



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



**IMPLEMENTASI ALGORITMA NAÏVE BAYES
UNTUK KLASIFIKASI BERITA TEKNOLOGI
INFORMASI**

LAPORAN SKRIPSI

**POLITEKNIK
NEGERI
JAKARTA**

**ALVRILA ZURRIYA ABRAR
4617010002**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2021



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



**IMPLEMENTASI ALGORITMA NAÏVE BAYES
UNTUK KLASIFIKASI BERITA TEKNOLOGI
INFORMASI**

LAPORAN SKRIPSI

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan untuk
Memperoleh Diploma Empat Politeknik**

**POLITEKNIK
NEGERI
JAKARTA**

ALVRILA ZURRIYA ABRAR

4617010002

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2021



HALAMAN PERNYATAAN ORISINALITAS

Skripsi/Tesis/Disertasi ini adalah hasil karya saya sendiri dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Alvrila Zurriya Abrar
NIM : 4617010002
Tanggal : Kamis, 03 Juni 2021
Tanda Tangan :

- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh :

Nama : Alvrila Zurriya Abrar
NIM : 4617010002
Program Studi : Teknik Informatika
Judul Skripsi : Implementasi Algoritma *Naïve Bayes* untuk Klasifikasi Berita TI

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Senin, Tanggal 14, Bulan Juni, Tahun 2021 dan dinyatakan **LULUS**.

Disahkan oleh

Pembimbing : Dr. Dewi Yanti Liliana, S.Kom., M.Kom.
Penguji I : Risna Sari, S.Kom., M.Ti.
Penguji II : Hata Maulana, S.Si., M.Ti.
Penguji III : Euis Oktavianti, S.Si., M.Ti.

()
()
()
()

Mengetahui :

Jurusan Teknik Informatika dan Komputer

Ketua



Mauldy Laya, S.Kom., M.Kom.
NIP. 197802112009121003



KATA PENGANTAR

Puji syukur atas kehadiran Allah SAW yang telah memberikan rahmat dan hidayahNya, sehingga penulis dapat menyelesaikan laporan skripsi dengan judul **RANCANG BANGUN SISTEM KLASIFIKASI BERITA TEKNOLOGI INFORMASI MENGGUNAKAN *MACHINE LEARNING* BERBASIS *WEBSITE***. Penulisan laporan skripsi ini ditujukan dalam rangka memenuhi salah satu syarat untuk memperoleh gelar Sarjana Terapan Politeknik. Penulis menyadari bahwa, tanpa bimbingan dan dukungan dari beberapa belah pihak sejak masa perkuliahan sampe pada penyusunan laporan tugas akhir.

Oleh karena itu penulis mengucapkan terimakasih kepada:

- a. Ibu Dr. Dewi Yanti Liliana, S.Kom, M.Kom, selaku dosen pembimbing skripsi yang telah meluangkan waktu, tenaga dan pikiran untuk membimbing dan membantu penulis dalam Menyusun laporan skripsi.
- b. Pihak Kominfo yang telah membantu penulis dalam memenuhi kebutuhan sistem dan melakukan validasi data penelitian.
- c. Orang tua penulis yang setiap saat mendoakan penulis, serta memberikan bantuan moral maupun material.
- d. Sahabat dan teman-teman yang telah mendukung penulis dalam menyelesaikan laporan skripsi ini.

Akhir kata, penulis berharap semoga Allah SWT membalas segala kebaikan semua pihak yang telah membantu. Semoga laporan skripsi ini dapat bermanfaat bagi pengembangan ilmu.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Alvrila Zurriya Abrar
NIM : 4617010002
Program Studi : Teknik Informatika
Jurusan : Teknik Informatika dan Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta **Hak Bebas Royalti Noneklusif (*Non-exclusive Royalty – Free Right*)** atas karya ilmiah saya yang berjudul:

Rancang Bangun Sistem Klasifikasi Berita Teknologi Informasi Menggunakan *Machine Learning* Berbasis *Website*.

