

Klasifikasi *Intent* pada Chatbot Terapi
Menggunakan *Multinomial Naive Bayes* dan *Mutual Information*

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

Nur Sanifa Arisya

NIM : 09021181621027

JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA

2021

LEMBAR PENGESAHAN SKRIPSI

**KLASIFIKASI *INTENT* PADA CHATBOT TERAPI
MENGUNAKAN *MULTINOMIAL NAIVE BAYES* DAN
*MUTUAL INFORMATION***

Oleh :


Nur Sanifa Arisya

NIM : 09021181621027

Palembang, 23 Agustus 2021

Mengetahui,

Ketua Jurusan Teknik Informatika,


Alvi Syahmi Utami, M.Kom
NIP. 197812222006042003

Pembimbing


Dr. Abdiansah, S.Kom, M.Cs
NIP. 198410012009121005

TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari selasa tanggal 26 Juli 2021 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

N a m a : Nur Sanifa Arisya

N I M : 09021181621027

Judul : Klasifikasi *Intent* Pada Chatbot Terapi Menggunakan *Multinomial Naive Bayes* dan *Mutual Information*

1. Ketua Penguji

Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003



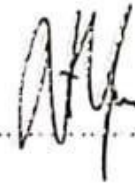
2. Pembimbing I

Dr. Abdiansah., S.Kom., M.Cs
NIP. 198410012009121005



3. Penguji I

Novi Yusliani, M.T.
NIP. 198211082012122001



4. Penguji II

M Qurhanul Rizqie, S.Kom, M.T.
NIP. 1671060312870008



Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Nur Sanifa Arisya
NIM : 09021181621027
Program Studi : Teknik Informatika Reguler
Judul Skripsi : Klasifikasi *Intent* Pada Chatbot Terapi Menggunakan
Multinomial Naive Bayes dan *Mutual Information*

Hasil Pengecekan Software *iThenticate /Turnitin* : 17 %

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 23 Agustus 2021



Nur Sanifa Arisya

NIM. 09021181621027

MOTTO DAN PERSEMBAHAN

“Sesungguhnya sesudah kesulitan itu ada kemudahan.”

(QS. Al-Insyirah, 94:6)

“Iman tanpa ilmu bagaikan lentera di tangan bayi. Namun ilmu tanpa iman bagaikan lentera di tangan pencuri”

(Buya Hamka)

Kupersembahkan Karya Tulis ini kepada :

- Orang Tuaku
- Adikku
- Dosen Pembimbing
- Sahabat-sahabat
- Fakultas Ilmu Komputer, Unsri

ABSTRACT

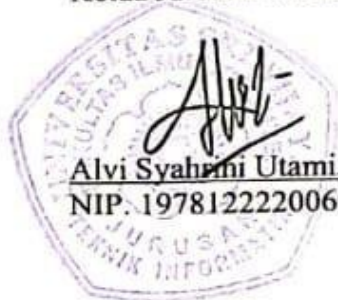
Chatbot is a program that designed to be able to interact with humans through messages or voice. The chatbot will identify the intent that will understand and recognize the statements that the user enters into the chatbot. However, the problem that often arises in chatbots is that sometimes chatbots respond to inappropriate dialogue interactions. Therefore, the classification of intents is one way to be able to categorize an intent that helps provide an appropriate response to dialogue interactions. The Multinomial Naive Bayes (MNB) method is a method that is widely used, especially in document classification. However, the more data, the more features will be processed, so the MNB process will take longer. To overcome this problem, the Mutual Information method is used to reduce the number of features. The purpose of this study was to determine the performance of the MNB classification by selecting the Mutual Information feature and the MNB classification method without selecting the Mutual Information feature. The results of the tests that have been carried out show that the MNB method by selecting the Mutual Information feature has better performance than the MNB method without the Mutual Information feature selection with an accuracy of 0,27 % and 0,53 %

Keyword : Intent Classification, Mutual Information, Stres, Multinomial Naive Bayes, Raelated Discussion

