa
79	47	0	0	0	0	1	0	2	Produk Rekayasa
80	48	0	4	0	0	0	1	1	Produk Rekayasa
81	51	0	3	1	0	0	0	1	Produk Rekayasa
82	52	0	0	4	0	0	0	0	Produk Rekayasa

No	id ke-	#1	#2	#3	#4	#5	#6	#7	Class
83	57	0	0	1	0	0	0	1	Produk Rekayasa
84	58	0	0	7	0	0	0	1	Produk Rekayasa
85	59	0	5	0	0	0	0	0	Produk Rekayasa
86	74	0	3	0	0	0	0	1	Produk Rekayasa
87	77	0	0	2	0	0	0	0	Produk Rekayasa
88	82	0	1	4	0	0	0	0	Produk Rekayasa
89	86	0	2	4	0	0	0	0	Produk Rekayasa
90	87	0	2	0	0	0	0	2	Produk Rekayasa

Class	Jumlah	Pci	Ni	Laplace
Analisa	26	$= 26/90$ $= 0,2889$	130	132
Media Pembelajaran	10	$= 10/90$ $= 0,1111$	170	172
Penelitian Pendidikan	24	$= 24/90$ $= 0,2667$	341	343
Produk Rekayasa	30	$= 30/90$ $= 0,3333$	157	159
Total	90			

Class	Media		Sistem		Alat	
	Sum	Prior	Sum	Prior	Sum	Prior
Analisa	1	$= 1/132$ $= 0,0076$	38	$= 38/132$ $= 0,2879$	17	$= 17/132$ $= 0,1288$
Media Pembelajaran	59	$= 59/172$ $= 0,3430$	4	$= 4/172$ $= 0,0233$	7	$= 7/172$ $= 0,0407$
Penelitian Pendidikan	2	$= 2/343$ $= 0,0058$	17	$= 17/343$ $= 0,0496$	2	$= 2/343$ $= 0,0058$
Produk Rekayasa	4	$= 4/159$ $= 0,0252$	70	$= 70/159$ $= 0,4403$	43	$= 43/159$ $= 0,2704$

Class	Pembelajaran		Analisis	
	Sum	Prior	Sum	Prior
Analisa	8	$= 8/132$ $= 0,0606$	28	$= 28/132$ $= 0,2121$
Media Pembelajaran	46	$= 46/172$ $= 0,2674$	5	$= 5/172$ $= 0,0291$
Penelitian Pendidikan	95	$= 95/343$ $= 0,2770$	20	$= 20/343$ $= 0,0583$
Produk Rekayasa	1	$= 1/159$ $= 0,0063$	9	$= 9/159$ $= 0,0566$

Class	Ajar		Metode	
	Sum	Prior	Sum	Prior
Analisa	1	$= 1/132$ $= 0,0076$	44	$= 44/132$ $= 0,3333$
Media Pembelajaran	44	$= 44/172$ $= 0,2558$	12	$= 12/172$ $= 0,0698$
Penelitian Pendidikan	180	$= 180/343$ $= 0,5248$	32	$= 32/343$ $= 0,0933$
Produk Rekayasa	4	$= 4/159$ $= 0,0252$	33	$= 33/159$ $= 0,2075$

Class	P(Ci)	P(Mda)	P(Stm)	P(Alt)	P(Pmb)	P(Anl)	P(Ajr)	P(Mtd)
Analisa	0,2889	0,0076	0,2879	0,1288	0,0606	0,2121	0,0076	0,3333
Media Pembelajaran	0,1111	0,3430	0,0233	0,0407	0,2674	0,0291	0,2558	0,0698
Penelitian Pendidikan	0,2667	0,0058	0,0496	0,0058	0,2770	0,0583	0,5248	0,0933
Produk Rekayasa	0,3333	0,0252	0,4403	0,2704	0,0063	0,0566	0,0252	0,2075

LEMBAR PERNYATAAN
VALIDASI KATEGORI ABSTRAK SECARA MANUAL

Lampiran: 100 Isi Abstrak

Dengan ini, saya sebagai validator semantik menyatakan bahwa:

Nama : Nur Indah Pratiwi

No. Registrasi : 5235107412

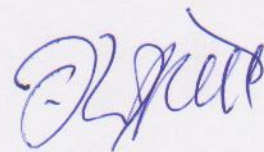
Program Studi : Pendidikan Teknik Informatika dan Komputer

Judul Skripsi : Klasifikasi Dokumen Karya Akhir Mahasiswa Menggunakan Naïve Bayes Classifier (NBC) Berdasarkan Abstrak Karya Akhir di Jurusan Teknik Elektro Universitas Negeri Jakarta.

Telah melakukan penelitian, mengklasifikasi 100 dokumen abstrak menjadi 4 kategori: Media Pembelajaran, Analisa, Produk Rekayasa, dan Penelitian Pendidikan secara manual dan hasil klasifikasi telah divalidasi.

Jakarta, Juli 2014

Validator Semantik



Widodo, M.Kom

NIP.197203252005011002

TENTANG PENULIS



Nur Indah Pratiwi, lahir pada tanggal 5 Agustus 1992 di Jakarta. Lahir dari keluarga biasa tetapi dibesarkan dan mendapat didikan dengan cara yang luar biasa.

Sejarah pendidikan yang ditempuh oleh penulis adalah sebagai berikut: pada tahun 1998-2004 menempuh jenjang pendidikan Sekolah Dasar di SDN 01 Pagi Jakarta. Tahun 2004-2007 menempuh jenjang menengah pertama di SMPN 27 Jakarta. Serta, tahun 2007-2010 menempuh jenjang menengah atas di SMAN 71 Jakarta.

Penulis memiliki minat bidang keilmuan yang berbeda dengan dua saudara kandung lainnya. Semenjak diperkenalkan komputer pada tahun 2003, anak ketiga dari tiga bersaudara ini tertarik pada komputer dan bentuk teknologi di dalamnya. Perkenalan penulis dengan tim *Web Developer* pada lingkungan perkuliahan semakin membuat penulis tertarik pada bidang profesi *System Analyst* dan *Programmer*. Penulis dapat dihubungi di nurindahpratiwi71@gmail.com dan dapat di temui pada laman <http://wiwaaw.com>. Terima kasih kepada para pembaca yang telah bersedia membaca karya akhir ini. Semoga bermanfaat.