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneklusif ini Politeknik Negeri Jakarta berhak menyimpan mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik HakCipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Depok, Pada tanggal: 02 Mei 2021

Yang menyatakan

(Alvrila Zurriya Abrar)

- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Sistem Klasifikasi Berita TI Menggunakan *Machine Learning* Berbasis Website

Abstrak

Kementerian Komunikasi dan Informatika (Kominfo) adalah salah satu kementerian yang bergerak di bidang komunikasi dan informatika, yang mempunyai tugas menyelenggarakan urusan di bidang komunikasi dan informatika dalam pemerintahan. Dengan terus bertambahnya jumlah berita yang beredar, kominfo juga dihadapkan dengan berbagai macam permasalahan. Salah satunya adalah berita TI ataupun Non TI. Untuk itu diperlukan sistem yang mampu menangani berita dibidang TI ataupun Non TI. Sistem yang dibangun mampu memberikan berita yang valid. Berita valid atau tidak valid tersebut diperoleh dari proses pengumpulan data dari situs-situs berita yang sudah disediakan oleh kominfo sendiri. Kemudian data tersebut diproses menggunakan algoritma Naïve Bayes. Naïve Bayes biasa digunakan untuk mengklasifikasi beberapa bidang, seperti mencari tingkat keakuratan yang tinggi. Algoritma Naïve Bayes tersebut diimplementasikan ke dalam aplikasi website untuk menentukan keakuratan konten berita. Dengan demikian implementasi algoritma Naïve Bayes ini mampu menjadi solusi dalam mengatasi masalah yang ada. Hasil evaluasi kinerja model klasifikasi berita TI dengan algoritma Naive Bayes menghasilkan nilai *precision* 91.28%, *accuracy* 79.56% dan *recall* 91.28%.

Kata kunci: Berita TI, Klasifikasi, Naïve Bayes

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS	ii
LEMBAR PENGESAHAN	iii
KATA PENGANTAR	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	v
ABSTRAK	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan dan Manfaat	2
1.5 Metode Pelaksanaan Skripsi	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Penelitian Sejenis	5
2.2 Klasifikasi	9
2.3 Text Mining	9
2.4 Text Preprocessing.....	10
2.5 Naïve Bayes	11
2.6 Berita.....	12
2.7 Black Box.....	12
2.8 Confusion Matrix	13
2.9 Representasi Fitur Teks.....	14
2.10 Rapid Application Development	14
BAB III PERANCANGAN DAN REALISASI	16
3.1 Perancangan Program Aplikasi.....	16
3.2 Cara Kerja Aplikasi.....	16
3.2.1 Deskripsi Program Aplikasi.....	22

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

3.2.1.1 Data <i>Preprocessing</i>	22
3.2.1.2 Analisis Kebutuhan Sistem	23
3.2.2 Rancangan Program Aplikasi	23
3.2.3 Desain <i>User Interface</i>	24
3.3 Realisasi Program Aplikasi.....	28
3.3.1 Realisasi Model Klasifikasi.....	28
3.4 Realisasi Model Aplikasi	33
3.4.1 Tampilan Utama	33
3.4.2 Tampilan Klasifikasi Teks Berita	34
3.4.3 Tampilan Hasil Klasifikasi Teks Berita <i>Naïve Bayes</i>	35
3.4.4 Tampilan Klasifikasi URL Berita <i>Naïve Bayes</i>	36
3.4.5 Tampilan Hasil Klasifikasi URL Berita <i>Naïve Bayes</i>	36
3.4.6 Tampilan Evaluasi Model.....	37
BAB IV PEMBAHASAN.....	38
4.1 Pengujian.....	38
4.2 Deskripsi Pengujian	38
4.3 Prosedur Pengujian	38
4.4 Data Hasil Pengujian.....	39
4.5 Analisa Data/Evaluasi.....	47
4.5.1 Evaluasi Model.....	47
4.5.2 Evaluasi Aplikasi	50
BAB V PENUTUP.....	51
5.1 Kesimpulan.....	51
5.2 Saran	51
DAFTAR PUSTAKA	52
DAFTAR RIWAYAT HIDUP	