Palembang, 23 Agustus 2021

Mengetahui,

Ketua Jurusan Teknik Informatika,



Alvi Syahuni Utami, M.Kom
NIP. 197812222006042003

Pembimbing,



Dr. Abdiansah., S.Kom., M.Cs
NIP. 198410012009121005

ABSTRAK

Chatbot merupakan sebuah program yang dirancang untuk dapat berinteraksi dengan manusia melalui pesan atau suara. Chatbot akan mengidentifikasi intent yang akan memahami dan mengenali isi pernyataan yang dimasukkan oleh pengguna kepada chatbot. Namun masalah yang sering muncul dalam chatbot yaitu terkadang chatbot memberikan balasan inetraksi dialog yang tidak sesuai. Maka dari itu, klasifikasi *intent* merupakan salah satu cara untuk dapat mengkatagorikan suatu *intent* yang membantu memberikan balasan interaksi dialog yang sesuai. Metode *Multinomial Naive Bayes* (MNB) merupakan metode yang banyak di gunakan khususnya pada klasifikasi dokumen . Namun semakin banyak data maka semakin banyak pula jumlah fitur yang akan di proses sehingga semakin lama pula proses MNB. Untuk mengatasi masalah tersebut, digunakan metode *Mutual Information* untuk mengurangi jumlah fitur. Tujuan penelitian ini adalah untuk mengetahui kinerja yang dihasilkan metode klasifikasi MNB dengan seleksi fitur *Mutual Information* dan metode klasifikasi MNB tanpa seleksi fitur *Mutual Information*. Hasil pengujian yang telah dilakukan menunjukkan bahwa metode MNB dengan seleksi fitur *Mutual Information* memiliki kinerja yang kurang baik dibandingkan dengan metode MNB tanpa seleksi fitur *Mutual Information* dengan akurasi 0,27 % dan 0,53 %.

Kata Kunci : Klasifikasi *intent*, *Mutual Information*, Stres, *Multinomial Naive Bayes*, Diskusi Terkait

Palembang, 23 Agustus 2021

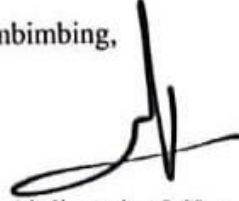
Mengetahui,

Ketua Jurusan Teknik Informatika,



Alvi Syahfani Utami, M.Kom
NIP. 197812222006042003

Pembimbing,



Dr. Abdiansah., S.Kom., M.Cs
NIP. 198410012009121005

KATA PENGANTAR



Puji syukur kepada Allah SWT atas segala rahmat dan karunia-Nya penulis dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas Akhir ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan program Strata-1 Program Studi Teknik Informatika pada Fakultas Ilmu Komputer, Universitas Sriwijaya.

Penulis ingin mengucapkan terima kasih kepada seluruh pihak yang telah memberikan bantuan dan dukungan baik dalam materil maupun moril selama proses pembuatan tugas akhir ini. Adapun penulis merujuk secara khusus pihak yang telah membantu sebagai berikut :

1. Ayahku Rokani dan Ibuku Kartini, selaku orang tuaku tersayang, dan Adikku Soni. Serta keponakan-keponakanku yang selalu menjadi teman serta penghibur dikala sulit.
2. Bapak Jaidan Jauhari, S.pd., M.T. selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Alvi Syahrini Utami, M.Kom selaku Ketua Jurusan Teknik Informatika.
4. Bapak Muhammad Fachrurrozi, S.SI., M.T. selaku pembimbing akademik yang telah memberikan bimbingan serta arahan dalam urusan akademik.
5. Dr. Abdiansyah., S.Kom., M.Cs. selaku pembimbing tugas akhir yang telah memberikan arahan, bimbingan, kritik dan saran kepada penulis untuk menyelesaikan tugas akhir.
6. Ibu Novi Yusliani, M.T. dan Bapak M. Qqurhanul Rizqie, S.Kom., M.T selaku Dosen Penguji I dan Dosen Penguji II yang telah memberikan masukan dan arahan kepada penulis dalam menyempurnakan tugas akhir ini.
7. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu kepada penulis selama masa perkuliahan.
8. Staff Jurusan Teknik Informatika beserta staff Fakultas Ilmu Komputer yang telah membantu dalam urusan administrasi dan akademik selama masa perkuliahan.

