

# SENASTIKA Universitas Malikussaleh

---

## DETEKSI RISIKO DEPRESI DAN KECEMASAN MAHASISWA MENGGUNAKAN ALGORITMA NAIVE BAYES

Putri Agustina Dewi<sup>1</sup>, Rizka Aulia<sup>2</sup>, Murniaty<sup>3</sup>, Anwa'il Khairi<sup>4</sup>, Munirul Ula<sup>5</sup>

Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh

Email : <sup>1</sup>[putri.210170094@mhs.unimal.ac.id](mailto:putri.210170094@mhs.unimal.ac.id), <sup>2</sup>[rizka.210170095@mhs.unimal.ac.id](mailto:rizka.210170095@mhs.unimal.ac.id),  
<sup>3</sup>[murniaty.210170093@mhs.unimal.ac.id](mailto:murniaty.210170093@mhs.unimal.ac.id), <sup>4</sup>[anwail.210170092@mhs.unimal.ac.id](mailto:anwail.210170092@mhs.unimal.ac.id), <sup>5</sup>[munirulula@unimal.ac.id](mailto:munirulula@unimal.ac.id)

### Abstrak

Kesehatan mental mahasiswa merupakan isu yang semakin mendesak di kalangan institusi pendidikan tinggi, mengingat meningkatnya tingkat stres dan tekanan akademik yang dialami oleh mahasiswa. Penelitian ini bertujuan untuk menganalisis hasil survei kesehatan mental mahasiswa dengan menggunakan algoritma Naive Bayes untuk mendeteksi risiko depresi dan kecemasan. data diambil dari Kaggle yang mengacu pada penelitian menggunakan data numerik atau data yang dapat diukur. Setelah melakukan preprocessing, data dibagi menjadi data latih dan data uji. Model Naive Bayes dilatih dan dievaluasi menggunakan data tersebut. Hasil analisis menunjukkan bahwa model Naive Bayes dapat dengan efektif mengklasifikasikan mahasiswa ke dalam kategori risiko rendah, sedang, dan tinggi terkait depresi dan kecemasan, dengan akurasi yang signifikan. Temuan ini memberikan wawasan penting bagi institusi pendidikan dalam merancang program intervensi yang lebih baik untuk mendukung kesehatan mental mahasiswa, serta meningkatkan kesadaran akan pentingnya perhatian terhadap kesehatan mental di lingkungan kampus.

Kata Kunci : *Depresi, Klasifikasi, Kesehatan mental, Kecemasan, Mahasiswa, Naive Bayes*

---

## 1. PENDAHULUAN

Perkembangan Teknologi Informasi (IT) sekarang ini sangat pesat dan telah digunakan diberbagai aspek kehidupan baik di bidang pemerintahan, perbankan, sosial budaya, industri, pendidikan, bahkan kesehatan.<sup>1</sup> Kesehatan mental merupakan aspek penting dari kesejahteraan individu yang kerap diabaikan, terutama di lingkungan akademis. Mahasiswa, sebagai kelompok yang mengalami tekanan akademik, sosial, dan psikologis yang cukup tinggi, sering kali menghadapi masalah kesehatan mental, seperti depresi dan kecemasan. Sebuah laporan dari World Health Organization (WHO) menyatakan bahwa sekitar 20% populasi remaja dan dewasa muda di seluruh dunia mengalami masalah kesehatan mental, dan angka ini terus meningkat di kalangan mahasiswa. Kondisi ini mengarah pada penurunan kualitas hidup, kinerja akademik yang buruk, serta meningkatnya risiko putus kuliah. Penelitian terkait kesehatan mental mahasiswa di Indonesia semakin berkembang seiring dengan meningkatnya perhatian terhadap dampak kesehatan mental pada kehidupan akademik dan sosial mahasiswa. Namun, tantangan utama dalam menangani masalah kesehatan mental adalah keterbatasan dalam mendeteksi gejala-gejala awal, seperti depresi dan kecemasan, yang sering kali tidak dilaporkan. Oleh karena itu, diperlukan upaya sistematis untuk mengidentifikasi mahasiswa yang berisiko mengalami gangguan kesehatan mental melalui metode yang efektif dan efisien. Usia remaja sering kali lebih rentan terkena masalah gangguan kecemasan dikarenakan adanya fenomena dunia modern yang menjadi pemicu peningkatan kasus tersebut, dan kini dengan adanya media sosial rasa cemas tersebut semakin menjadi jadi karena prasangka yang muncul dari diri sendiri, seperti merasa sendirian, terkucilkan, dan bahkan dipermalukan.<sup>2</sup>