DAFTAR GAMBAR

Gambar 2.1 Proses <i>text mining</i>	10
Gambar 2.2 Tahap text preprocessing.....	10
Gambar 2.3 Teknik <i>black box</i>	12
Gambar 2.4 Confusion matrix	13
Gambar 3.1 Perancangan aplikasi.....	16
Gambar 3.2 Flowchart system pada aplikasi.....	17
Gambar 3.3 Flowchart process pada penggunaan aplikasi	18
Gambar 3.4 Flowchart process <i>preprocessing</i>	19
Gambar 3.5 Flowchart klasifikasi berita	20
Gambar 3.6 Flowchart klasifikasi kategori	21
Gambar 3.7 <i>Use case diagram</i>	23
Gambar 3.8 Halaman <i>home</i>	24
Gambar 3.9 Halaman klasifikasi URL berita NB	25
Gambar 3.10 Halaman hasil klasifikasi NB	26
Gambar 3.11 Halaman klasifikasi URL berita NB	27
Gambar 3.12 Halaman hasil klasifikasi URL berita NB	28
Gambar 3.13 Tampilan file CSV	29
Gambar 3.14 Tampilan fungsi preprocessing	30
Gambar 3.15 Tampilan <i>code NB modeling</i> 1	31
Gambar 3.16 Tampilan NB modelling 1	31
Gambar 3.17 Tampilan fungsi modelling 1	31
Gambar 3.18 Tampilan fungsi modelling 2	32
Gambar 3.19 Tampilan Utama Pada Website	33
Gambar 3.20 Tampilan Klasifikasi Teks Berita Naïve Bayes	34
Gambar 3.21 Tampilan code backend Naïve Bayes.....	34
Gambar 3.22 Tampilan Hasil Klasifikasi Teks Berita NB.....	35
Gambar 3.23 Tampilan Klasifikasi URL Berita NB	36
Gambar 3.24 Tampilan Hasil Klasifikasi URL Berita NB	36
Gambar 3.25 Tampilan Hasil Evaluasi Model NB	37
Gambar 4.1 Tampilan Hasil Evaluasi Model NB	49
Gambar 4.2 Tampilan Hasil Evaluasi Model NB	49

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR TABEL

Tabel 1 Penelitian sejenis	5
Tabel 2 Jumlah Data Perlabel	29
Tabel 3 Data hasil pengujian	39
Tabel 4 Hasil evaluasi <i>confusion matrix</i>	47
Tabel 5 Hasil evaluasi <i>confusion matrix</i> kategori	48
Table 6 Nilai <i>precision</i> dan <i>recall</i>	48
Table 7 Tampilan F-1 <i>score</i>	49



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



BAB I PENDAHULUAN

1.1 Latar Belakang

Berita adalah teks yang berisi tentang segala peristiwa, isi pernyataan dan keterangan mengenai suatu peristiwa, yang berisi informasi aktual dan terkini. Berita biasanya disebarkan melalui berbagai macam media informasi, seperti radio, televisi, internet, situs web dan banyak sekali berita yang beredar bisa kita ketahui melalui media sosial (Shabrina Alfari, 2018).

Text mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dimana *text mining* merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar. Selain klasifikasi *text mining* juga digunakan untuk menangani masalah clustering information dan extraction dan information retrieval (Bambang Kurniawan, Syahril Effendi, Opim Salim Sitompul, 2012).

Kementerian Kominfo merupakan perangkat Pemerintah Republik Indonesia ini membidangi urusan yang ruang lingkupnya disebutkan dalam Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 yaitu Informasi dan Komunikasi. Kementerian Komunikasi dan Informatika mempunyai tugas menyelenggarakan urusan pemerintah di bidang komunikasi dan informatika untuk membantu Presiden dalam menyelenggarakan pemerintahan negara.

Metode yang digunakan untuk mengidentifikasi berita yang berhubungan dengan pemilihan presiden 2019 adalah *Naïve Bayes Classifier* (NBC). Metode NBC dipilih karena dapat mengidentifikasi berita dengan akurasi sebesar 91.36% dengan tambahan metode mutual information dan information gain sebagai seleksi fitur dan unigram sebagai ekstraksi fiturnya. Data yang digunakan dalam penelitian ini akan diambil dari Twitter dengan crawling data menggunakan Bahasa Pemrograman Python (Azizah Zain, 2020).

Dengan latar belakang diatas, pihak tim Aptika sangat memerlukan pengklasifikasian kategori berita secara otomatis. Maka, dibuat penelitian ini yang

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

berjudul “Implementasi Algoritma Naïve Bayes untuk Klasifikasi Berita TI” untuk menghasilkan sistem yang dapat mengklasifikasi kategori berita di bidang TI secara otomatis dengan menggunakan algoritma Naïve Bayes. Agar pencarian berita sudah bisa menggunakan klasifikasi kategori berita, yang menghasilkan berita tersebut termasuk dalam berita TI atau Non TI.

1.1 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat dirumuskan masalah dalam pembuatan sistem ini adalah:

- a. Bagaimana merancang sistem klasifikasi berita TI berbasis *website*?
- b. Bagaimana mengimplementasikan klasifikasi berita TI menggunakan algoritma *Naive Bayes*?
- c. Bagaimana mengevaluasi kinerja model klasifikasi berita TI dengan algoritma *Naive Bayes*?

1.2 Batasan Masalah

Batasan masalah dalam pembuatan sistem ini adalah:

- a. Artikel berita menggunakan bahasa Indonesia.
- b. Artikel berita diambil dari situs Ditjen Aptika dan 5 situs portal berita Indonesia yang teratas berdasarkan situs <https://www.similarweb.com/> per 10 Maret 2021, yaitu Kompas, Tribunnews, Detik, JPNN, dan IDN Times.
- c. Sistem website dikembangkan menggunakan framework Flask yang berbasis bahasa Python.