9. Sahabat sekaligus teman tetanggaku, Mbak Dita dan Suci Adelia. Terima kasih sudah bersedia menemani dan mau berbagi keluh kesah selama masa kos dan perkuliahan ini. Semoga Allah melancarkan urusan kita dan senantiasa mengeratkan persahabatan kita.
10. Sahabat-sahabat ku selama masa perkuliahan, Syamila Hanun, Rima Melati, Najla Afifah dan Suhartini. Terima kasih sudah senantiasa menemani selama masa perkuliahan, memberikan semangat serta motivasi dalam proses penulisan tugas akhir ini.
11. Teman-teman Teknik Informatika Reguler 2016, terima kasih atas jalinan pertemanannya selama ini, dan senantiasa berbagi informasi bersama. Semoga Allah memudahkan untuk menyelesaikan studi S-1 ini.
12. Semua pihak lainnya yang tidak dapat disebutkan satu persatu yang telah berperan bagi penulis dalam menyelesaikan tugas akhir, Terima Kasih banyak atas semuanya.

Semoga tugas akhir ini dapat bermanfaat bagi penulis khususnya maupun pembaca pada umumnya. Serta dapat menjadi referensi dan rujukan bagi hal-hal yang bermanfaat. Penulis menyadari bahwa dalam proses penyelesaian tugas akhir ini, terdapat beberapa penjelasan yang kurang sempurna. Oleh karena itu penulis mengharapkan kritik dan saran yang membangun agar tugas akhir ini menjadi karya tulis yang sempurna supaya terciptanya bekal pengetahuan yang baik bagi peneliti di masa depan.

Palembang, 23 Agustus 2021



Nur Sanifa Arisya

DAFTAR ISI

Halaman

HALAMAN JUDUL	i
HALAMAN PENGESAHAN.....	ii
DAFTAR ISI.....	iii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xii
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah.....	I-4
1.4 Tujuan Penelitian.....	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-6
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Stres	II-1
2.2.2 Klasifikasi <i>Intent</i>	II-1
2.2.3 Chatbot.....	II-1
2.2.4 Text Pre-processing	II-2

2.2.3 Pembobotan TF-IDF	I-4
2.2.4 <i>Mutual Information</i>	II-5
2.2.5 <i>Naive Bayes Classifier : Multinomial Naive Bayes</i>	II-6
2.2.6 <i>K-Fold Cross Validation</i>	II-7
2.2.7 <i>Confusion Matrix</i>	II-8
2.2.8 <i>Rational Unified Process</i>	II-9
2.3 Penelitian lain yang Relevan	II-11
2.4 Kesimpulan	II-12
BAB III METODE PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data.....	III-1
3.2.1 Jenis dan Sumber Data.....	III-1
3.2.2 Metode Pengumpulan Data	III-1
3.3 Tahapan Penelitian	III-3
3.3.1 Menentukan Kerangka Kerja Penelitian	III-3
3.3.2 Menentukan Kriteria Pengujian.....	III-7
3.3.3 Menetapkan Format Data Pengujian	III-7
3.3.4 Menentukan Alat yang digunakan dalam Penelitian	III-10
3.3.5 Melakukan Pengujian Penelitian	III-10
3.3.6 Melakukan Analisis Hasil Pengujian dan Membuat Kesimpulan Penelitian	III-11
3.4 Metode Pengembangan Perangkat Lunak.....	III-11
3.4.1 Fase Insepsi	III-11
3.4.2 Fase Elaborasi	III-12
3.4.3 Fase Konstruksi	III-12

3.4.4 Fase Transisi.....	I-13
3.5 Rencana Jadwal Penelitian	III-13
3.6 Kesimpulan	III-15
BAB IV PENGEMBANGAN PERANGKAT LUNAK	IV-1
4.1 Pendahuluan	IV-1
4.2 Fase Insepsi	IV-1
4.2.1 Pemodelan Bisnis	IV-1
4.2.2 Kebutuhan Sistem.....	IV-2
4.2.2.1 Fitur <i>Pre-processing</i> Data.....	IV-4
4.2.2.2 Fitur Klasifikasi MNB.....	IV-4
4.2.2.3 Fitur Klasifikasi MNB + MI	IV-4
4.2.3 Analisis dan Desain	IV-4
4.2.3.1 Analisis Kebutuhan Perangkat Lunak	IV-5
4.2.3.2 Analisis Data.....	IV-5
4.2.3.3 Analisis <i>Pre-processing</i>	IV-6
4.2.3.4 Analisis Seleksi Fitur <i>Mutual Information</i>	IV-8
4.2.3.5 Analisis Pembobotan TF-IDF	IV-10
4.2.3.6 Analisis Klasifikasi <i>Multinomial Naive Bayes</i>	IV-13
4.2.3.7 Analisis <i>Confusion Matrix</i>	IV-19
4.2.3.6 Desain Perangkat Lunak.....	IV-21
4.3 Fase Elaborasi.....	IV-1
4.3 Fase Elaborasi.....	IV-28
4.3.1 Pemodelan Bisnis	IV-28
4.3.2 Perancangan Data	IV-29