Dalam upaya untuk mengatasi masalah ini, teknologi komputasi modern, seperti pembelajaran mesin (machine learning), menawarkan pendekatan yang menjanjikan. Salah satu algoritma yang dapat digunakan untuk klasifikasi risiko kesehatan mental adalah **Naive Bayes**, yang dikenal memiliki kecepatan dan akurasi tinggi dalam mengolah data survei. Dengan menerapkan algoritma ini, analisis survei kesehatan mental dapat digunakan untuk memprediksi risiko depresi dan kecemasan pada mahasiswa. Pendekatan ini diharapkan mampu memberikan hasil yang signifikan dalam mengklasifikasikan tingkat risiko, sehingga intervensi dini dapat dilakukan secara lebih efektif. Salah satu contoh penelitian yang telah terdahulu dilakukan oleh Mayy M. Al-Tahrawi (2015) melakukan penelitian tentang kategorisasi teks bahasa Arab menggunakan regresi Support Vector Machine, Naive Bayes, dan Sistem Informasi Geografis. Hasilnya, Support Vector Machines mengungguli algoritma statistik lainnya, seperti GIS, Naive Bayes, dan regresi logistik, dalam hal akurasi.<sup>3</sup> Penelitian ini bertujuan untuk menganalisis hasil survei kesehatan mental mahasiswa dengan menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat

risiko depresi dan kecemasan. Dengan pendekatan ini, diharapkan dapat memberikan kontribusi terhadap upaya institusi pendidikan dalam mendeteksi dan menangani masalah kesehatan mental di kalangan mahasiswa secara lebih tepat sasaran.

## 2. METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif karena data diambil dari Kaggle yang mengacu pada penelitian menggunakan data numerik atau data yang dapat diukur seperti hasil survei kesehatan mental mahasiswa akan berupa data angka, misal seperti skor depresi, tingkat kecemasan, atau solusi aktivitas apa yang harus dilakukan untuk menghilangkan stres. Data ini kemudian dianalisis menggunakan algoritma Naive Bayes, yang merupakan metode statistik untuk memprediksi klasifikasi berdasarkan data numerik. Algoritma naive bayes mengklasifikasi risiko depresi dan kecemasan pada mahasiswa berdasarkan hasil survei Kesehatan mental. Secara sistematis, naïve bayes didasarkan pada Teorema Bayes:

$$p(c|x) = \frac{p(x|c) \cdot p(c)}{p(x)}$$

Pada akhirnya, penelitian ini berfokus pada kuantifikasi hasil berapa persen mahasiswa yang berisiko depresi atau kecemasan dan memprediksi berdasarkan data serta memberikan solusi aktivitas apa yang harus dilakukan untuk menghilangkan stres sesuai dengan kategori risiko rendah, sedang, dan tinggi. Metode penelitian ini juga terdiri dari sejumlah tahap yang akan diselesaikan untuk mempermudah mendeteksi risiko depresi dan kecemasan.

### 2.1 Pengumpulan Data

Pengumpulan data adalah Teknik yang digunakan untuk mengumpulkan berbagai data dan informasi yang nantinya akan berguna sebagai bukti pendukung untuk penjelasan studi penelitian. Pada penelitian ini data diambil dari Kaggle ([Link Data](#)) yaitu platform online yang menyediakan berbagai layanan untuk ilmuwan data dan praktisi pembelajaran mesin.

### 2.2 Subyek dan Obyek Penelitian

Subyek penelitian ini diambil dari kumpulan data yang berisi respons survei dari mahasiswa TI, yang berfokus pada stres akademis, kesehatan mental, dan faktor gaya hidup. Data ini diambil dari Kaggle milik Avinash Bunga yang berjudul “Survey Kesehatan Mental Mahasiswa”. Adapun obyek penelitiannya itu ada mencakup beberapa variabel penting, seperti:

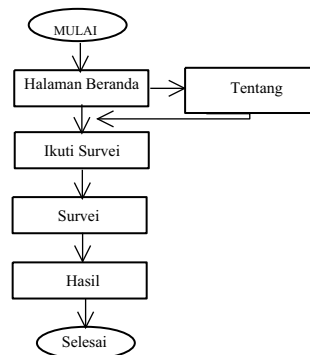
- Kepuasan studi
- Beban kerja akademis
- Tekanan akademis
- Masalah finansial
- Hubungan sosial
- Depresi
- Kecemasan
- Isolasi
- Ketidakamanan masa depan

### 2.3 Teknik Analisis Masalah

Seperti yang kita ketahui banyak sekali mahasiswa yang mengalami masalah kesehatan mental diakibatkan kombinasi tekanan akademis, masalah finansial, perasaan terisolasi, dan perubahan gaya hidup. Beban tugas yang berat, ujian, tuntutan prestasi, serta ketidakpastian karir sering memicu stress dan kecemasan, jauh dari keluarga dan sulitnya membangun hubungan social juga membuat mereka merasa kesepian, pola hidup yang tidak sehat dan kurangnya dukungan mental memperburuk kondisi, sementara tekanan social dari lingkungan dan media sosial menambah beban emosional.

### 2.4 Desain

Desain adalah tahap dimana gambaran lengkap tentang apa yang perlu dilakukan dan bagaimana sistem tersebut bekerja. Disini kami hanya menggunakan flowchart dalam tahap desain ini. Adapun flowchart adalah sekelompok symbol yang digunakan dalam diagram notasi untuk mewakili urutan proses dan aliran data dalam suatu system.



## 2.5 Normalisasi

Pada survei kesehatan mental mahasiswa, terdapat 9 variabel yang diukur dengan skala Likert (1-5). Jika terdapat perbedaan skala antar variabel atau perbedaan yang signifikan dalam rentang nilai, model klasifikasi seperti Naive Bayes bisa memberikan bobot yang lebih besar pada variabel dengan skala yang lebih besar, meskipun tidak seharusnya demikian. Misalnya seperti kepuasan Studi dan Tekanan Akademis sama-sama diukur dalam skala 1-5, tetapi jika ada variabel lain dengan skala berbeda, model mungkin tidak memperlakukannya secara adil.

Normalisasi bertujuan untuk menyamakan rentang nilai antar variabel sehingga semua fitur memiliki skala yang serupa dan berpengaruh secara seimbang dalam proses klasifikasi. Ini membuat algoritma Naive Bayes dapat bekerja dengan lebih baik, khususnya ketika perbedaan skala antar variabel bisa mempengaruhi performa model.

Adapun jenis normalisasi yang diterapkan adalah Min-Max:

- Setiap nilai dalam fitur diubah kedalam rentang 0 hingga 1.
- Rumus :

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

- Misalnya jika nilai tekanan akademis di skala 1-5, hasil normalisasinya akan tetap dalam rentang 0 hingga 1.

## 2.6 Klasifikasi

Klasifikasi adalah yang mengacu pada metode yang dikenal sebagai pembelajaran yang diawasi yaitu membutuhkan penentuan jarak antara input dan atribut target. Klasifikasi pada penelitian ini bertujuan untuk mengelompokkan mahasiswa berdasarkan hasil survey mereka menjadi kategori berisiko atau tidak berisiko terhadap depresi dan kecemasan. Model klasifikasi Naive Bayes melibatkan beberapa tahap penyelesaian, dimulai dari pelatihan data (training) dan diakhiri dengan pengujian (testing), dengan tujuan menghasilkan keputusan yang akurat.<sup>4</sup>

## 2.7 Algoritma Naive Baiyes

Naive Bayes adalah algoritma yang digunakan untuk klasifikasi data, khususnya dalam menentukan apakah seorang mahasiswa berisiko mengalami depresi atau kecemasan berdasarkan hasil survei kesehatan mental. Algoritma ini bekerja dengan menghitung probabilitas dari setiap gejala atau faktor (seperti tekanan akademis, hubungan sosial, isolasi) dan kemudian memprediksi kemungkinan mahasiswa mengalami masalah mental. Naive Bayes memiliki keunggulan yaitu Sederhana dan cepat dalam memproses data serta akurat untuk data yang memiliki banyak variabel atau faktor. Sedangkan kelemahannya ialah mengasumsikan bahwa setiap faktor (seperti isolasi, tekanan, dll.) tidak saling memengaruhi, meskipun dalam kenyataannya mungkin saja saling berkaitan.