1.3 Tujuan dan Manfaat

Tujuan pembuatan sistem ini adalah:

- a. Merancang sistem klasifikasi berita TI berbasis *website*.
- b. Mengklasifikasi berita TI menggunakan algoritma *Naive Bayes*.
- c. Mengevaluasi kinerja model klasifikasi berita TI dengan algoritma *Naive Bayes*.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Manfaat yang dapat diperoleh dari pembuatan sistem ini di antaranya adalah:

- a. Menghasilkan sistem klasifikasi berita TI secara otomatis yang dapat diakses melalui *browser* internet.
- b. Mempermudah tim Aptika melakukan klasifikasi kategori berita TI.

1.4 Metode Pelaksanaan Skripsi

Pelaksanaan skripsi dilakukan dengan dua metode, yaitu metode untuk pengumpulan data dan metode untuk pengembangan sistem.

a. Metode Pengumpulan Data

Pengumpulan data dilaksanakan dengan melakukan studi kepustakaan untuk mempelajari penelitian-penelitian yang terdahulu dan teori-teori yang dapat mendukung penelitian dan mengumpulkan data artikel berita di situs portal berita Indonesia menggunakan *library* BeautifulSoup4 pada Python.

b. Metode Pengembangan Sistem

Pengembangan sistem dilakukan dengan metode *Rapid Application Development* atau RAD, sebuah metodologi yang menekankan pada proses pembuatan aplikasi berdasarkan pembuatan prototype, iterasi dan feedback yang berulang-ulang. Dengan begitu, aplikasi yang dibuat bisa dikembangkan dan diperbaiki dengan cepat. Berikut tahapan-tahapan dalam metode RAD di antaranya:

- Menentukan kebutuhan proyek
Pada tahap ini, tim perlu menentukan kebutuhan yang ingin dipenuhi dari sebuah proyek. Tahap awal ini berguna untuk memberikan gambaran luas pada proyek yang ingin dikerjakan.
- Membuat *Prototype*
Setelah proyek telah diberi batasan kebutuhan, tim mulai membangun model dan *prototype* awal aplikasi.
Tujuannya untuk mengecek apakah *prototype* yang dibuat sudah sesuai dengan kebutuhan klien. Tahapan ini bisa saja dilakukan berulang-ulang. Karena, melibatkan *user* untuk *testing* dan memberikan feedback.
- Proses pengembangan dan pengumpulan *feedback*
Tahap ini adalah ketika *coding* aplikasi, pengujian sistem, dan integrasi unit

terjadi mengubah *prototype* menjadi model yang bekerja. *Feedback* yang dimaksud disini mencakup fitur, fungsi, visual dan juga *interface* program yang sedang dikembangkan. Setelah itu, *prototype* akan dikembangkan lagi sampai klien memberikan persetujuan untuk finalisasi produk. Seperti yang disebutkan sebelumnya, kedua tahapan ini akan diulang terus-menerus, sampai hasilnya sesuai dengan keinginan klien.

- Implementasi dan finalisasi produk
Tahap ini adalah ketika tim pengembang mengkoreksi kekurangan yang mungkin terjadi ketika proses pengembangan aplikasi. Tugas ini termasuk melakukan optimasi untuk stabilitas aplikasinya, memperbaiki *interface* hingga melakukan *maintance* dan menyusun dokumentasi.



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta





BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan proses implementasi dan pembahasan sistem diatas, maka kesimpulan yang dapat diambil yaitu:

- a. Pengambilan data yang dipakai diambil dari *dataset* telah berhasil.
Serta sistem klasifikasi berita TI berbasis *website* yang dikembangkan sudah bekerja sesuai dengan yang diinginkan.
- b. Penelitian ini menghasilkan sistem klasifikasi konten berita yang digunakan untuk melakukan prediksi terhadap suatu berita dengan menerapkan Naïve Bayes dan *Library* dengan berbasis bahasa *Python* untuk pembuatan model tersebut.
- c. Hasil kinerja model klasifikasi berita TI dengan algoritma Naïve Bayes tersebut menghasilkan nilai *precision* 91.28%, *accuracy* 79.56% dan *recall* 91.28%.