4.3.3 Perancangan Antarmuka	V-29
4.3.4 Kebutuhan Sistem.....	IV-30
4.3.5 Diagram Aktivitas	IV-30
4.3.6 Diagram Alur	IV-32
4.4 Fase Kontraksi	IV-33
4.4.1 Kebutuhan	IV-33
4.4.2 Implentasi.....	IV-34
4.4.2.1 Implemntasi Kelas.....	IV-34
4.4.2.2 Implementasi Antarmuka	IV-35
4.5 Fase Transisi.....	IV-36
4.6 Kesimpulan	IV-39
BAB V HASIL DAN ANALISIS PENETIAN.....	V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Penelitian.....	V-1
5.2.1 Konfigurasi Percobaan	V-1
5.2.2 Data Hasil Konfigurasi I.....	V-1
5.2.2 Data Hasil Konfigurasi II	V-3
5.2.3 Data Hasil Konfiguasi III	V-5
5.3 Analisis Hasil Penelitian	V-5
5.3.1 Analisis Percobaan.....	V-5
5.3.2 Analisis Waktu Komputasi	V-6
5.3.3 Analisisi Jumlah Fitur	V-7
5.4 Kesimpulan	V-9
BAB IV KESIMPULAN.....	I-1

6.1	Pendahuluan	I-1
6.2	Kesimpulan	VI-1
6.3	Saran	VI-2
DAFTAR	PUSTAKA	IV-1

DAFTAR TABEL

Tabel II-1. Tabel kemunculan <i>term</i> pada <i>Mutual Infomation</i>	I-5
Tabel II-2. Tabel <i>K-Fold Cross Validation</i>	II-8
Tabel II-3. Tabel <i>Confusion Matrix</i>	II-9
Tabel III-1. Tabel <i>Confusion Matrix</i> Pada MNB.....	III-8
Tabel III-2. Tabel <i>Confusion Matrix</i> pada klasifikasi MNB+ MI	III-8
Tabel III-3. Rancangan Tabel Hasil Klasifikasi	III-9
Tabel III-4. Rancangan Tabel Hasil Analisis Klasifikasi.....	III-11
Tabel III-5. Rencana Jadwal Penelitian	III-14
Tabel IV-1. Kebutuhan Fungsional	IV-3
Tabel IV-2. Kebutuhan Non-Fungsional.....	IV-3
Tabel IV-3. Tabel Contoh Data Diskusi Terkait <i>Stress</i>	IV-6
Tabel IV-4. Hasil <i>Tokenizing</i>	IV-7
Tabel IV-5. Hasil <i>Flitering</i>	IV-8
Tabel IV-6. Tabel <i>Confusion Matrix Mutual Informtion</i>	IV-9
Tabel IV-7. Hasil <i>Mutual Information</i>	IV-9
Tabel IV-8. Hasil TF dan IDF dari Contoh Data Diskusi Terkait	IV-11
Tabel IV-9. Hasil Pembobotan Kata Dari Contoh Data Diskusi Terkait.....	IV-12
Tabel IV-10. Hasil Pembobotan Kata dari Hasil MI	IV-13
Tabel IV-11. Hasil Perhitungan Probabilitas Data Training MNB	IV-15
Tabel IV-12. Hasil Perhitungan Probabilitas Data Training MNB+MI	IV-16
Tabel IV-13. Data Testing.....	IV-16
Tabel IV-14. Tabel Confusion Matrix MI Data Testing	IV-17