$$p(c|x) = \frac{p(x|c) \cdot p(c)}{p(x)}$$

Keterangannya:

- $P(C|X)$  : Probabilitas bahwa kelas **C** (misalnya berisiko atau tidak berisiko depresi/kecemasan) diberikan data fitur **X** (hasil survei mahasiswa).
- $P(X|C)$  : Probabilitas munculnya data **X** (misalnya kepuasan studi, tekanan akademis, dll.) untuk kelas **C**.
- $P(C)$  : Probabilitas awal dari kelas **C** (probabilitas depresi/kecemasan terjadi pada populasi)
- $P(X)$  : Probabilitas keseluruhan dari data fitur **X** tanpa mempertimbangkan kelas.

### 2.8 Implementasi

Website ini dirancang untuk membantu mahasiswa melakukan survey mengenai kesehatan mental mereka dengan menggunakan algoritma Naive Bayes untuk menganalisis hasilnya. Di halaman utama, mahasiswa dapat menemukan informasi tentang pentingnya Kesehatan mental dan akses ke formulir survey. Setelah mengisi survey yang mencakup berbagai pertanyaan tentang kepuasan studi, beban kerja, dan faktor lain yang mempengaruhi kesehatan mental, data mereka akan diproses oleh system. Hasil analisis akan menampilkan apakah mereka berada pada risiko tinggi atau rendah terhadap depresi dan kecemasan. Dengan pendekatan ini, mahasiswa dapat lebih memahami kondisi mental mereka dan mencari bantuan jika diperlukan. Website ini juga bisa dikembangkan lebih lanjut untuk memberikan fitur analisis lebih mendalam dan dukungan bagi mahasiswa yang membutuhkannya.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Hasil Model Naive Bayes

Model naive bayes dibuat dengan menggunakan bantuan library python yaitu sklearn untuk mengetahui performanya. Pembuatan model naive bayes dilakukan dengan menggunakan 88 data, dari data yang digunakan 20% untuk pengujian, sementara 80% sisanya digunakan untuk melatih model. Pembelajaran mesin yang diawasi biasanya dimulai dengan serangkaian fitur prediktor yang bobotnya disetel ke nilai arbitrer yang setara. Dalam setiap evaluasi berikutnya, bobot yang terkait dengan fitur tertentu bergeser hingga tingkat tertentu. Hal ini berlanjut selama beberapa iterasi pada kumpulan data pelatihan yang sama hingga bobot telah dioptimalkan sepenuhnya, yang mencerminkan pengaruh agregat dari fitur yang diwakilinya.<sup>5</sup>

	precision	recall	f1-score	support
Creative Outlets	0.00	0.00	0.00	1
Nothing	0.00	0.00	0.00	1
Online Entertainment	0.25	1.00	0.40	1
Online Entertainment, Sleep	0.00	0.00	0.00	1
Outdoor Activities	0.00	0.00	0.00	1
Religious Activities	0.00	0.00	0.00	0
Religious Activities, Creative Outlets, Social Connections	0.00	0.00	0.00	1
Religious Activities, Online Entertainment	0.00	0.00	0.00	1
Religious Activities, Online Entertainment, Outdoor Activities	0.00	0.00	0.00	1
Religious Activities, Sleep	0.00	0.00	0.00	1
Religious Activities, Social Connections, Online Entertainment	0.00	0.00	0.00	1
Religious Activities, Social Connections, Online Entertainment, Sleep	0.00	0.00	0.00	1
Religious Activities, Sports and Fitness, Creative Outlets, Online Entertainment	0.00	0.00	0.00	1
Social Connections, Online Entertainment, Outdoor Activities, Sleep	0.00	0.00	0.00	1
Sleep	0.00	0.00	0.00	1
Social Connections	0.00	0.00	0.00	1
Social Connections, Online Entertainment, Sleep	0.00	0.00	0.00	1
Social Connections, Sleep	0.00	0.00	0.00	2
accuracy			0.06	18
macro avg	0.01	0.06	0.02	18
weighted avg	0.01	0.06	0.02	18