5.2 Saran

Dari penelitian yang dilakukan, saran yang dapat diberikan adalah:

- a. Pemilihan data yang digunakan sebagai data uji dan data latih harus lebih akurat untuk melihat hasil dari *Naïve Bayes Classifier*.
- b. Melakukan klasifikasi dengan metode lain agar sistem bisa menggunakan pencarian teks berita dengan menggunakan judul berita, tidak harus dengan isi dari teks berita tersebut.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR PUSTAKA

[Apa itu Text Mining? \(binus.ac.id\)](https://binus.ac.id)

Delta Sierra. 2019. Algoritma TF-IDF ([Algoritma TF—IDF. Tulisan ini adalah bagian dari project... | by Delta Sierra | Medium](#))

Handi, Ujang. *Pengujian Dan Implementasi Sistem*, https://elib.unikom.ac.id/files/disk1/440/jbptunikompp-gdl-ujanghandi-21990-5-unikom_u-v.pdf, diakses pada 02 Juni 2021.

Hanifah, Umi. 2016. *Penggunaan Metode Black Box Pada Pengujian Sistem Informasi Surat Keluar Masuk..* Universitas Pembangunan Nasional “Veteran” Jawa Timur.

HKALABS. 2018. Berkenalan dengan scikit-learn (part 1) – preparations ([Berkenalan dengan scikit-learn \(Part 1\) – Preparations – Just Type and Run \(wordpress.com\)](#))

Ibrahim, Muhammad. 2020. “Penerapan Algoritma Naïve Bayes Classifier Untuk Mendeteksi Tingkat Kredibilitas Hoax/Fake News Pada Sosial Media di Indonesia Berbasis Android (Studi Kasus : Kantor Tribun Medan) dalam RESOLUSI : Rekayasa Teknik Informatik dan Informasi Volume 1 (hlm. 9 – 17).

Kayser, V. and Blind, K. (2017) ‘Extending the knowledge base of foresight: The contribution of text mining’, *Technological Forecasting and Social Change*, 116, pp. 208–215. doi: 10.1016/j.techfore.2016.10.017.

Kuliahkomputer. 2018. Pengujian Dengan Confusion Matrix ([Pengujian Dengan Confusion Matrix - Kuliah Komputer Blogger](#))

Kurniawan, Bambang. 2012. *Klasifikasi Konten Berita Dengan Metode Text Mining* dalam JURNAL DUNIA TEKNOLOGI INFORMASI Volume 1 (hlm. 14 – 19).

[Mengapa & Bagaimana menggunakan algoritma Naive Bayes dalam industri yang diatur dengan sklearn | Python + kode \(ichi.pro\)](#)

Nabamita Deb, Vishesh Jha, Alok K Panjiyar, Roshan Kr Gupta. 2020. A Comparative Analysis Of News Categorization Using Machine Learning Approaches. VOLUME 9

RA Juperta. BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM. 14.10113084_REDIT ADRITO JUPERTA_BAB 4.pdf (unikom.ac.id)

Risa Wati. 2020. PENERAPAN ALGORITMA NAIVE BAYES DAN PARTICLE SWARM OPTIMIZATION UNTUK KLASIFIKASI

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BERITA HOAX PADA MEDIA SOSIAL. VOL. 5. NO. 2

Shabrina Alfari. 2018. Teks Berita: Pengertian, Ciri-Ciri & Struktur | Bahasa Indonesia Kelas 8. [Teks Berita: Pengertian, Ciri-Ciri & Struktur | Bahasa Indonesia Kelas 8 \(ruangguru.com\)](#)

TEORI TEXT MINING DAN WEB MINING (binus.ac.id)

[TEXT MINING CENTER: 4 TAHAP PROSES TEXT MINING \(textmining-center.blogspot.com\)](#)

Utami, P. D., & Sari, R. 2018. *Filtering Hoax Menggunakan Naive Bayes Classifier*. *MULTINETICS*, 4(1), 57–61. <https://doi.org/10.32722/multinetics.v4i1.1179>

Wahyu Nurjaya WK , Yusrina Adani. 2018. PENERAPAN ALGORITMA NAÏVE BAYES UNTUK MEMPREDIKSI KEPUTUSAN CALON NASABAH DAN NASABAH TETAP BANK BRI SYARIAH MENERIMA PENAWARAN PROGRAM DEPOSITO BERJANGKA. Vol 8 No 1

Zain, Azizah. 2020. *Identifikasi Tweet Hoax yang Berhubungan dengan Pemilihan Presiden 2019 Menggunakan Naive Bayes Classifier*. <https://openlibrary.telkomuniversity.ac.id/pustaka/156894/identifikasi-tweet-hoax-yang-berhubungan-dengan-pemilihan-presiden-2019-menggunakan-na-ve-bayes-classifier.html>



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

POLITEKNIK
NEGERI
JAKARTA