Tabel IV-15. Tabel Hasil MI Data Testing	V-17
Tabel IV-16. Tabel Hasil Threshold MI Pada Data Testing	IV-18
Tabel IV-17. Tabel Klasifikasi MNB Tanpa MI	IV-19
Tabel IV-18. Tabel Klasifikasi MNB dengan MI.....	IV-19
Tabel IV-19. Tabel Confusion Matrix	IV-20
Tabel IV-20. Tabel Definisi Aktor	IV-22
Tabel IV-21. Tabel Definisi Use Case	IV-22
Tabel IV-22. Tabel Skenario Use Case Proses Data	IV-24
Tabel IV-23. Skenario Klasifikasi MNB	IV-25
Tabel IV-24. Skenario Klasifikasi MNB + MI.....	IV-27
Tabel IV-25. Tabel Implementasi Kelas dalam Python	IV-34
Tabel IV-26. Tabel Rencana Pengujian Use Case Proses Data	IV-37
Tabel IV-27. Rencana Pengujian Use Case Klasifikasi MNB	IV-37
Tabel IV-28. Rencana Pengujian Use Case MNB + MI.....	IV-37
Tabel IV-29. Pengujian Use Case Proses Data	IV-38
Tabel IV-30. Pengujian Use Case Klasifikasi MNB	IV-38
Tabel IV-31. Pengujian Use Case MNB + MI	IV-38
Tabel V-1. Hasil Evaluasi Klasifikasi <i>Multinomial Naive Bayes</i>	V-2
Tabel V-2. Hasil Evaluasi Klasifikasi <i>Multinomial Naive Bayes</i> + MI	V-4
Tabel V-3. Perbandingan Rata-Rata Hasil Klasifikasi MNB dan MNB+ MI ..	IV-5

DAFTAR GAMBAR

Gambar II-1. Tahapan yang dilakukan pada <i>pre-processing</i>	I-2
Gambar II-2. Proses <i>tokenizing</i>	II-3
Gambar II-3. Proses <i>filtering</i>	II-3
Gambar II-5. <i>Rational Unified Process</i>	II-10
Gambar III-1. Data yang akan digunakan	III-2
Gambar III-2. Diagram alur proses umum perangkat lunak	III-4
Gambar III-3. Proses pengambilan data.....	III-5
Gambar IV-1. Diagram Use Case	IV-21
Gambar IV-2. Rancangan Antarmuka Perangkat Lunak	IV-29
Gambar IV-3. Diagram Aktivitas Proses Data	IV-31
Gambar IV-4. Diagram Klasifikasi MNB	IV-31
Gambar IV-5. Diagram Klasifikasi MNB+MI	IV-31
Gambar IV-6. Diagram Alur Proses Data	IV-32
Gambar IV-7. Diagram Alur Klasifikasi MMNB	IV-33
Gambar IV-8. Diagram Alur Klasifikasi MNB + MI	IV-33
Gambar IV-9. Diagram Kelas Perangkat Lunak	IV-34
Gambar IV-10. Implementasi Antarmuka Perangkat Lunak	IV-35
Gambar V-1. Grafik Hasil Pengujian 10 <i>Fold</i> pada <i>Multinomial Nive Bayes</i> ...	V-3
Gambar V-2. Grafik Hasil Pengujian 10 <i>Fold</i> pada MNB + MI.....	V-4
Gambar V-3. Perbandingan Hasil Evaluasi.....	V-6
Gambar V-4. Diagram Perbandingan Waktu Komputasi Metode Klasifikasi	V-7
Gambar V-5. Perbandingan Jumlah Fitur yang Dihasilkan Setiap Metode	V-8

DAFTAR LAMPIRAN

Lampiran 1. Kode Program.....	L-1
-------------------------------	-----

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini memberikan penjelasan umum tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah yang menjadi dasar dalam penelitian ini.

1.2 Latar Belakang

Chatbot merupakan sebuah program yang dirancang untuk dapat berinteraksi dengan manusia melalui pesan atau suara. Chatbot membekali dirinya dengan *Artificial Intelligence* (kecerdasan buatan) dan pemrosesan bahasa alami atau *Natural Language Processing* (NLP) yang dapat membuat chatbot menjadi cerdas dan dapat menjawab pertanyaan yang diberikan oleh pengguna (Eka Yuniar, 2019). Chatbot terapi merupakan chatbot yang dapat memberikan informasi dan berharap bisa membantu pengguna merasa lebih baik mengenai kesehatan, medis, kecemasan, depresi dan stres. Pengguna dapat mengungkapkan isi hati atau tekanan yang ada didalam dirinya.