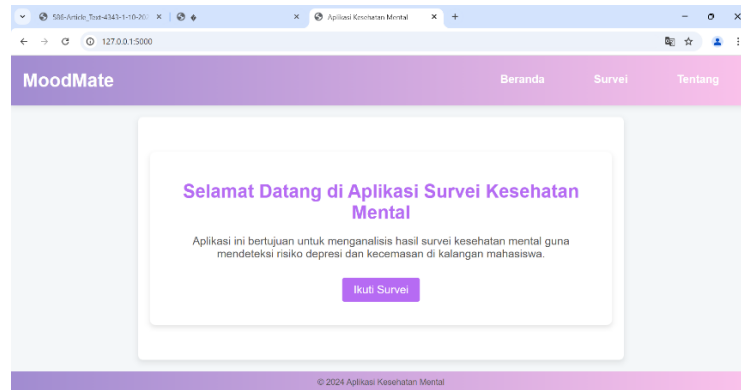
Gambar 1 Hasil Model Naive Bayes

### 3.2 Implementasi Desain Website

Adapun hasil dari tampilan desain website yang telah dibuat sebagai berikut:

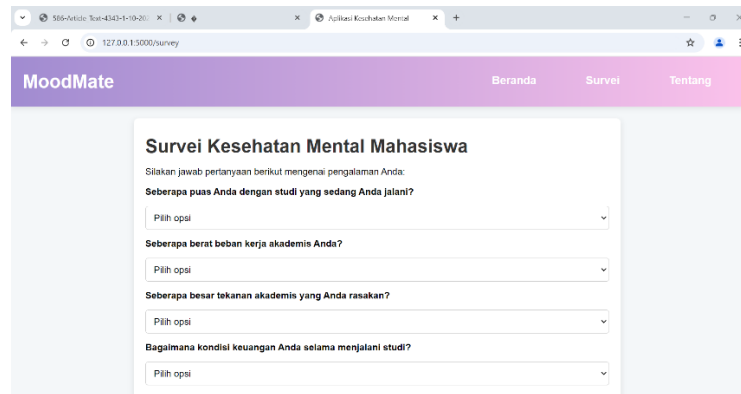
#### 1. Halaman Beranda

Halaman beranda adalah halaman utama dari sebuah situs web yang berfungsi sebagai pintu gerbang bagi pengunjung.



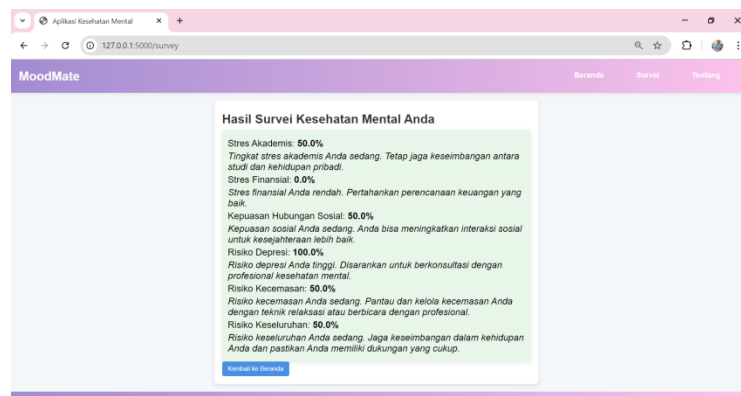
## 2. Halaman Survei

Halaman survei adalah bagian dari situs web yang dirancang untuk mengumpulkan data an informasi dari responden melalui serangkaian pertanyaan.



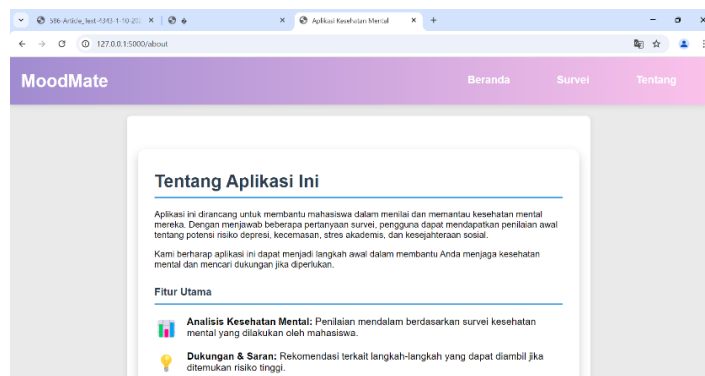
## 3. Halaman Hasil Survei

Halaman hasil survei yaitu bagian dari situs web yang menyajikan analisis dan informasi yang diperoleh dari data survei yang telah dilakukan. Dari halaman hasil survei ini juga dapat diketahui pengguna mendapatkan risiko sedang, rendah atau tinggi dan juga terdapat solusi cara mengatasinya.