Secara umum komponen penyusun chatbot terdiri dari bot program dan *brain file*. Bot program adalah program utama pada chatbot yang bertujuan untuk mengakses masukan dari pengguna dan selanjutnya akan melakukan parsing. Bot program terdiri dari komponen scanner dan parsing. *Brain file* merupakan otak

dari sebuah chatbot yang akan berfikir dan memberikan respon. Semua kosakata, kepribadian dan pengetahuan dari chatbot disimpan di dalam *brain file* (Baiti dan Nugroho, 2013). Chatbot akan mengidentifikasi *user*, *agent* dan *intent* (Oktavia, 2020). User akan mengirim pesan ke *agent*, lalu *agent* mencari ekspresi kata yang sesuai. Dalam *agent* terdiri dari *intent* dan entitas, *intent* merupakan proses dimana chatbot akan memahami dan mengenali isi pernyataan yang dimasukan oleh pengguna kepada chatbot. Sedangkan entitas merupakan konteks spesifik dari suatu tersebut. Setelah menemukan *intent* yang sesuai maka respon akan diberikan. Masalah yang sering muncul dalam chatbot yaitu terkadang chatbot tidak mengerti maksud dari pengguna, yang membuat chatbot memberikan balasan interaksi dialog yang tidak sesuai. Maka klasifikasi *intent* merupakan salah satu cara untuk mengkatagorikan suatu *intent*.

Metode yang dapat digunakan dalam klasifikasi yaitu *Naive Bayes*, *K-Nearest Neighbor*, dan *Support Vector Machine* dan masih banyak lagi. *Support Vector Machine* memiliki kelebihan yakni mampu mengidentifikasi hyperplane atau batas keputusan terpisah yang memaksimalkan margin antar dua kelas yang berbeda sedangkan kekurangan yang dimiliki oleh *Support Vector Machine* yaitu jika dalam pemilihan fitur dan parameter tidak sesuai maka akan dapat mengurangi akurasi (Yunita, 2016). *K-Nearest Neighbor* mempunyai kelebihan yaitu pelatihan yang sangat cepat, sederhana, mudah dipelajari, kebal terhadap noise dan efektif dalam data pelatihan yang sangat besar. Sedangkan kekurangan dari *K-Nearest Neighbor* yaitu keterbatasan memori, mudah terperdaya dengan atribut yang tidak relevan (Mutrofin et al., 2014). Kelebihan yang dimiliki oleh

Naive Bayes yaitu kecepatan dan kesederhaan dalam pengklasifikasian dan memiliki kinerja yang baik dalam klasifikasi yang mengandung angka maupun teks (Pandhu dan Agus, 2016), sedangkan kekurangan dari *Naive Bayes* yaitu kadang ada celah untuk mengurangi keefektifan dari suatu metode yang berakibat dokumen melenceng dari kelas yang sebenarnya (Natalius, 2010). Pada penelitian yang dilakukan (Pandhu dan Agus, 2016) yang membandingkan kinerja metode *Naive Bayes* dan *K-Nearest Neighbor* untuk klasifikasi artikel berbahasa Indonesia, bahwa kinerja *Naive Bayes* lebih unggul dibandingkan *K-Nearest Neighbor*.

Metode *Naive Bayes* bisa dipakai di berbagai bidang, khususnya pada klasifikasi dokumen atau teks (Isnaini, 2017). *Multinomial Naive Bayes* merupakan variasi dari *Naive Bayes* (Bunga, Bertha, dan Yelly 2018). *Multinomial Naive Bayes* adalah salah satu metode *bayes* yang akan memperhitungkan frekuensi kemunculan kata dalam suatu dokumen dan probabilitas (Diandra, Firdaus, dan Kurniati 2010). Metode *Multinomial Naive Bayes Classifier* adalah mudah diimplemetasi, tingkat akurasi yang tinggi dan dapat mengatasi jumlah kosakata yang besar (Vidya, Maharani, dan Yulita 2012).

Mutual Information merupakan metode seleksi fitur yang cukup baik dalam pemilihan fitur pada suatu dokumen, sehingga dapat berpengaruh dalam proses pengklasifikasian agar bekerja secara maksimal (Nurfikri dan Mubarak, 2018). Penelitian yang dilakukan oleh (Irham, Adiwijaya, dan Wisesty 2019) *Mutual Information* membantu dalam menyeleksi fitur, akurasi terbaik diperoleh

ketika menggunakan *Mutual Information* sebesar 93,81% dengan waktu komputasi yang sedikit. Tetapi penelitian yang dilakukan oleh (Karunia, 2017) bahwa hasil yang dilakukan oleh *mutual information* mengalami penurunan karena adanya fitur yang terbuang dari proses seleksi fitur *mutual information*. Berdasarkan latar belakang diatas maka dibangun sistem dengan menggunakan *Multinomial Naive Bayes* dan *Mutual Information*

1.3 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka dibuat rumusan masalah dalam penelitian ini adalah

1. Bagaimana cara melakukan klasifikasi *intent* pada chatbot terapi menggunakan *Multinomial Naive Bayes* dan *Mutual Information*?
2. Berapa hasil presentase tingkat akurasi yang dihasilkan klasifikasi *intent* dengan menggunakan *Multinomial Naive Bayes* dan *Mutual Information*?

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Membuat perangkat lunak klasifikasi *intent* pada chatbot terapi menggunakan *Multinomial Naive Bayes* dan *Mutual Information*.
2. Mengetahui tingkat akurasi dari hasil klasifikasi *intent* menggunakan *Multinomial Naive Bayes* dan *Mutual Information*.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah :

1. Dapat mengetahui informasi mengenai hasil klasifikasi *intent* menggunakan algoritma *Multinomial Naive Bayes* dan *Mutual Information*.
2. Menjadi referensi atau rujukan untuk penelitian berikutnya yang lebih relevan.

1.6 Batasan Masalah

Untuk membatasi penelitian supaya tidak keluar dari topik bahasan, maka diberikan beberapa batasan masalah penelitian sebagai berikut :

1. Data yang digunakan adalah data yang diambil langsung dari diskusi terkait Alodokter.com mengenai stres.
2. Tidak dapat membaca kalimat yang berisikan emoticon atau emoji.
3. Data diambil sebanyak 300 berbahasa Indonesia.
4. Terdiri dari 20 intent yaitu sakit kepala, nafsu makan, tidur tidak nyenyak, merasa takut, merasa (cemas, tegang, khawatir), tangan gemetar, gangguan pencernaan, sulit berfikir, merasa tidak bahagia, sering menangis, sulit untuk menikmati kehidupan sehari-hari, mengalami kesulitan untuk mengambil keputusan, tugas/aktivitas sehari-hari terbengkalai, merasa tidak mampu berperan dalam kehidupan, kehilangan minat, merasa tidak berharga, pikiran untuk mengakhiri hidup, merasa lelah sepanjang waktu, merasa tidak enak di perut dan mudah merasa lelah.
5. Tidak berlaku probabilitas kondisionalnya nol, apabila nol maka probabilitasnya akan bernilai nol juga.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya, antara lain:

BAB I. PENDAHULUAN

Pada bab ini diuraikan mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah/ruang lingkup, dan sistematika penulisan.

BAB II. KAJIAN LITERATUR

Bab ini membahas seluruh dasar-dasar teori yang digunakan mulai dari definisi sistem, informasi mengenai domain, dan semua yang digunakan pada tahapan analisis, perancangan, dan implementasi.

BAB III. METODE PENELITIAN

Pada bab ini akan dibahas mengenai tahapan yang akan dilaksanakan pada penelitian ini. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci dengan mengacu pada suatu kerangka kerja. Di akhir bab ini berisi perancangan manajemen proyek pada pelaksanaan penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini membahas mengenai analisis dan perancangan perangkat lunak yang akan digunakan sebagai alat penelitian. Dimulai dari pengumpulan dan analisa kebutuhan, rancangan dan konstruksi perangkat lunak serta pengujian untuk memastikan semua kebutuhan pengembangan perangkat lunak sesuai dengan kebutuhan.

BAB V. HASIL DAN ANALISA PENELITIAN

Pada bab ini diuraikan hasil pengujian berdasarkan langkah-langkah yang telah direncanakan. Tabel hasil pengujian serta analisisnya disajikan sebagai basis dari kesimpulan yang akan diambil dalam penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga saran yang diberikan berdasarkan hasil dari penelitian.

1.8 Kesimpulan

Dengan demikian berdasarkan masalah yang telah diuraikan diatas akan dilakukan penelitian terhadap data pertanyaan dan kalimat yang akan digunakan. Kemudian melakukan implementasi dengan *Natural Language Processing* menggunakan *Multinomial Naive Bayes* dan seleksi fitur *Mutual Information*.