## 4. Halaman Tentang

Halaman tentang adalah bagian dari situs web yang memberikan informasi mengenai tujuan dari situs tersebut.



#### 4. DISKUSI

Penelitian ini menggunakan algoritma Naïve Bayes untuk mengklasifikasikan hasil survei Kesehatan mental mahasiswa dengan tujuan mendeteksi risiko depresi dan kecemasan. Data yang digunakan mencakup variabel seperti tekanan akademis, masalah finansial, hubungan social, dan tingkat isolasi. Naïve bayes dipilih karena kesederhanaan dan efisiennya, meskipun memiliki keterbatasan dan asumsi independensi antar fitur. Hasil penelitian menunjukkan bahwa model ini mampu memberikan prediksi dengan tingkat akurasi, precision, recall, dan F1-score yang layak. Namun, tantangan seperti ketidakseimbangan kelas dan hubungan antar fitur menjadi pusat perhatian yang perlu ditangani untuk meningkatkan kinerja model. Deteksi dini yang dihasilkan dari analisis ini dapat membantu universitas mengambil Tindakan preventif dalam mendukung Kesehatan mental mahasiswa.

#### 5. KESIMPULAN

Kesimpulan dari analisis klasifikasi hasil survei Kesehatan mental mahasiswa menggunakan algoritma naïve bayes adalah bahwa metode ini mampu memberikan prediksi yang cukup efektif dalam mendeteksi risiko depresi dan kecemasan berdasarkan factor-faktor seperti tekanan akademis, hubungan social, dan kondisi emosional lainnya. Naïve bayes dengan pendekatan probabilitiknya yang sederhana namun kuat dapat mengolah data survei untuk menentukan kategori rendah, sedang, atau tinggi dengan memperhitungkan semua fitur secara independent. Dengan hasil prediksi yang akurat, algoritma ini dapat digunakan sebagai alat pendukung dalam memantau Kesehatan mental mahasiswa dan memberikan rekomendasi intervensi yang lebih cepat dan tepat sasaran. Namun, seperti pada setiap model prediktif, performanya tergantung pada kualitas dan keberagaman data yang digunakan serta asumsi independensi antar fitur yang mungkin tidak selalu berlaku di setiap kasus.

#### UCAPAN TERIMA KASIH

Tim penulis mengucapkan terima kasih yang mendalam kepada Bapak Munirul Ula, S.T., M.Eng., Ph.D selaku dosen pengampu mata kuliah Riset Teknologi Informasi dan memberikan dukungan dan bimbingan selama proses penelitian ini.

#### DAFTAR PUSTAKA

- [1] Anjarsari, T., Astutik, I. R. I., & Indahyanti, U. (2022). Deteksi Dini Gangguan Kecemasan Menggunakan Metode Naive Bayes. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 7(4), 1198–1210. <https://doi.org/10.29100/jupi.v7i4.3197>
- [2] Dwi Putra, H., Khairani, L., Hastari, D., Studi Sistem Informasi, P., Sains dan Teknologi, F., & Author, C. (2023). SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Comparison of Naive Bayes Classifier and Support Vector Machine Algorithms for Classifying Student Mental Health Data Perbandingan Algoritma Naive Bayes Classifier dan Support Vector Machine . *SENTIMAS: Seminar Nasional Penelitian Dan Pengabdian Masyarakat*, 120–125. <https://journal.irpi.or.id/index.php/sentimas>
- [3] Krieg, A., Okamura, K. H., & Higa-McMillan, C. K. (2019). *Identifying Situational, Individual, and Demographic Determinants of Social Anxiety Using a Naive Bayesian Classification Algorithm*. 81(080), 41.
- [4] Rayuwati, Husna Gemasih, & Irma Nizar. (2022). IMPLEMENTASI AIGORITMA NAIVE BAYES

UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID. *Jurnal Riset Rumpun Ilmu Teknik*, 1(1), 38–46. <https://doi.org/10.55606/jurritek.v1i1.127>

- [5] Siska, S., Saputra, G. A., Rohmat, C. L., & Sidik, F. (2023). Implementasi Metode Naive Bayes pada Prediksi Penyakit Seliak. *KOPERTIP Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 7(1), 8–13. <https://doi.org/10.32485/kopertip.v7i1.325>