DAFTAR PUSTAKA

- Baiti, Z. N., & Nugroho, F. (2013). Aplikasi Chatbot “Mi3” Untuk Informasi Jurusan Teknik Informatika Berbasis Sistem Pakar Menggunakan Metode Forward Chaining. *Matics*. <https://doi.org/10.18860/mat.v0i0.2425>
- Boutell, C. (1865). Milton. *Notes and Queries*, s3-VII(182), 504. <https://doi.org/10.1093/nq/s3-VII.182.504a>
- Bunga, M. T. H., Djahi, B. S., Kom, S., Nabuasa, Y. Y., Cs, M., Komputer, J. I., Sains, F., Teknik, D., & Cendana, U. N. (2018). Multinomial Naive Bayes Untuk Klasifikasi Status Kredit Mitra Binaan Di Pt. Angkasa Pura I Program Kemitraan. *J-Icon*, 6(2), 30–34.
- Eka Yuniar, H. P. (2019). *IMPLEMENTASI CHATBOT “ALITTA” ASISTEN VIRTUAL DARI BALITTAS SEBAGAI PUSAT INFORMASI DI BALITTAS*. 12(1).
- Ferlin, J. (2019). *Klasifikasi Customer Intent Melalui Text Mining Untuk Mengetahui Tingkat Kepuasan Pelanggan Berdasarkan Komentar Pelanggan Menggunakan Metode Support Vector Machine Pada Restoran Bakso President*. 3(10), 9867–9875.
- Irham, L. G., Adiwijaya, A., & Wisesty, U. N. (2019). Klasifikasi Berita Bahasa Indonesia Menggunakan Mutual Information dan Support Vector Machine. *Jurnal Media Informatika Budidarma*, 3(4), 284. <https://doi.org/10.30865/mib.v3i4.1410>

- Isnaini, Z. (2017). *Vol. 16, No. 1, Juni 2017. 16(1), 1–13.*
- Karunia, S. A. (2017). *Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection. 6(1).*
- Maskur, M. (2016). Perancangan Chatbot Pusat Informasi Mahasiswa Menggunakan Aimi Sebagai Virtual Assistant Berbasis Web. *Kinetik, 1(3), 123.* <https://doi.org/10.22219/kinetik.v1i3.47>
- Natalius, S. (2010). *Makalah II2092 Probabilitas dan Statistik-Sem. I Tahun. 3.*
- Nurfikri, F. S., & Mubarak, M. S. (2018). *Klasifikasi Topik Berita Menggunakan. 5(1), 1579–1588.*
- Oktavia, C. A. (2020). Implementasi Chatbot Menggunakan Dialogflow dan Messenger Untuk Layanan Customer Service Pada E-Commerce. *J I M P - Jurnal Informatika Merdeka Pasuruan, 4(3), 36–40.* <https://doi.org/10.37438/jimp.v4i3.230>
- Pakpahan, D., Widyastuti, H., Politeknik,), & Batam, N. (2014). Aplikasi Opinion Mining dengan Algoritma Naïve Bayes untuk Menilai Berita Online. *Jurnal Integrasi, 6(1), 1–10.*
- Pandhu, A., & Agus, H. (2016). Naive Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government. *Journal of Applied Intelligent System, 1(1), 48-55–55.*
- Photovoltaic, I. (2018). *Hitta denna un uni in un at kaiutti uniti. 1.*
- Rahman, A. (2017). Online News Classification Using Multinomial Naive Bayes.

Itsmart, 6(1), 32–38. <https://doi.org/10.1177/1096348015584441>

Sukadiyanto. (2010). Stress dan Cara mengatasinya. *Cakrawala Pendidikan*, 29(1), 55–66.

Widiastuti, L. (2018). Pemilihan Fitur Pada Analisis Sentimen Review Travel Online Menggunakan Algoritma Naïve Bayes Dalam Penerapan Mutual Information Dan Particle Swarm Optimization (PSO). *IJCIT (Indonesian Journal on Computer and Information Technology)*, 3(1), 10.

Yunita, N. (2016). Analisis Sentimen Berita Artis Dengan Menggunakan Algoritma Support Vector Machine dan Particle Swarm Optimization. *Jurnal Sistem Informasi STMIK Antar Bangsa*, 5(2), 104–112.